

**АЛФА УНИВЕРЗИТЕТ
ФАКУЛТЕТ ИНФОРМАЦИОНИХ ТЕХНОЛОГИЈА**



Алфа БК Универзитет

**Модел заснован на модификованом
DBSCAN алгоритму за идентификацију
и анализу “црних тачака” на путевима
Републике Србије
- Докторска дисертација -**

кандидат

Јордан М. Атанасијевић

ментор

проф. др Војкан Николић

Београд, 2026. године

**ALFA BK UNIVERSITY
FACULTY OF INFORMATION TECHNOLOGIES**



Алфа БК Универзитет

**A model based on a modified DBSCAN
algorithm for identifying and analyzing
“black spots” on the roads of the Republic
of Serbia**

- Doctoral Dissertation -

candidate

Jordan M. Atanasijević

mentor

prof. Dr Vojkan Nikolić

Belgrade, 2026.

Захвалница

“Несебична помоћ потенцијалном квалитету да се докаже, вреди више од свих могућих личних узлета”, крилатица је у коју може стати захвалност коју су ми у току школовања и студирања пружили наставници, професори, колеге и пријатељи.

Својим саветима, сугестијама, предлозима и подршком учинили су да ова дисертације добије садржај и облик који има и да у сваком смислу представља круну мог школовања, па им се овом приликом срдечно захваљујем.

Дисертацију посвећујем својој ћерки Дарији, са љубављу и захвалношћу.

Сажетак

Модерни начин живота и рада и темпо пословања готово да подразумева присуство информационих технологија у свим сферама наших живота. Развој технологије унапређује наш живот и стандард, иако са друге стране генерише огромну количину података. Живимо у времену када, нажалост, нисмо у могућности да искористимо потенцијал свих тих података, а интуитивно закључујемо да се у тим подацима крију одређене вредности од којих сви можемо да имамо само користи.

Циљ примене информационих технологија је да се подаци који се масовно прикупљају обраде на такав начин да генеришу нову вредност, било за друштво, било за пословање.

Претходно наведено иницирало је појаву концепта машинског учења, које се односи на алгоритме који се ослањају на математику, статистику, физичке и друге науке, а користе се са циљем да се на основу података извуку законитости. Законитост је све оно што није случајност. Изучавање и примена статистичких тестова служи за јасно раздвајање законитости и случајности, а алгоритми машинског учења тај процес обављају аутоматски.

Циљ ове дисертације је да се сагледа начин рада алгоритма машинског учења за кластеровање просторних података и да се његовом модификацијом прошири могућност његове примене, а да се добијени резултати примене за идентификацију “црних тачака” на путевима у Републици Србији, чиме би био остварен како научни, тако и друштвени допринос.

Кључне речи: машинско учење, алгоритам кластеровања, пословна интелигенција, складиште података, “црна тачака”

Abstract

The modern way of life and work and the pace of business almost implies the presence of information technologies in all spheres of our lives. The development of technology improves our lives and standards, although on the other hand it generates a huge amount of data. We live in a time when, unfortunately, we are not able to use the potential of all that data, and we intuitively conclude that certain values are hidden in that data from which we can all only benefit.

The goal of applying information technologies is to process data that is collected en masse in such a way as to generate new value, either for society or for business.

The above-mentioned initiated the emergence of the concept of machine learning, which refers to algorithms that rely on mathematics, statistics, physical sciences and other sciences, and are used with the aim of extracting regularities from data. Regularity is everything that is not a coincidence. The study and application of statistical tests serves to clearly separate regularity from randomness, and machine learning algorithms perform this process automatically.

The goal of this dissertation is to examine the operation of a machine learning algorithm for clustering spatial data and to expand its application possibilities by modifying it, and to apply the obtained results to the identification of “black spots” on roads in the Republic of Serbia, thereby achieving both scientific and social contribution.

Keywords: machine learning, clustering algorithm, business intelligence, data warehouse, “black spot”

Садржај

| | |
|---|-----------|
| 1. УВОД..... | 1 |
| 1.2. ПРЕДМЕТ И ЦИЉ ИСТРАЖИВАЊА | 7 |
| 1.2.1. Предмет истраживања | 7 |
| 1.2.2. Циљ истраживања | 11 |
| 1.3. ХИПОТЕЗЕ..... | 13 |
| 1.4. ОЧЕКИВАНИ РЕЗУЛТАТИ И НАУЧНИ ДОПРИНОС..... | 14 |
| 1.5. МЕТОДЕ ИСТРАЖИВАЊА | 16 |
| 1.6. СТРУКТУРА РАДА | 18 |
| 2. ПРЕГЛЕД ЛИТЕРАТУРЕ У ПОГЛЕДУ ДЕФИНИСАЊА ОПАСНИХ МЕСТА – „ЦРНИХ ТАЧАКА“ | 22 |
| 2.1. ПРЕГЛЕД ПОЈМОВА И ДЕФИНИЦИЈА ОПАСНОГ МЕСТА – „ЦРНЕ ТАЧКЕ“ | 22 |
| 2.2. ПРЕГЛЕД ОДАБРАНЕ ЛИТЕРАТУРЕ У КОЈОЈ СУ ОБРАЂИВАНА ОПАСНА МЕСТА – „ЦРНЕ ТАЧКЕ“ НА ПУТЕВИМА – РАЗВОЈ НАУЧНЕ МИСЛИ..... | 25 |
| 2.3. ПРЕГЛЕД СТРАНИХ ИСКУСТАВА У ПОГЛЕДУ ДЕФИНИСАЊА ОПАСНИХ МЕСТА ПО РАЗЛИЧИТИМ МЕТОДОЛОГИЈАМА | 27 |
| 2.3.1. Немачка..... | 27 |
| 2.3.2. Мађарска..... | 27 |
| 2.3.3. Норвешка | 28 |
| 2.3.4. Аустрија..... | 28 |
| 2.3.5. Данска | 29 |
| 2.3.6. Белгија (Фландрија)..... | 29 |
| 2.3.7. Португал..... | 30 |
| 2.3.8. Швајцарска..... | 30 |
| 3. ПРЕГЛЕД ДОСАДАШЊИХ РЕАЛИЗОВАНИХ СТУДИЈА, ПРОЈЕКТА И ИСТРАЖИВАЊА КОЈА СУ ОБРАЂИВАЛА ОПАСНА МЕСТА – „ЦРНЕ ТАЧКЕ“ НА ПУТЕВИМА РЕПУБЛИКЕ СРБИЈЕ | 33 |
| 3.1. ИДЕНТИФИКАЦИЈА ОПАСНИХ МЕСТА НА ДРЖАВНИМ ПУТЕВИМА РЕПУБЛИКЕ СРБИЈЕ СА ПРЕДЛОГОМ МЕРА САНАЦИЈЕ | 33 |
| 3.2. ИДЕНТИФИКАЦИЈА ОПАСНИХ МЕСТА И ПУЊЕЊЕ БАЗЕ ПОДАТАКА О ОПАСНИМ МЕСТИМА НА ДРЖАВНИМ ПУТЕВИМА | 34 |
| 3.3. УПОРЕДНА АНАЛИЗА ОПАСНИХ МЕСТА И РИЗИЧНИХ ДЕОНИЦА НА ДРЖАВНИМ ПУТЕВИМА..... | 35 |
| 3.4. ИДЕНТИФИКАЦИЈА И КЛАСИФИКАЦИЈА ОПАСНИХ МЕСТА „ЦРНИХ ТАЧАКА“ У САОБРАЂАЈУ НА ПОДРУЧЈУ ГРАДА БЕОГРАДА, СА ПРЕДЛОГОМ САНАЦИОНИХ МЕРА | 35 |
| 3.5. МЕТОДОЛОГИЈА ИДЕНТИФИКАЦИЈЕ И УПРАВЉАЊА ОПАСНИМ МЕСТИМА (ЦРНИМ ТАЧКАМА) НА ДРЖАВНИМ ПУТЕВИМА РЕПУБЛИКЕ СРБИЈЕ | 37 |
| 3.6. РЕКАПИТУЛАЦИЈА | 37 |
| 4. ИНТЕГРИСАНИ СИСТЕМИ ЗА УПРАВЉАЊЕ И АНАЛИЗУ ПОДАТАКА | 39 |
| 4.2. ОТВОРЕНИ ПОДАЦИ | 39 |
| 4.2. ПОСЛОВНА ИНТЕЛИГЕНЦИЈА..... | 41 |
| 4.3. СКЛАДИШТА ПОДАТАКА..... | 47 |

| | |
|---|------------|
| 5. МОДИФИКОВАНИ <i>DENSITY-BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE</i> АЛГОРИТАМ | 51 |
| 5.1. ПОЈАМ КЛАСТЕРА ЗАСНОВАНИХ НА ГУСТИНИ | 58 |
| 5.2. ОСНОВНИ ПОЈМОВИ | 60 |
| 5.3. АЛГОРИТАМ | 62 |
| 5.4. ОТКРИВАЊЕ ПАРАМЕТАРА EPS И MINPTS | 63 |
| 5.5. ПРИМЕР | 65 |
| 6. ДЕТАЉАН ОПИС МАТЕМАТИЧКОГ МОДЕЛА ЗА ОДРЕЂИВАЊЕ „ЦРНИХ ТАЧАКА“ (ОДНОСНО ДЕОНИЦА, ЗОНА)..... | 78 |
| 6.1. ФОРМУЛАЦИЈА МАТЕМАТИЧКОГ МОДЕЛА | 78 |
| 6.2. ПРИМЕНА МОДИФИКОВАНОГ АЛГОРИТАМА НА ПРИМЕРУ СА САОБРАЋАЈНИМ НЕЗГОДАМА | 87 |
| 6.3. КОМПАРАТИВНА АНАЛИЗА ДОБИЈЕНИХ РЕЗУЛТАТА..... | 92 |
| 7. АРХИТЕКТУРА И КОНСТРУКЦИЈА ОКРУЖЕЊА ЗА ИДЕНТИФИКАЦИЈУ “ЦРНИХ ТАЧАКА” НА ПУТЕВИМА У РЕПУБЛИЦИ СРБИЈЕ КОРИШЋЕЊЕМ МОДИФИКОВАНОГ АЛГОРИТАМА И СИСТЕМА ПОСЛОВНЕ ИНТЕЛИГЕНЦИЈЕ..... | 93 |
| 7.1. ПРИКУПЉАЊЕ ЗАХТЕВА | 94 |
| 7.2. ПРИКУПЉАЊЕ ПОДАТАКА | 94 |
| 7.3. ЕКСТРАКЦИЈА, ТРАНСФОРМАЦИЈА И УПИСИВАЊЕ ПОДАТАКА У СКЛАДИШТЕ ПОДАТАКА | 102 |
| 7.3.1. <i>Имплементација димензије времена</i> | 102 |
| 7.3.2. <i>Имплементација димензије Врста саобраћајне незгоде</i> | 105 |
| 8. ДОКАЗ ХИПОТЕЗА..... | 116 |
| 9. НАУЧНИ ДОПРИНОС И ДИСКУСИЈА РЕЗУЛТАТА..... | 124 |
| 9.1. МОДИФИКАЦИЈЕ У ОДНОСУ НА СТАНДАРДНИ DBSCAN..... | 125 |
| 9.2. НОВИ ЕЛЕМЕНТИ И ПАРАМЕТРИ МОДЕЛА..... | 125 |
| 9.3. РАЗЛИКЕ У ОДНОСУ НА ПОСТОЈЕЋА РЕШЕЊА..... | 126 |
| 9.4. КОНКРЕТНА ПОБОЉШАЊА..... | 126 |
| 9.5. ДОПРИНОС РАДА..... | 128 |
| 9.6. ПОТВРДА НАУЧНОГ ДОПРИНОСА КРОЗ ПРАКТИЧНУ ПРИМЕНУ | 129 |
| 9.7. ОГРАНИЧЕЊА АКТУЕЛНОГ ИСТРАЖИВАЊА И МОГУЋИ ПРАВЦИ ЗА БУДУЋЕ ИСТРАЖИВАЊЕ | 134 |
| 10. ЗАКЉУЧАК..... | 136 |
| ЛИТЕРАТУРА | 137 |
| ПРИЛОЗИ | 145 |
| СПИСАК СЛИКА | 145 |
| СПИСАК ТАБЕЛА | 148 |

1. Увод

Изразито брз развој информационих технологија условио је њихово присуство у свим сферама свакодневног живота и рада. Готово да данас не постоји ни једна област која не користи или није ослоњена и усмерена на одређена софтверска решења која прикупљају, генеришу или обрађују податке, а све са циљем да информацију по било ком критеријуму корисник добије у приближно реалном времену, а која ће притом бити прецизна, сажета и разумљива. Претходно наведено временом је условило да наше одлуке, како у свакодневном животу, тако и у пословању, постану зависне од коришћења информационих технологија. На самом почетку развоја, идеја о примени информационих технологија везивала се искључиво за повећање продуктивности, ефикасности и ефективности. Тада су информациони системи пројектовани и имплементирани тако да подрже текући начин рада и пословања. Данас, они представљају средство у свакодневном раду од кога се очекује да у што краћем времену обраде све већу количину података која се свакодневно генерише. То представља велики проблем, али и велики изазов.

Временска димензија је подразумевана при анализи података. Готово да нема извештаја и одлуке која се не везује за неку временску одредницу. Поред времена, ништа мање није битна ни просторна димензија. Простор се може дефинисати као све оно у чему се физички простире материја или се одвија неки догађај. Динамички развој комуникационих технологија у последњих неколико десетина година донео је нову димензију у погледу геоинформационих састава и геоинформационих технологија, тј. преносивост. Преносивост омогућава једноставније и брже прикупљање података као и обраду и презентацију истих независно од подручија прикупљања. Са перспективе доносиоца одлука, веома је важан увид о локацији догађаја, људи и простора о којима се одлучује. Већина информација, са којима се корисници сусрећу имају везе са географском локацијом. Најпопуларније средство за приказ геоинформација је карта. Подаци са карте примењују се у оријентацији и навигацији, за визуелизацију, управљање и планирање као и за разне анализе.

Постоје разне потребе за анализом просторних податка, тј. података који су везани за простор. Све већи број података се добија из сателитских снимака, кристалографије, рендгенских зрака или из других аутоматских

система. Просторни подаци користе се у управљању ресурсима, имовинско правним поступцима, просторном планирању, картографији, инфраструктурним пројектима итд. Потреба за откривањем законитости у просторним подацима постаје све већа.

Данашње базе података ефикасно одговарају на сва питања која у себи садрже реч “ШТА?”. Ахилова пета данашњих база података је недостатак информације о простору. И најмоћније базе се блокирају када се нађу пред питањем “ГДЕ?”. То је велики недостатак, обзиром да оваква питања нису безначајна. Циљ дисертације је да се применом савремених технологија обезбеди компромис између географије и информатике, са стварањем могућности да се добије одговор на оба питања. Још значајније за подршку одлучивању је то што ће крајњи резултат рада обезбедити и одговоре на питања „Шта је где ?“ и „Где је шта?“

Намене примене савремених информационо комуникационих технологија у раду са просторним подацима су бројне, а могу се поделити у три основне категорије:

- визуелизација података,
- повезивање географских и атрибутских обележја,
- могућност интердисциплинарног одлучивања.

Могућност да се велика количина података о простору сачуваних у меморији рачунара прикаже у визуелном, једноставном и корисном облику, омогућава визуелизацију и анализу података на нов начин, откривајући притом скривене везе међу њима.

Просторно-временска компонента је кључна у нелинеарном, асиметричном окружењу које се карактерише високим темпом активности и брзом променом ситуације на терену. Нарастање броја доступних података у времену доводи до проблема њихове трансформације у информације корисне за подршку у доношењу одлука по принципу „праве информације у право време“.

Доносиоци одлука захтевају добро форматиране податке за доношење одлука. Више од 75% потребних података у процесу доношења одлука садрже просторне или географске компоненте које се могу визуелно представити уз помоћ карте. На основу експеримената утврђено је да је за право разумевање

података и уочавање њихове међусобне повезаности битан начин на који су приказани. Читање специфичних података је лакше ако су подаци табеларно представљени, али је за уочавање разлике између одређених података у неком временском интервалу бољи графички приказ. Ово је и разумљиво када се узме у обзир да људски визуелни систем представља можда најбоље изграђен спознајни систем јер заузима око 20% мождане масе [1].

Подаци представљени на карти захтевају мање додатног знања за сагледавање ситуације у односу на табеларни приказ. Аналогно томе, време доношења одлуке драматично расте при приказу података табеларно у поређењу са случајем кад се ти подаци приказују на карти.

Објекти у природи и пословању су у великом броју случајева међусобно повезани, па је неопходно препознати начин њиховог груписања и открити њихове међусобне релације и компетитивност. Разлози могу бити различити, а најчешће се ради о функционалним и структурним зависностима.

Циљ кластеровања је откривање тих структурних и функционалних законитости, или патерна, код података. За потребе њихове идентификације се користе кластеринг алгоритми који спадају у групу алгоритама ненадгледаног учења [2].

Енормне количине података које се свакодневно генеришу представљају и велики изазов да се организују и ускладиште. Велику олакшицу у данашње време представља развој хардвера, па проблем меморије није толико изражен, јер је релативно једноставно обезбедити меморијски простор за њихово чување. Алгоритми који користе алате за обраду података имају циљ да те податке обраде у приближно реалном времену и из њих одреде нова знања која ће корисницима служити у процесу одлучивања.

Драстичан феномен експоненцијалног раста количине података у претходном периоду представља велики изазов, а посебно покушај да се симбиозом база података и алгоритама машинског учења, огромне количине ускладиштених података обраде и представе на такав начин који ће менаџерима и доносиоцима одлука бити погодан за аналитику и одлучивање [3]. На тај начин је могуће открити нове повезаности и законитости и доносити боље и праведније одлуке, на основу којих ће пословни модели боље пријављати условима и захтевима корисника и тржишта [4].

Изузетно брз развој информационих технологија довео је до

глобализације светског тржишта, јачања конкуренције, повезивања, брзог протока информација, стварања нових производа, што је узроковало промену свакодневице у којој менаџери доносе стратешке, тактичке и оперативне одлуке. Савремено пословање диктира присуство информационих технологија у свим пословним процесима [5].

Последица комплексности и непредвидивости пословних захтева, њихових међусобних релација и неизбежне неусаглашености циљева појединих саставних делова организације, складиште података треба да задовољи одређене унапред дефинисане критеријуме како би на оптималан начин били подржани процеси одлучивања, планирања и усклађивања у датом пословном систему. Стога је неопходно да складиште података буде јасно усмерено и вођено ка појединачним пословним циљевима и организацији у целини на тај начин да недвосмислено може да одреди стање система у датом тренутку времена. Такође, складиште података треба да обухвати и низ стања система у одређеном временском периоду, како би била омогућена анализа и упоређивање историјских стања, као и предвиђање и планирање будућих стања система. Поред тога, неопходно је да складиште података има могућност интеграције података услед хетерогености извора, при чему интеграцију треба извршити на такав начин да се подаци могу прикупљати без круцијалних измена, односно без промене семантике увезених података [6].

Почетак 21. века означио је прекретницу у технолошком развоју, обзиром да нове технологије укључују питања комплексности и ефикасности. Већа брзина Интернета и робусност је довела до појаве Рачунарства у облаку (*Cloud*, енгл.) који омогућује појаву система пословне интелигенције (*Business Intelligence - BI*, енгл.) која је доступнија компанијама.

У последњој деценији, *BI* и системи за складиштење података еволуирају у односу на свој почетни облик. Савремени системи врше обраду великог броја различитих упита тако да морају бити континуирано доступни различитим групама корисника.

У складу са наведеним, подручје истраживања овог рада представља сагледавање могућности унапређења и проширивања примене алгорита машинског учења за идентификацију кластера над просторним подацима. Ваљаност резултата је проверена над просторним и временским подацима који су карактеристични за саобраћајне незгоде, са циљем идентификације “црних

тачака”, а ради друштвене добити. Сагледане су до сада коришћене методологије за идентификацију опасних места на путевима и дат је предлог коришћења модификованог алгоритма машинског учења за кластеровање над просторним подацима и за идентификацију “црних тачака” на путевима. Алгоритам је коришћен у оквиру система пословне интелигенције, а као извор података коришћено је складиште података. Подаци о саобраћајним незгодама прикупљени су из више извора, и ускладиштени у јединствени формат, са циљем њиховог јединственог сагледавања и анализе. Складиште података представља основу пословне интелигенције. На основу централизованих и форматираних података, модификовањем *DBSCAN (Density-Based Spatial Clustering of Applications with Noise)* алгоритма машинског учења, створена је могућност за развијање новог приступа, који представља предлог за идентификацију опасних места на путевима. На тај начин омогућено је да се кроз анализу велике количине података добију нове информације које до тада нису биле познате [7]. На основу добијених информација могуће је доносити праведније одлуке, са циљем унапређења постојећег стања.

Приступ решавању потенцијалних проблема је сложен, јер иако су многи подаци генерисани са аспекта простора и времена случајни догађаји, детаљном анализом могуће је успоставити корелацију између истих догађаја и уочити међузависности. Применом алгоритама машинског учења могуће је идентификовати потенцијалне и скривене корелације између изабраних атрибута, и мерити њихов међусобни утицај [8]. На тај начин могуће је добијање информација из структурираних и неструктурираних података, из интерних и екстерних извора које, уз помоћ предложених технологија, генеришу у нову вредност, у знање, и могу се користити у процесу одлучивања, визуелном представљању кориснику.

У случају разматрања опасних места на путевима, у литератури је наведено више појмова, па се опасна места називају и местима високог ризика, локацијама са великом вероватноћом настанка незгода, црним тачкама, жариштима, итд. У Републици Србији се у последњој деценији, усталио појам “црне тачке” или опасног места на путу. На сличан начин дефинисане су и методе идентификације опасних места, обзиром да оне нису међу државама стандардизоване и постоји значајна разлика између постојеће праксе и савремених приступа [9].

Енормна количина података представља и велики изазов како за науку, тако и за човечанство, обзиром да постоји низ питања која се узрочно-последично покрећу, а то су питања етике, правних устројстава и других. Коришћење велике количине података покреће и низ механизма којима ће се наведена питања довести у везу.

Адекватно снабдевање информацијама које садрже просторно-временску компоненту у реалном времену, или скоро реалном, омогућава доносиоцима одлука бољу процену ситуације, одређивање адекватне врсте активности (прегруписавање, ослобађање ресурса, избор циљева и слично) и правовремено реаговање и извршавање додељених задатака [1].

1.2. Предмет и циљ истраживања

Теоријско одређење предмета истраживања представља сагледавање просторних, временских и других димензија које су карактеристичне за сваку саобраћајну незгоду, сагледавање до сада коришћених методологија за идентификацију опасних места на путевима и примене кластерована (груписања) алгоритмом машинског учења. На тај начин могуће је проширити могућност његове примене, а као један од примера његове употребе је одређивање опасних места, тј. “црних тачака” или деоница пута. Модификовани алгоритам користи параметризовани математички модел за њихову идентификацију, употребом података о географској локацији на којој се незгода догодила и њеном врстом (саобраћајне незгоде са погинулим и повређеним лицима), а на основу њихових пондерисаних вредности. Овако модификован и имплементиран, алгоритам представља прилог унапређења постојеће методологије за идентификацију црних тачака на путевима, при чему је коришћен у оквиру система пословне интелигенције. Направљено је јединствено складиште података, које представља основу пословне интелигенције. На тај начин је створена могућност да се кроз анализу велике количине података добију информације које до сада нисмо знали, а на основу којих можемо да доносимо праведније и рационалније одлуке и у којима модели боље одговарају идентификацији опасних места, а све са циљем унапређења постојећег стања.

1.2.1. Предмет истраживања

Предмет истраживања докторске дисертације обухвата подручје науке о рачунарству (*Computer Science*, енгл.) [10]. Наука о рачунарству је подручје у коме се, користећи научне принципе и прилазе, долази до нових типова дизајна, решења и имплементација у софтверском домену, као и до нових приступа и развоја оптималних решења везаних за неки од постојећих проблема који се решавају употребом рачунара, односно рачунарске технике.

У складу са литературом рачунарство је могуће поделити на следећих пет дисциплина: рачунарски хардвер и архитектура, системска инфраструктура, софтверске методе и технологије, апликационе технологије и организациона питања и информациони системи. Фокус науке о рачунарству је на: системској

инфраструктури, софтверским методама и технологијама и апликационим технологијама, док врло мало залази у дисциплине рачунарски хардвер и архитектура и организациона питања и информациони системи [10]. У складу са наведеним, предмет докторске дисертације је базиран на поменутих рачунарским дисциплинама које су представљене као окосница науке о рачунарству. Коришћени су и развијени теоријски принципи као основно полазиште из кога даље проистиче развој решења намењених конкретној примени било да се ради о новим апликативним решењима или побољшању кроз одговарајуће иновирање, реконфигурисање и слично.

Подручје које обухвата наука о рачунарству поседује дефинисану базу знања (*Body of Knowledge – BoK*, енгл.) у оквиру које је 2013. године дефинисано укупно осамнаест подручја знања (*Knowledge Areas – KA*, енгл.) формираних на основу одговарајућих области које се проучавају и користе у оквиру науке о рачунарству [11]. Могуће је увидети да је докторска дисертација примарно ослоњена на шест од поменутих осамнаест подручја знања, а то су: рачунарска наука (*Computational Science – CC*, енгл.), графика и визуелизација (*Graphics and Visualization – GV*, енгл.), управљање информацијама (*Information Management – IM*, енгл.), интелигентни системи (*IS*, енгл.), оперативни системи (*Operating Systems – OS*, енгл.) и софтверско инжењерство (*Software Engineering – SE*, енгл.). У оквиру докторске дисертације, а у складу са литературом [11], рачунарска наука пружила је основе за моделовања и симулације, адекватно процесуирање података, као и за сагледавања података и генерисање информација из истих са различитих аспеката, нумеричке методе и слично [12]. Графика и визуелизација је омогућила адекватну визуелизацију добијених резултата и реализованих активности над подацима. Приликом израде докторске дисертације коришћено је више подручја знања, те је тешко подвући јасну црту у коришћењу сваке од области у датом моменту примене.

Предложена тема докторске дисертације припада актуелним областима науке о подацима. Предмет дисертације је примена кластеровања (груписања) алгоритмом машинског учења за одређивање тј. “црних тачака” или деоница пута, чиме се проширује могућност његове примене, а који користи параметризовани математички модел за њихову идентификацију, употребом података о географској локацији на којој се незгода догодила и њеном врстом

(саобраћајне незгоде са погинулим и повређеним лицима), а на основу њихових пондерисаних вредности и који треба да унапреди и прошири постојеће методологије за идентификацију опасних места на путевима у Републици Србији, ради ефикасног доношења оптималних одлука. Истраживачки задатак, предложен у дисертацији, је сложен и мултидисциплинаран, обзиром да укључује области система пословне интелигенције, науке о подацима и интернет технологија.

Период који се користи за идентификацију опасних места варира између једне и пет година.

Анализа досадашњих искустава различитих држава у идентификацији и класификацији „дрних тачака“ на путевима (нпр. Аустрија, Данска, Немачка, Белгија, Норвешка, Португал, Швајцарска, Мађарска) показала је да су у праксама анализираних држава развијени индивидуални приступи у идентификацији „дрних тачака“, који се базирају на објективном методу идентификације. У поређењу са најновијим техникама за идентификацију опасних места, већина практичних приступа који се користе у посматраним државама су засновани на евидентираном броју саобраћајних незгода [13].

Идентификација „дрних тачака“ на путевима у Републици Србији је још увек у почетној фази развоја, зато и даље не постоји национална дефиниција „дрне тачке“, као што је то случај у претходно поменутих државама [9].

Све у свему, максимално искоришћење доступних података, и откривање законитости које у њима владају може донети велику корист и значајно увећати наше знање, док са друге стране такав подухват представља и велики изазов.

Унапређење примене постојећих технологија у инжењерским наукама, па тако и алгорита машинског учења, прати неколико корака од идеје до реализације. Почетни корак обухвата анализу постојећег алгорита машинског учења који се односи на кластероване над просторним подацима, и сагледавање могућности проширења. У савременом пословању информације представљају кључни ресурс којим се генерише конкурентска предност или као друштвена корист, која је у овом случају доминантна јер се на тај начин могу сачувати људски животи. Међутим, дефинисати информацију као ресурс неке организације или друштва представља веома тежак задатак. Ресурси се обично

прецизирају као добра која носе мерљиву или потенцијалну корист кориснику. Информација се може сматрати врстом нематеријалног добра, иако и она сама има материјалне карактеристике [14]. Када се информација комбинује са другим информацијама, може имати далеко већу вредност од иницијалне [15]. Као пример ваљаности предложеног решења, дат је предлог проширења могућности примене постојећег алгоритма машинског учења, од које можемо имати само корист.

У раду су за анализу коришћени отворени подаци. Разлог за то је што су отворени подаци јавно доступни, машински читљиви, тј. могу се форматирати и прилагођавати за процес обраде од стране рачунара и слободни су за употребу у било које сврхе, комерцијалне и некомерцијалне.

Подаци добијени синтезом система пословне интелигенције и предложеног алгоритма могу да буду веома значајани за прецизнију идентификацију опасних места, у односу на до сада коришћене поступке. Они омогућавају уочавање законитости које је иначе веома тешко уочити, или је за њихово уочавање потребно веома много времена, а оно што је најбитније је да се резултати приказују на пригодан начин, што олакшава посао доносиоцу одлуке. Пронађене законитости у подацима, као пронађено знање, треба приказати на одговарајући визуелни начин. Зато је важна визуелизација података и резултата. На пример, различитим моделима и техникама анализе добијају се нумерички подаци. Међутим, људско око ће далеко брже и прецизније препознати резултате анализе уколико су они приказани графички, а разумевање уочене појаве ће бити боље. Зато је посебан акценат био на визуелизацији анализираних података.

У оквиру ове докторске дисертације прво је обрађен теоријски приступ датој проблематици, а потом је извршено постепено кретање ка практичној реализацији и примени.

Операционално одређење предмета истраживања у суштини представља конкретизацију теоријског одређења предмета истраживања докторске дисертације.

Временско одређење предмета истраживања односи се на временски доступне податке о саобраћајним незгодама који су доступни на порталу Отворених података, и омогућавају период истражива од 2018. до 2024. године

Просторно одређење предмета истраживања, подразумева одређивање

простора којим ће истраживање бити обухваћено. У том смислу, просторни оквир који обухвата предмет истраживања докторске дисертације односи се на Републику Србију, њене градове, њихове општине до нивоа географске тачке на којој се догодила одређена саобраћајна незгода идентификована координатама географске дужине и географске ширине.

1.2.2. Циљ истраживања

Проблему огромне количине података из којих треба извући конкретну информацију у најкраћем временском периоду, треба приступити на посебан начин, тако да је ова тема и даље актуелна и нуди доста простора за рад и унапређење постојећих решења и резултата.

У модерном пословном окружењу информација има велику важност. Могућност и брзина сазнавања одређене информације директно утиче на одлуке које се свакодневно доносе. Уколико се са друге стране узму у обзир одлуке које директно утичу на људске животе, онда те информације имају још већи значај. Структуре са више димензија могу бити изузетно комплексне за разумевање [14]. Подаци приказани нумерички у табелама, као и резултати спроведених анализа, су егзактни, али их човек споро (и тешко) перципира. Визуелизација података помаже кориснику да сазна есенцију неког модела података у најкраћем временском периоду.

Пословни субјекти улажу напоре у сврху проналаска начина организације и разумевања мноштва података који су генерисани кроз њихово пословање. Могућност и брзина прикупљања података утиче директно на пословне одлуке, а самим тим и на бољитак организације. Потреба за брзим прикупљањем података, али и за анализом тих података, су могућности којој теже пословни субјекти. Информације добијене обрадом података из база података и складишта података требало би да буду приказане у једноставном визуелном облику јер он пружа јаснији увид од бројчаног табеларног приказа. Бројни произвођачи софтверских пословних решења препознали су потребу за визуелизацијом података, па су у своје пословне системе уврстили функционалности визуелизације кроз корисничке интерфејсе. Даљим развојем пословних система визуелизација је постала једна од важнијих функционалности да би на послетку били развијени посебни алати који служе искључиво за визуелизацију података. Управо такви алати се могу користити за визуелни приказ података, што представља крајњи циљ ове докторске

дисертације, тј. да се након идентификације опасних места она и визуелно прикажу на мапи.

У складу са претходно изложеним, циљеви који су реализовани у оквиру докторске дисертације су следећи:

- Примарни циљ је доказ да се предложеним и модификованим алгоритмом може проширити могућност примене *DBSCAN* алгоритма и добити оптималнији резултати, на основу којих ће бити доношене праведније одлуке, а на основу идентификованих кластера. Наведени алгоритам је узет као репрезентативни представник алгоритама који користе методе засноване на густини за идентификацију кластера. То су методе које претпостављају да кластери представљају подручја у простору која су гушћа у поређењу са остатком простора. Предложени модификовани алгоритам полази од претпоставке да тачке за које се траже кластери (групе), поред просторне идентификације и јединствености, имају и своју вредност (“тежину”). Ту вредност добија параметрима математичког модела којима се реализује. На тај начин добија се оптималније циљано решење. Предложени алгоритам не захтева да се унапред познаје број кластера, што је његова предност. Пожељно је да параметри које алгоритам користи буду што прецизнији, како би били добијени веродостојнији резултати.
- Основни је примена кластерована алгоритмом машинског учења за идентификацију “црних тачака” или деоница пута, који користи параметризовани математички модел за њихову идентификацију, употребом података о географској локацији где се незгода догодила и њеном врстом, а на основу њихових пондерисаних вредности, за прецизнију идентификацију опасних места на путевима, проширењем тренутно примењене методологије, коришћењем и применом система пословне интелигенције;
- Специфичан циљ је идентификација и визуелни приказ опасних места на путевима;
- Посебан циљ дисертације је примена кластерована алгоритмом машинског учења за одређивање опасних места, тј. “црних тачака” или деоница пута, који користи параметризовани математички модел за њихову идентификацију.

1.3. Хипотезе

а) Општа хипотеза

Приликом разматрања предмета докторске дисертације, узети су у обзир проблем и циљ истраживања, као и могућности примене кластеровања алгоритмом машинског учења за одређивање “црних тачака” или деоница пута, који користи параметризовани математички модел за њихову идентификацију, употребом података о географској локацији на којој се незгода догодила и њеном врстом, а на основу њихових пондерисаних вредности,. Коришћењем система пословне интелигенције, унапредиће се досадашњи поступци идентификације опасних места на путевима у Републици Србији. Општа хипотеза гласи:

Применом модификованог алгоритма машинског учења за идентификацију кластера над просторним подацима проширује се могућност примене стандардног DBSCAN алгоритма и успешно су идентификована опасна места, “црне тачке” на путевима на територији Републике Србије, коришћењем јавно доступних података, што се успешно може користити за доношење одлука значајних за управљање.

б) Посебне хипотезе

Посебне хипотезе су везане за карактеристике које има предложени и модификовани алгоритам машинског учења за одређивање опасних места, тј. “црних тачака” или деоница пута, који користи параметризовани математички модел за њихову идентификацију, употребом података о географској локацији где се незгода догодила и њеном врстом (саобраћајне незгоде са погинулим и повређеним лицима), а на основу њихових пондерисаних вредности. Користи за унапређење постојеће методологије и мора да поседује карактеристике које одговарају њеној намени како би одговорио намени, а произилазе из опште хипотезе и односе се на одређивање предмета истраживања и гласе:

- X1: *Модификовани DBSCAN алгоритам користи параметре математичког модела и не захтева унапред познат број кластера, што доводи до прецизнијих резултата;*
- X2: *Модификовани DBSCAN алгоритам се може применити у системима пословне интелигенције заснованим на концепту складишта података;*
- X3: *Примена модификованог DBSCAN алгоритма побољшава идентификацију опасних тачака на путевима и омогућава бољу анализу података за доношење одлука;*
- X4: *Примена модификованог DBSCAN алгоритма омогућава лакше креирање и дељење знања о опасним тачкама међу доносиоцима одлука;*
- X5: *Модификовани DBSCAN алгоритам побољшава идентификацију опасних тачака и омогућава скалабилну примену на већим сетовима података.*

1.4. Очекивани резултати и научни допринос

Научни допринос овог рада лежи у предлогу оригиналног приступа да се применом кластерована модификованим алгоритмом машинског учења за одређивање “црних тачака” или деоница пута, који користи параметризовани математички модел идентификације, употребом података о географској локацији где се незгода догодила и њеном врстом, а на основу њихових пондерисаних вредности, са циљем унапређења постојеће методологије који укључује детекцију опасних места на путевима у Републици Србији. Сагледавањем литературе која се бави наведеном проблематиком креиран је нов приступ за идентификацију „црних тачака”. Коришћењем система за управљање базама података, јавно доступним подацима пружа се нова улога у идентификацији опасних места, првенствено коришћењем предложеног модификованог, итеративног алгоритма који се користи у оквиру система пословне интелигенције, за елиминацију одређене локације, да би се на крају, изолованим поступцима дошло до најучесталијих места на којима се догађају саобраћајне незгоде. Међусобна компарација вршена је према последицама које су изазване саобраћајном незгодом. Једна од битних карактеристика је та што се за одређивање опасних места нису узимале искључиво локације на којима су се догодиле саобраћајне незгоде са погинулим лицима, као што је уобичајено, већ је опсег посматрања проширен и на места на којима се учестало догађају и саобраћајне незгоде са повређенима, које су по последицама рангиране испод саобраћајних незгода са страдалим лицима. Ово представља знатно унапређење, али са друге стране и знатно усложњава поступак идентификације опасних места, обзиром да је потребно извршити анализу огромне количине података.

Као продукт предложеног научно-истраживачког рада, остварено је више општих, стручних, практичних и научних резултата и доприноса:

1. Прикупљање, систематизација и критичка анализа постојећих решења, знања, искустава и научних резултата који се односе на идентификацију опасних места на путевима у Републици Србији;
2. Дефинисање и примена кластерована модификованим алгоритмом машинског учења за одређивање “црних тачака” или деоница пута, који користи параметризовани математички модел за њихову

идентификацију, употребом података о географској локацији где се незгода догодила и њеном врстом, а на основу њихових пондерисаних вредности, а који ће представљати унапређење постојеће методологије за идентификацију опасних места на путевима у Републици Србији;

3. Развој и примена статистичких модела за анализу података.

Анализом добијених резултата приказано је тумачење поређењем у различитим временским периодима и интервалима, ради провере поузданости предлога. Такође, остварен је допринос ове тезе у области анализе саобраћајних незгода кроз интеграцију свих расположивих података и утврђивање њихове корелације. Поред тога, допринос и додатна вредност се огледају у безбедности у саобраћају, саобраћајној култури, образовању и васпитању.

Практичан допринос лежи у чињеници да резултат дисертације представља проширење могућности примене стандардног алгоритма машинског учења за кластеровање просторних података, што представља унапређење постојеће методологије и омогућава анализу и откривање знања у оквиру доступних података. То подразумева стављање информационих ресурса на располагање доносиоцима одлука који могу у сваком тренутку, интерактивно добити жељене извештаје уз минималну подршку информатичке службе. Тиме се значајно смањује временско кашњење између идентификације евентуалних проблема и предузимања неопходних акција како би се ти проблеми решили.

Дисертација је писана у складу са резултатима спровођења процеса научног истраживања, прегледа области истраживања, анализе постојећих решења и приступа, развоја нових приступа пројектовања и коришћења (са свим њиховим саставним деловима), и приказом примене имплементираног новог приступа. На тај начин се остварује добијање информација из структурираних и неструктурираних података, из различитих извора који се уз помоћ предложених технологија генеришу у нову вредност, у знање, и могу се користити у процесу одлучивања визуелном представом кориснику. Једна од неизоставних техника која прати све етапе обраде података од чишћења, преко претпроцесирања и претраживања података је визуелизација података. Визуелизација података може на једноставан и ефикасан начин да укаже на основне смернице за даљу анализу у било којој етапи истраживања. Појам визуелизације се најчешће објашњава као „ментална слика”. У

техничком смислу она се бави приказивањем понашања односно стања, посебно при комплексним условима, на начин разумљив људском оку/уму. Она представља једну од популарних метода упознавања с подацима, информацијама, па и једноставнијим анализама.

1.5. Методе истраживања

Систематизован преглед литературе у теоријском делу рада биће реализован коришћењем дескриптивних и компаративних метода, у циљу проналажења бољих заједничких карактеристика сродних појава у досадашњим истраживањима, синтезом и анализом добијених података. Подаци потребни за имплементацију и примену кластеровања алгоритмом машинског учења за одређивање “црних тачака” или деоница пута, који користи параметризовани математички модел за њихову идентификацију, употребом података о географској локацији где се незгода догодила и њеном врстом (саобраћајне незгоде са погинулим и повређеним лицима), а на основу њихових пондерисаних вредности, биће преузети са портала отворених података и релевантних репозиторијума. Резултати добијени емпиријским истраживањем биће статистички обрађени стандардизованим статистичким методама.

Предвиђено је креирање одговарајућег складишта података. За поступке прикупљања, складиштења, обраде и извештавања биће коришћени алати пословне интелигенције. За организацију базе података биће коришћени *OLAP* технологија (за интерактивно аналитичко процесирање, као и за сакупљање, обраду и анализу), као и алати за извештавање и колаборативно доношење одлука. Приликом реализације докторске дисертације коришћене су следеће методе:

1. Метода прикупљања референтних података са портала отворених података Републике Србије, у електронском облику, са циљем формирања у облик погодан за анализу;
2. Метода дизајнирања, развоја и имплементације модела складишта података. Моделом је представљена упрошћена и реална слика оригинала. Упрошћавање се не односи на елементе истраживања. Напротив, настојало се да они буду што је могуће више верни оригиналу;
3. Метода складиштења података уз помоћ система за управљање

- базама података (СУБП) у сврху трајног или привременог смештаја података коришћењем *SQL* сервера;
4. Методе анализе и синтезе за анализу прикупљених података и за рашчлањивање сложених појмова и проблема на једноставније саставне делове и обрнуто;
 5. Метода дескрипције за опис фактора и чинилаца саобраћајних незгода;
 6. Компаративни методи који служе за проверу оправданости и упоређивање тренутног поступка идентификације опасних места на путевима у Републици Србији са резултатима предложене методологије и подацима добијеним коришћењем система пословне интелигенције;
 7. Метода индукције за анализу појединачних и међусобно повезаних утицаја фактора саобраћајних незгода са аспекта времена;
 8. Статистичке методе за статистичку обраду података, са акцентом на примену функција за одређивање међузависности просторне и временске димензије на саобраћајне незгоде;
 9. Методе предиктивне аналитике за повезивање података из различитих извора, са циљем да се боље сагледа тренутна ситуација и донесу правилне пословне одлуке;
 10. Методе софтверског инжињерства;
 11. Елементи теоријског одлучивања и алгоритми одлучивања.

1.6. Структура рада

Рад садржи увод, осам поглавља и закључак.

Уводни део представља осврт на циљ истраживања, опис проблема и сажет садржај рада. Након тога је у теоријском делу рада дат сажет преглед појмова и дефиниција везаних за алгоритме машинског учења, са акцентом на кластеровање просторних података.

У поглављу **Преглед литературе у погледу дефинисања опасних места – „црних тачака“ на путевима** сагледано је тренутно стање обзиром да литература указује на више појмова која дефинишу опасна места, као што су места високог ризика, локације са великом вероватноћом настанка незгода, „црним тачкама“, „жариштима“, итд. У Републици Србији се последњих година усталио појам „црне тачке“ или опасног места на путу. У овом делу сагледана је референтна литература у погледу дефинисања опасних места – „црних тачака“ на путевима и поступак њихове идентификације.

У оквиру подпоглавља преглед одабране литературе у којој су обрађивана опасна места – „црне тачке“ на путевима – развој научне мисли дат је осврт на досадашња истраживања, која су карактеристична и узимају се као полазна основа јер садрже квалитетан преглед критеријума и метода за идентификацију опасних места на путевима у осам европских држава[9].

Подпоглавље **Преглед досадашњих реализованих студија, пројеката и истраживања која су обрађивала опасна места – „црне тачке“ на путевима Републике Србије** представља одговор на основну хипотезу везану за предлог примене кластеровања алгоритмом машинског учења за одређивање „црних тачака“ или деоница пута, који користи параметризовани математички модел за њихову идентификацију, употребом података о географској локацији где се незгода догодила и њеном врстом (саобраћајне незгоде са погинулим и повређеним лицима), а на основу њихових пондерисаних вредности, са циљем проширења постојеће методологије за идентификацију опасних места.

Поглавље **Интегрисани системи за управљање и анализу података** даје осврт на начин имплементације који омогућава повезивање различитих извора података, у конкретном случају јавно доступних података са портала Отворених података, њихово формирање и централно складиштење ради ефикаснијег управљања и формирања. Коришћењем система Пословне

интелигенције и складишта података, наведени подаци су ефикасно припремљени за коришћење у оквиру модификованог алгоритма.

Подпоглавље **Отворени подаци** описује концепт у коме су настали и у коме се развијају отворени подаци. Отворени подаци представљају концепт у коме одређени подаци треба да буду слободно (јавно) доступни свима, на коришћење и поновну употребу, без ауторских и других ограничења. У складу са принципима других иницијатива/покрета, као што су Софтвер отвореног кода, подаци отвореног садржаја и других, који промовишу предности отвореног приступа и дељења информација, тако и иницијатива отворених података дефинише основне форме комерцијалне и некомерцијалне употребе и размене података, уз поштовање одређених правила кроз одговарајуће лиценцирање сваког скупа података.

У наредном подпоглављу, приказан је концепт функционисања **система пословне интелигенције**. Пословна интелигенција представља начин пословног понашања који омогућава, да се пословне одлуке на свим нивоима одлучивања (стратешком, тактичком и оперативном) доносе и заснивају на релевантним и ажурним пословним информацијама, а не на предосећају и субјективном утиску. У овом раду, предложени алгоритам се заснива на резултатима и подацима добијеним коришћечем система пословне интелигенције.

Откривене законитости у подацима, као откривено знање, треба приказати на одговарајући графички начин. Из тог разлога је важна визуелизација података. Различитим моделима и техникама анализе нпр. добијају се бројчани подаци. Међутим, људско око ће много брже и прецизније препознати резултате анализе ако су резултати приказани графички уместо бројчано, а разумевање уочене појаве ће бити боље. Зато је у раду посебан акценат дат на визуелизацију анализираних података.

У наредном подпоглављу разматрана су **Складиште података** (*Data Warehouse*, енгл.), из разлога што оно чини основу система пословне интелигенције. Складиште података одржава скуп историјских података који представљају низ стања система проузрокованих догађајима или активностима које се обављају и систему који се прати. Ови подаци се периодично прикупљају, најчешће из више извора, како би се вршила одређена анализа, са циљем праћења трендова и дефинисања стратешких одлука.

Следеће поглавље под називом **Модификовани *Density-Based Spatial Clustering of Applications with Noise*** описује постојећи алгоритам кластеровања који се примењује на просторне податке *DBSCAN* и предлог једне његове модификације са детаљним описом итерација. *DBSCAN* је алгоритам кластеровања који кластере одређује на основу густине распореда елемената скупа података у простору. *DBSCAN* је коришћен као инспирација за развој многих других алгоритама заснованих на густини. За *DBSCAN* биће приказан школски пример ради сагледавања његовог рада. Затим је алгоритам примењен на реалном случају, а то су подаци о саобраћајним незгодама. Над истим скупом података је примењен и модификовани алгоритам ради приказа боље идентификације кластера. Затим је приказана примена модификованог алгоритма за одређивање опасних места на основу података о локацијама саобраћајних незгода, са описом кластеровања алгоритмом машинског учење за одређивање “црних тачака” или деоница пута, који користи параметризовани математички модел за њихову идентификацију, употребом података о географској локацији где се незгода догодила и њеном врстом (саобраћајне незгоде са погинулим и повређеним лицима), а на основу њихових пондерисаних вредности.

У поглављу **Детаљан опис математичког модела за одређивање „црних тачака“** је дат опис математичког модела. Модел је нестационаран, обзиром да садржи променљиве које се мењају у току времена. Променљиве које фигуришу у моделу, а подлежу промени у току времена су број и последице саобраћајних незгода. Поред тога, саобраћајне незгоде се различито позиционирају на мрежи путева и улица у току времена, те се и места „нагомилавања“ саобраћајних незгода могу евентуално променити.

Након преидентификације, односно, коначне идентификације „црних тачака“, потребно је рангирати опасна места. Рангирање је изведено на основу збира пондерисаних вредности за сваку од саобраћајних незгода која се догодила у пречнику од 500 метара, у последњем трогодишњем периоду, према [83]. Критеријуми су детаљно објашњени у поглављима која описују алгоритам и математички модел.

Поглавље **Архитектура и конструкција окружења за идентификацију “црних тачака” на путевима у Републици Србије коришћењем модификованог алгоритма и система пословне**

интелигенције описује процес аутоматског креирања складишта података и апликације за идентификацију опасних места на путевима у Републици Србији, где корисник бира и уноси параметре које користи модификовани *DBSCAN*, а као излаз добија кластере који представљају “црне тачке” на путевима.

У поглављу **Доказ хипотеза** је дискутовано о општој и постављеним посебним хипотезама.

Поглавље **Научни допринос и дискусија резултата** представља синтезу резултата истраживања и јасно формулише оригиналне елементе рада који га издвајају од постојећих решења у литератури. У овој целини приказан је научни допринос са аспекта развоја модификованог *DBSCAN* алгоритма са пондерисаним параметрима, увођења математичког модела за процену значаја кластера, интеграције алгоритма са системом пословне интелигенције и складиштем података, као и методологије за идентификацију и визуелизацију „црних тачака“ на путевима. На тај начин, поглавље обједињује сажет преглед теоријских и практичних аспекта доприноса, показујући како предложени модел не само да унапређује постојеће технике кластеризације, већ и пружа конкретну примену у реалним условима, чиме се обезбеђује нова вредност у области пословне аналитике и управљања подацима, са акцентом на научни допринос.

Закључни део дисертације разматра целокупну сврсисходност истраживања. Осим закључака рада размотрене су и могућности даљег унапређивања и испитивања у обрађеној области. Дат је критички осврта на рад, уз издвајање теоријских и практичних доприноса рада. У овом делу дисертације анализирани су могућности практичне примене развијеног алгоритма, као и представљање планираних праваца даљег развоја и истраживања.

Резултати истраживања презентовани су текстуално, описивањем, објашњавањем и дискутовањем, уз табеларни и графички приказ. На основу прегледа садржаја докторске дисертације могуће је закључити да примењене научне методе и технике по свом значају и структури одговарају проблему, предмету и циљевима дисертације.

2. Преглед литературе у погледу дефинисања опасних места – „црних тачака“

Бројна истраживања о утицају појединих фактора безбедности саобраћаја на настанак саобраћајних незгода показала су значајан утицај пута, самостално или у спрези са другим факторима, на настанак саобраћајних незгода [18]. Сама појава „нагомилавања“ саобраћајних незгода са сличним карактеристикама омогућује квалитетне информације, као што су места где треба применити одговарајуће инжењерске мере. Са тим у вези, у раду је примењен алгоритам машинског учења за кластеровање као предлог за идентификацију опасних места на путевима. Новија истраживања у Великој Британији показала су да су утицајни фактори који се односе на пут и окружење пута заступљени у 13% пријављених саобраћајних незгода са настрадалим лицима [19].

Анализе евидентираних утицајних фактора у Републици Србији показале су да су утицајни фактори који се односе на пут и путну околину заступљени у 16% саобраћајних незгода са настрадалим лицима [20].

Управљање опасним местима – „црним тачкама“ се састоји од следећих активности [21]:

- 1) дефинисање и идентификација опасних места на путевима,
- 2) анализа саобраћајних незгода и фактора ризика на опасним местима, у циљу идентификације фактора који доприносе саобраћајним незгодама и предлога одговарајућег „третмана“ опасних места, и
- 3) имплементација и евалуација примењених третмана на опасним местима.

С обзиром да су различите државе оствариле индивидуалне приступе у примени овог алата, нема усаглашеног приступа, а свака држава која почиње примену, чини то у сагласности са актуелним могућностима [22]. Приликом идентификације користе се различити показатељи нивоа ризика, као и различити критеријуми идентификације, што је детаљније објашњено у тексту који следи.

2.1. Преглед појмова и дефиниција опасног места – „црне тачке“

Просторна дистрибуција саобраћајних незгода није равномерно распоређена по путу или путној мрежи. На релативном малом броју места дешава се релативно велики број саобраћајних незгода. Део пута на коме се у

одређеном временском периоду догоди већи број незгода од просечног на сличној деоници пута, најчешће се назива опасна деоница, опасно место, опасан сектор, опасан километар, опасан пут или опасна зона. То су места са високом учесталости саобраћајних незгода [23].

У литератури се среће више појмова код описа опасних места, као што су: места високог ризика, локације са великом вероватноћом настанка незгода, „црне тачке“, „жаришта“, итд [24]. У Републици Србији се последњих година усталио појам „црне тачке“ или опасног места на путу. Слична је ситуација и када су у питању дефиниције и методе идентификације опасних места, с обзиром да оне међу државама нису стандардизоване и да постоји значајна разлика између постојеће праксе и савремених приступа [9].

Генерално, најчешће дефиниције „црне тачке“ се могу разликовати на следећи начин [25]:

- 1) нумеричке дефиниције,
- 2) статистичке дефиниције, и
- 3) дефиниције засноване на моделима.

Нумеричке дефиниције „црне тачке“ су засноване на броју саобраћајних незгода, стопи незгода, или њиховој комбинацији [26].

Статистичке дефиниције „црних тачака“ се базирају на поређењу евидентираног броја саобраћајних незгода са просечним бројем незгода за сличну врсту локације [27].

Дефиниције „црне тачке“ засноване на моделима изведене су на основу мултиваријантног модела за предвиђање саобраћајних незгода [28].

Према ауторима у [29] “црне тачке” на путевима су локације или деонице пута на којима су се догодиле саобраћајне незгоде, а које се дефинишу као догађаји на путу који су неочекивани и ненамерни и у којима учествују возила са или без других учесника у саобраћају, што резултира људским жртвама и/или губитком имовине.

Аутори [30] дефинишу да је “црна тачка” деоница пута са високим ризиком од саобраћајне несреће.

У свом истраживању [31] аутори наводе да је “црна тачка” деоница на путу дужине око 500 метара на којој се догодило 5 саобраћајних незгода са смртним исходом и тешким повредама или 10 смртних случајева у периоду од три

Године.

2.2. Преглед одабране литературе у којој су обрађивана опасна места – „црне тачке“ на путевима – развој научне мисли

У истраживању [32] приказан је преглед критеријума и метода за идентификацију опасних места на путевима у осам европских држава.

Систематичан преглед дефиниција опасних места према анализираним критеријумима у одабраним европским државама приказан је у Табели 1 [32]. У четири државе опасна места на путу се идентификују као места која имају значајно већи број незгода у односу на нормалан број. Период који се користи за идентификацију опасних места варира између једне и пет година, у зависности од броја незгода на конкретној локацији.

Табела 1. Преглед дефиниција опасних места на путевима у одабраним европским државама [32]

| Држава | Узима у обзир популацију локација | Примена „клизећег прозора“ | Узима у обзир нормални ниво безбедности | Евидентирани или очекиван број незгода | Узима у обзир тежину незгода | Период посматрања |
|---------------------|---|---|---|---|--|---|
| Аустрија | Не | Да, 250 m | Да, на основу критичне вредности стопе незгода | Евидентирани, минимална критична вредност 3 – у функцији интензитета саобраћаја | Не | 3 године |
| Данска | Да, детаљна категоризација елемената пута | Да, за деонице пута – променљиве дужине | Да, применом модела за предикцију незгода | Евидентирани, заснован на статистичком тесту – најмање 4 незгоде | Не | 5 година |
| Фландрија (Белгија) | Не | Да, 100 m | Не | Евидентирани, пондерисан по тежини | Да, на основу пондера | 3 године |
| Немачка | Не | Не, користе се мапе незгода | Не | Евидентирани, минимална вредност 3 или 5 | Да, на основу различитих критичних вредности | Једна година (све незгоде) или 3 године (незгоде са настрадалима) |
| Мађарска | Не | Да, 100 или 1000 m | Не | Евидентирани, минимум 4 | Не | 3 године |
| Норвешка | Не када се идентификују црне тачке | Да, 100 m (тачка) или 1000 m (деоница) | Да, на основу нормалне стопе незгода за елементе пута | Евидентирани, већи од нормалног броја на основу статистичког теста, минимална вредност 4 (тачке) или 10 (деонице) | Да, на основу процене трошкова незгода и потенцијалних уштеда | 5 година |
| Португал | Да, за једну дефиницију, а не за другу | Да, за једну дефиницију, а не за другу | Да, за једну дефиницију, а не за другу | Евидентирани у једној дефиницији (минимум 5), очекивани у другој | Да, у једној дефиницији (на основу тежинских пондера), не у другој | Једна или 5 година |
| Швајцарска | Да, отворени путеви и раскрснице | Не, фиксне деонице различите дужине | Да | Евидентирани, низ критичних вредности | Да, на основу различитих критичних вредности | 2 године |

У истраживању [33] аутори су спровели експериментално оцењивање различитих метода за идентификацију „црних тачака“. Истраживање је нарочито третирао различите нивое поузданости модела, посебно у погледу идентификације лажно „негативних“ (идентификација високоризичног места као безбедног) и лажно „позитивних“ опасних места (идентификација безбедног места као високоризичног).

Истраживање аутора [34] имало је за циљ да у проблематици избора „црних тачака“ представи приступ заснован на вишеструком тестирању.

У свом истраживању [35] аутор је закључио да емпиријски Бајесов метод има бољи учинак у односу на остале методе идентификације „црних тачака“.

У свом истраживању [36] аутори су закључили да традиционални приступ „црних тачака“ није подобан метод, нарочито на путној мрежи са малим

саобраћајним оптерећењем и „расипањем“ саобраћајних незгода. Са друге стране, аутори сматрају да је индекс ризика, заснован на провери безбедности пута, оправдана алтернатива, са перформансама које су упоредиве са најновијим емпиријским Бајесовим методом. Стога се њихова обострана примена препоручује као одговарајућа замена за традиционално управљање „црним тачкама“, али и као искорак ка проактивном управљању безбедношћу путне мреже.

Аутори [37] у раду *Analysis Of Road Traffic Accident And Identification Of Blackspot Locations* примењују статистичку анализу саобраћајних незгода и методе просторног кластерованја како би идентификовали „црне тачке“ на путевима, користећи учесталост и тежину несрећа као главне критеријуме за локацију ризичних зона.

Аутори [38] у раду *Urban Road Accident Black Spot Identification and Classification* користе комбинацију класификационих модела и просторне анализе, при чему се црне тачке идентификују на основу густине незгода и њихове категорије (смртне, тешке, лакше), што омогућава систематско рангирање ризичних урбаних локација.

Аутори [39] у раду *Identifying Black Spot: A GIS Approach to Road Safety* примењују ГИС технологију за просторну визуализацију и анализу саобраћајних незгода, где се црне тачке одређују кроз географско преклапање података о учесталости несрећа и карактеристикама путне инфраструктуре.

Аутори [40] у раду *Spatial clustering of accident-prone zones using GIS and ML* комбинују методе просторног кластерованја и машинског учења, при чему се црне тачке идентификују кроз анализу образаца незгода и предиктивно моделирање ризичних зона, што омогућава динамичко откривање нових опасних локација.

Аутори [41] у раду *Bayesian approaches to black spot detection in urban traffic* примењују бајесовске статистичке моделе за процену вероватноће настанка незгода на одређеним локацијама, чиме се црне тачке идентификују кроз интеграцију историјских података и вероватносних предикција ризика у урбаним условима.

Аутори у [42] предлажу да се у сврху смањења штете од саобраћајних незгода на брзим путевима и побољшања тачности идентификације „црних тачака“ у саобраћајним незгодама, примени метод заснован на секундарној

подели путних јединица и емпиријској Бајесовој методи.

2.3. Преглед страних искустава у погледу дефинисања опасних места по различитим методологијама

2.3.1. Немачка

Приликом идентификације „црних тачака“ у Немачкој, разликују се појмови „црне тачке“, „црне деонице“ и „црне зоне“, а разлика између ових појмова се прави на основу анализе мапе саобраћајних незгода.

„Црна тачка“ је дефинисана на следећи начин: „Место се сматра „местом са великом концентрацијом незгода“ уколико се велики број саобраћајних незгода догађа на веома краткој деоници пута на мрежи путева, односно, уколико је одређени број незгода достигнут или премашен коришћењем приказа једногодишњег и/или трогодишњег периода саобраћајних незгода на мапи [9]. Уколико је нека од граничних вредности приказаних у Табели 2 достигнута или премашена, требало би спровести прелиминарну анализу одређеног „места са великом концентрацијом незгода“.

Табела 2. Критичне вредности за идентификацију „црних тачака“ у Немачкој [43]

| Извор података | Критичан број незгода | Дужина периода |
|----------------|--------------------------------------|----------------|
| 1-годишња мапа | 5 незгода сличног типа | 12 месеци |
| 3-годишња мапа | 5 незгода са повређеним лицима | 36 месеци |
| 3-годишња мапа | 3 незгоде са тешко повређеним лицима | 36 месеци |

2.3.2. Мађарска

У Мађарској су у употреби две дефиниције „црних тачака“. Ван насељених места, „црна тачка“ је дефинисана као место на коме је евидентирано најмање 4 саобраћајне незгоде у току три године на деоници пута дужине до 1.000 m [44].

У насељеним местима, „црна тачка“ је дефинисана као место на коме је евидентирано најмање 4 саобраћајне незгоде у току три године на деоници пута дужине до 100 m. Претрага се врши применом „клизног прозора“ дужине 1.000 m или 100 m.

Приликом рангирања, у обзир се узима и обим саобраћаја, како би биле идентификоване „црне тачке“ које имају већу стопу саобраћајних незгода од

нормалне. На тај начин је мађарски приступ у идентификацији „црних тачака“ практично близак методу са стопом и бројем саобраћајних незгода.

2.3.3. Норвешка

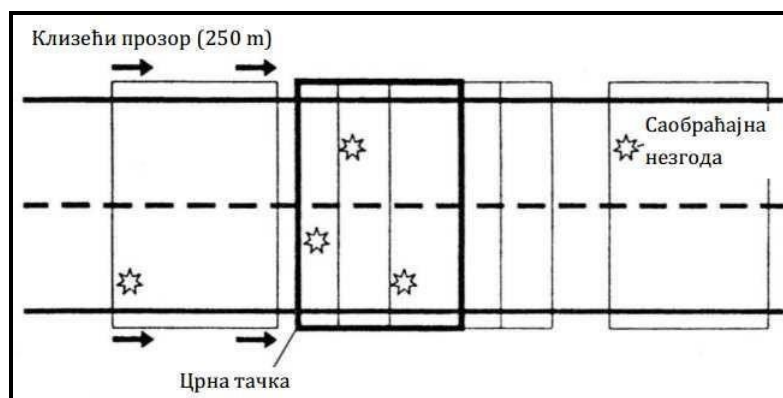
У Норвешкој се разликују појмови „црне тачке“ и „црне деонице“. „Црна тачка“ је свако место дужине до 100 m на коме је евидентирано најмање 4 саобраћајне незгоде са повређеним лицима током последњих 5 година. „Црна деоница“ је свака деоница пута дужине до 1.000 m на којој је евидентирано најмање 10 саобраћајних незгода са повређеним лицима током последњих 5 година [9]. Идентификација „црних тачака“ и „црних деоница“ се врши применом „клизећег прозора“ који се поставља на локације саобраћајних незгода. „Црне деонице“ се често састоје од неколико „црних тачака“ које се налазе у близини.

На местима која су високо рангирана врше се детаљније инжењерске анализе које обухватају детаљну анализу саобраћајних незгода, теренска истраживања, снимање понашања учесника у саобраћају, итд. На основу детаљне анализе, предлажу се мере за унапређење безбедности и процењују се њихови трошкови и ефекти.

2.3.4. Аустрија

У Аустрији, „црне тачке“ су дефинисане у Приручнику за планирање, изградњу и одржавање путева из 2002. године. Према Приручнику, места саобраћајних незгода се разврставају у „црне тачке“ и „опасна места“ у зависности од догађаја саобраћајних незгода у одређеном временском периоду [9].

За идентификацију „црних тачака“ примењује се тзв. „клизећи прозор“ дужине 250 m. Прозор прати путању анализираних саобраћајница и обележава сваку локацију где је испуњен један од два поменута критеријума (Слика 1).



Слика 1. Идентификација „црних тачака“ коришћењем приступа „клизећег прозора“ [43]

2.3.5. Данска

Дефиниција „црне тачке“ у Данској заснива се на детаљној класификацији путне мреже на различите врсте деоница путева и различите врсте раскрсница [9]. На државним путевима разликују се деонице путева, кружне раскрснице и остале раскрснице. Свака од ових група редом састоји се од неколико типова, као што су ауто-путеви, остали путеви са физички раздвојеним коловозним тракама, двотрачни путеви ван насеља, двотрачни путеви у насељу и сл.

За идентификацију „црних тачака“ користи се тест који се заснива на Поасоновој расподели. Претпоставља се да модели који су довољно добро дефинисани представљају систематску варијацију броја саобраћајних незгода у мери у којој се варијација може приписати обиму саобраћаја и променљивама које се користе за класификацију путева и раскрсница. Уколико одређено место има значајно већи број саобраћајних незгода него што се може објаснити случајном варијацијом, претпоставка је да неуобичајено већи број саобраћајних незгода макар делимично постоји услед локалних фактора ризика.

С обзиром на то, дефиниција „црне тачке“ која се користи у Данској подразумева место које има евидентирани број саобраћајних незгода већи од дефинисаног минималног броја и значајно већи од нормалног очекиваног броја саобраћајних незгода на сличном типу елемената пута (деоница, кружна раскрсница или раскрсница).

Минимални број саобраћајних незгода да би се одређено место идентификовало као „црна тачка“ је 4 саобраћајне незгоде које су се догодиле током петогодишњег периода [45].

2.3.6. Белгија (Фландрија)

У Фландрији је у употреби следећа дефиниција „црне тачке“, заснована на полицијским извештајима о саобраћајним незгодама:

У разматрање се узима свако место на којем се у последње три године догодило три или више саобраћајних незгода.

Свака локација која је идентификована као „црна тачка“ мора бити дужине до 100 m. За идентификацију „црних тачака“ користе се подаци о саобраћајним незгодама у трогодишњем периоду [41].

За разлику од осталих дефиниција „црних тачака“, дефиниција која се користи у Фландрији узима у обзир тежину саобраћајних незгода додељивањем већих пондера тешким и смртним повредама, у односу на лаке повреде. На тај начин, место на којем се догодило 2 саобраћајне незгоде са погинулим лицима и једна саобраћајна незгода са тешким телесним повредама имаће вредност приоритета 13, док ће место на којем се догодило 10 саобраћајних незгода са лаким телесним повредама имати вредност приоритета 10.

2.3.7. Португал

Идентификација „црних тачака“ у Португалу спроводи се на националној мрежи путева којом управља Португалска агенција за путеве. У употреби су две дефиниције „црних тачака“. Према једној од дефиниција, „црна тачка“ је деоница пута дужине до 200 m, са 5 или више саобраћајних незгода (укупан број саобраћајних незгода, било које врсте) и индексом тежине већим од 20, посматрано за период од последњих годину дана [47]. Притом, у овом случају не постоји разлика између саобраћајних незгода које су се догодиле на раскрсницама или ван раскрсница. Идентификација се врши применом „клизећег прозора“ који се помера дуж пута.

У другом методу идентификације „црних тачака“ усвојена је следећа дефиниција: „црна тачка“ је географско подручје где је очекивани број саобраћајних незгода већи него на сличним (не нужно суседним) подручјима, услед утицаја карактеристика пута специфичних за то подручје [48]. У пракси, ова дефиниција се различито примењује у зависности од тога да ли су се саобраћајне незгоде догодиле на раскрсницама или ван раскрсница.

2.3.8. Швајцарска

У Швајцарској је у употреби следећа дефиниција „црне тачке“:

„Црна тачка“ је дефинисана као деоница пута или раскрсница, на којој је број саобраћајних незгода „значајно изнад“ броја саобраћајних незгода на сличним местима. Слична места се идентификују на основу следеће процедуре [49]:

1. Деоница која се анализира се сврстава у једну од две различите врсте путева: потези отворених путева и потези који садрже раскрснице. Дужина потеза отворених путева зависи од врсте пута на начин да, што је пут веће важности, то је дужина потеза отворених путева већа (варира између 100 m и 500 m). Потези са раскрсницама обухватају велике (значајне) раскрснице.
2. Просечна стопа саобраћајних незгода на отвореном путу на читавој деоници израчунава се на основу саобраћајних незгода до којих долази на потезима отвореног пута.
3. Просечна стопа саобраћајних незгода на раскрсницама на одређеној деоници израчунава се на основу саобраћајних незгода до којих долази на потезима са раскрсницама.
4. Како за стопу саобраћајних незгода на отвореним путевима, тако и за стопу саобраћајних незгода на раскрсницама, израчунава се опсег. Израчунате стопе саобраћајних незгода требало би да се налазе у том опсегу (који у основи представља интервал поверења са вероватноћом грешке од 5%).
5. Овај опсег представља очекивану стопу саобраћајних незгода израчунату на основу анализираних интервала.
6. Следећи корак представља одређивање места саобраћајних незгода. Места саобраћајних незгода су потези отворених путева или потези са раскрсницама, на којима број саобраћајних незгода премашује процењени очекивани број саобраћајних незгода заснован на нормалним стопама незгода.
7. У последњем кораку, места саобраћајних незгода се пореде са тзв. граничним вредностима. Места саобраћајних незгода на којима број саобраћајних незгода премашује граничне вредности сматрају се „црним тачкама“. Граничне вредности варирају у зависности од врсте пута (Табела 3), и односе се на период од две године.

Табела 3. Граничне вредности за одређивање „црних тачака“ [43]

| Врста пута и места | Граничне вредности | | |
|--|--------------------|---|-----------------------|
| | Број свих СН | Број СН са повређенима и/или погинулима | Број СН са погинулима |
| Ауто-путеви (по смеровима) | 10 | 4 | 2 |
| Уливне/изливне рампе | 10 | 4 | 2 |
| Ванградски путеви (раскрснице и отворени путеви) | 8 | 4 | 2 |
| Главни путеви у насељима – отворени путеви | 8 | 5 | 2 |
| Главни путеви у насељима – раскрснице | 10 | 6 | 2 |

3. Преглед досадашњих реализованих студија, пројеката и истраживања која су обрађивала опасна места – „црне тачке“ на путевима Републике Србије

У тексту који следи приказан је преглед низа реализованих студија, резултата истраживања и пројеката који су обрађивали опасна места, односно „црне тачке“, на територији Републике Србије (Табела 4).

Табела 4. Преглед реализованих студија, пројеката и истраживања у којима је обрађивана тема идентификације опасних места – „црних тачака“ на путевима Републике Србије [9]

| Р. бр. | Назив студије | Извођач и година | Обрађивана област | Метод идентификације | Коришћени ризици | Коришћени појмови |
|--------|---|--|--|---|---|--|
| 1 | „Идентификација опасних места на државним путевима Републике Србије са предлогом мера санације“ | ЈП „Путеви Србије“, МУП, Министарство за инфраструктуру, службе надзора и предузећа за путеве, 2007. | Управљање црним тачкама | СУБЈЕКТИВАН – анкетно истраживање | ризик добијен на основу анкетног истраживања | „црна тачка“ (тачка у геометријском смислу) „опасан потез“ (променљиве дужине) |
| 2 | „Идентификација опасних места и пуњење базе података о опасним местима на државним путевима“ | Криминалистичко-полицијска академија, 2011. | Управљање црним тачкама | ОБЈЕКТИВАН – број и последице саобраћајних незгода | ПБСН | „црна тачка“ (потез дужине до 300 м, изузетно до 1.000 м) |
| 3 | „Упоредна анализа опасних места и ризичних деоница на државним путевима“ | Криминалистичко-полицијска академија, 2010. | Управљање црним тачкама; Мапирање ризика | СУБЈЕКТИВАН и ОБЈЕКТИВАН | ризик добијен на основу анкетног истраживања К | „опасна деоница“ (саобраћајна деоница променљиве дужине, једнокилометарска деоница) „црна тачка“ (тачка у геометријском смислу) „опасан потез“ (променљиве дужине) |
| 4 | „Идентификација и класификација опасних места „црних тачака“ у саобраћају на подручју града Београда, са предлогом санационих мера“ | Криминалистичко-полицијска академија, 2010. | Управљање црним тачкама | ОБЈЕКТИВАН – број и последице саобраћајних незгода | ПБСН КР | „опасно место“ (дужине до 100 м) „опасна деоница“ (дужине веће од 100 м) |
| 5 | „Методологија идентификације и управљања опасним местима (црним тачкама) на државним путевима Републике Србије“ | Саобраћајни факултет, 2010. | Управљање црним тачкама | ОБЈЕКТИВАН – број и последице саобраћајних незгода | ПБСН | „црна тачка“ (раскрсница са прилазима у дужини од 50 м или део пута дужине 100 м) |

ПБСН – пондерисани број саобраћајних незгода
К – коефицијент колективног ризика на основу броја погинулих и тешко повређених
КР – колективни ризик настанка незгода

3.1. Идентификација опасних места на државним путевима Републике Србије са предлогом мера санације

Приказана студија датира из 2007. године, када је ЈП „Путеви Србије“, у сарадњи са стручним лицима МУП-а, Министарства за инфраструктуру, Службе надзора и предузећа за путеве, спровело истраживање које је имало за циљ да идентификује опасна места на државним путевима Републике Србије и рангира их по критеријуму приоритета њихове санације.

The screenshot shows a detailed Excel spreadsheet for a dangerous location. Key sections include:

- Podaci o lokaciji OMT:** PZP AD "Novi Pazar-put" - Novi Pazar, Oznaka puta i deonice M-5, Dužina poteza, Kilometar puta u okviru deonice 641-500 km, Lokacija (odlono) Uzlazna M5 sa lokalnim putem u Adranima, GPS oznaka (waypoint) 10-1, X, Y i Z koordinate 7470902, 4847000, 211.
- Područje:** Područje, Put van naselja, Put u pravcu, Upr. skretanje (na levo, na desno), Vpr. skretanje, Razvazanje (T, Y, o, x, ...), Most, Prizni prelaz, Vozni.
- Vrsta podloge:** Asfal, Ostale (tucanik, beton, ...).
- Stanje puta:** Dobro, Udobno (rupe), Kolotezi, Ostalo (nagib, prijava).
- Preglednost/vidljivost:** Dobra.
- Regulisanje saobraćaja/oprema puta:** Svetlosna signalizacija, Vertikalna i horizontalna sig., Oprema puta (del ogr. i smer.), Ograničenje brzine (do 300 km/h), Pesaki tokovi.
- Skica/fotografija:** A photograph of a road intersection with a sign.
- Opisni predlog mera:** 1. Formiranje trake za levo skretanje iz pravca Kraljeva Kraljeva, 2. Izmeštanje autobuskog stajališta iz raskrsnice, 3. U krivini postaviti usmerivače, 4. Apsolutno zaustavljanje, 5. Označiti nesigurnom signalizacijom.
- Uzroci (ime preuze i mob. tel.):** 1. MUJ, 2. Nadzori, 3. Isp. pešača i/ili, 4. PZP.

Слика 2. Приказ екрана програма Microsoft Excel са детаљима опасног места [50]

Након спроведеног анкетања, вршена је верификација уочених опасних места путем теренских истраживања, након чега је формирана јединствена база података о опасним местима на мрежи државних путева. Приликом креирања базе, коришћена је *GIS* технологија заснована на *GPS* координатама опасних места и пратећим табеларним и мултимедијалним садржајима (фотографије и видео записи). База података са детаљним подацима о опасним местима формирана је у програму *Microsoft Office Excel* (Слика 2), док се у програму *MapSource* може изабрати ниво детаљности приказа путне мреже.

3.2. Идентификација опасних места и пуњење базе података о опасним местима на државним путевима

Током 2011. године у Републици Србији је реализована студија (р. бр. 2 у Табели 4) која је имала циљ да идентификује опасна места на државним путевима Републике Србије и да буде формирана база података која ће омогућити квалитетно похрањивање података о опасним местима [51].

За избор опасних места у овој студији примењен је модел за идентификацију „црних тачака“ коришћењем пондерисаног броја саобраћајних незгода. Овај модел издваја најопасније делове пута на основу објективног ризика страдања израженог бројем и последицама саобраћајних незгода, при чему врши додељивање тежинских коефицијената, тзв. Пондера, у зависности од тежине (врсте) саобраћајне незгоде. Пондери се одређују на основу укупних друштвених трошкова које носе поједине врсте страдања. С обзиром да Република Србија још увек није дефинисала вредности пондера за поједине врсте страдања, то су за потребе ове студије преузете вредности пондера дефинисане од стране Министарства саобраћаја Велике Британије, на следећи начин [9]:

- саобраћајним незгодама са лаким телесним повредама додељује се вредност пондера 1,
- саобраћајним незгодама са тешким телесним повредама додељује се вредност пондера 13, и
- саобраћајним незгодама са погинулим лицима додељује се вредност пондера 99.

3.3. Упоредна анализа опасних места и ризичних деоница на државним путевима

Студија р. бр. 3 у Табели 4, која је реализована током 2010. године, примарно је обухватала упоредну анализу опасних места која су идентификована анкетним истраживањем (субјективни ризик страдања у саобраћају) и ризичних деоница које су идентификоване на основу броја и последица саобраћајних незгода (објективни ризик страдања у саобраћају) на државним путевима Републике Србије [52]. Студија је имала за циљ да изврши усаглашавање метода за идентификацију „црних тачака“ примењених у два независна истраживања:

- 1) „Идентификација опасних места на државним путевима Републике Србије са предлогом мера санације“, и
- 2) „Мапирање ризика на мрежи државних путева I реда“.

Прво истраживање спроведено је у оквиру дате студије из 2007. године, која је претходно представљена у подпоглављу 4.1. Методологија и неки од закључака овог истраживања су приказани у поменутој тачки, па овде није посебно обрађена. Друго истраживање спроведено је у оквиру истоимене студије из 2008. године, која је имала циљ да квантификује и представи ризике на државним путевима I реда (претходни назив: магистрални путеви) на подручју Србије.

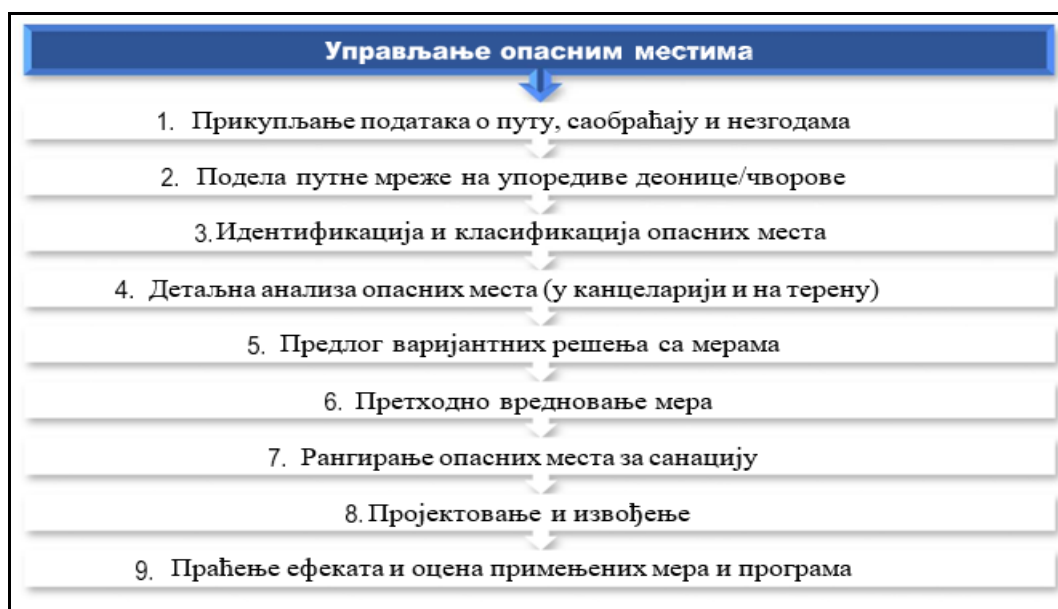
У овој студији као „црне тачке“ на путевима разматране су опасне деонице, при чему су посебно обухваћене опасне микродеонице (саобраћајне деонице променљиве дужине) и опасне једнокилометарске деонице (пут дужине једног километра).

Овом студијом идентификовано је 60 опасних места, од чега су 22 опасна места означена као приоритетна за санацију, док су 63 опасна места означена као места која не представљају објективно ризичне локације за учеснике у саобраћају [53].

3.4. Идентификација и класификација опасних места „црних тачака“ у саобраћају на подручју града Београда, са предлогом санационих мера

Студија р. бр. 4 у Табели 4, која је реализована 2010. године, представља

значајан корак ка развоју и дефинисању научно утемељеног метода идентификације и класификације „црних тачака“ на путевима. Студија је имала циљ да допринесе успостављању система управљања опасним местима – „црним тачкама“ у Београду и да омогући усмеравање средстава за санацију најугроженијих места. У студији је дефинисан алгоритам процеса управљања опасним местима, који се састоји од девет корака приказаних на Слици 3.



Слика 3. Алгоритам процеса управљања опасним местима у Београду

У алгоритму је наглашено да основу за квалитетно управљање „црним тачкама“ чине подаци о карактеристикама путне и уличне мреже, саобраћајном оптерећењу и саобраћајним незгодама. База података о саобраћајним незгодама је неопходан елемент за управљање „црним тачкама“. Такође, потребно је путну и уличну мрежу поделити на хомогене целине, односно, места са истим или сличним путним и саобраћајним условима (деонице сличних функција, чворови – раскрснице са сличним елементима, и сл.).

Трећи корак алгоритма је идентификација и класификација (рангирање) опасних места – „црних тачака“ према задатим критеријумима. У овој студији разматрана су два појма: опасна места (дужине до 100 m) и опасне деонице (дужине веће од 100 m), при чему су оба појма дефинисана у односу на број и тежину саобраћајних незгода. За потребе идентификације опасних места развијен је модел заснован на броју и последицама саобраћајних незгода, при чему је свакој саобраћајној незгоди додељен тежински индекс у зависности од врсте последица, и на тај начин за сваку од локација нагомилавања незгода

израчунат је пондерисани број саобраћајних незгода (ПБСН).

Индекси који су коришћени засновани су на трошковима саобраћајних незгода, и предложени од стране Међународног удружења путара [49]. Саобраћајне незгоде мултипликоване су на следећи начин:

- саобраћајне незгоде са лаким телесним повредама – тежинским индексом 1,
- саобраћајне незгоде са тешким телесним повредама – тежинским индексом 10, и
- саобраћајне незгоде са погинулим лицима – тежинским индексом 85.

Идентификована опасна места су рангирана према вредности показатеља ПБСН.

3.5. Методологија идентификације и управљања опасним местима (црним тачкама) на државним путевима Републике Србије

Студија р. бр. 5 у Табели 4 датира из 2010. године и представља још једну у низу студија које су имале за циљ да успоставе систем управљања „црним тачкама“ на државним путевима Републике Србије. У овој студији детаљно је сагледано и анализирано стање и тенденције у дефинисању и идентификацији „црних тачака“ у свету, али и анализирана правна регулатива у области безбедности саобраћаја у Србији и оцењене досадашње активности у вези са управљањем „црним тачкама“.

Важан резултат студије представља формирање Методолошког упутства о идентификацији и управљању „црним тачкама“, као и предлог конкретних активности које је потребно предузети у циљу увођења и практичне примене предложене методологије у Републици Србији.

3.6. Рекапитулација

Досадашњи развој научне мисли и најновија истраживања идентификације „црних тачака“ на путевима показали су да је емпиријски Бајесов метод најбољи приступ са теоријског становишта, односно, овај метод издваја се као најсавршенији приступ [33][54][32] [28][31][49]. Остали препоручени приступи спадају у смернице најбоље праксе, што је вероватно најважније за

реализацију овог пројекта.

Анализа досадашњих искустава различитих држава у идентификацији и класификацији „црних тачака“ на путевима (Аустрија, Данска, Немачка, Белгија, Норвешка, Португал, Швајцарска, Мађарска) показала је да су у праксама анализираних држава развијени индивидуални приступи у идентификацији опасних места – „црних тачака“, који се базирају на објективном методу идентификације. У поређењу са најновијим техникама за идентификацију опасних места, већина практичних приступа који се користе у посматраним државама су засновани на евидентираном броју саобраћајних незгода, иако теоретски најбољи приступ подразумева да би „црне тачке“ требало идентификовати на основу очекиваног броја саобраћајних незгода [32]. Једно од решења које се препоручује је и принцип идентификације заснован на традиционалном моделу у коме се евидентирани број саобраћајних незгода пореди са општим очекиваним бројем незгода [9].

Идентификација опасних места – „црних тачака“ на путевима у Републици Србији је још увек у почетној фази развоја, зато још увек не постоји национална дефиниција опасног места – „црне тачке“, као што је то случај у горе поменутих државама [55].

Последњих година покренуте су одређене активности, пре свега управљача путева, на развоју методологије за идентификацију „црних тачака“, али оне још увек нису у потпуности усаглашене тако да се могу применити на читавој путној мрежи Републике Србије.

4. Интегрисани системи за управљање и анализу података

У раду је коришћен јединствени интегрисани систем за управљање и анализу података који омогућава повезивање различитих извора података. У конкретном случају, коришћени су јавно доступни подаци са портала Отворених података, њихово форматирање и централно складиштење ради ефикаснијег управљања и анализе. Коришћењем система Пословне интелигенције и складишта података, наведени подаци су ефикасно припремљени за коришћење у оквиру модификованог алгорита.

4.2. Отворени подаци

По дефиницији, податак је било која чињеница, број или текст који може бити обрађен и ажуриран током пословне обраде. Данас, организације нагомилавају огромне количине података у различитим форматима и различитим базама. Ово укључује [56]:

- оперативне и трансакционе податке као што су подаци о продаји, трошковима, наплатама, рачуноводству итд.,
- неоперативне податке, као што је индустријска продаја, предвиђање података и макроекономски подаци,
- мета податке, тј. податке о подацима, као што је логички дизајн базе или речник базе података.

Подаци се преплићу кроз све аспекте савременог живота. Друштво их ствара у просвети, здравству, транспорту, бизнису и многим другим аспектима. Јавне институције историјски производе и обрађују највише скупова података из различитих области, од праћења саобраћајних незгода и квалитета ваздуха, до вођења статистике о привреди или образовању грађана.

У циљу проналажења заједничког одговора на глобалне проблеме и ширења знања које ови подаци носе са собом, настала је идеја да би подаци које прикупљају јавне институције требало да буду заједничко добро које свима припада. Да би могли сви да их користе, неопходно је да буду доступни у форми отворених података. Отворени подаци су јавно доступни, машински читљиви, тј. припремљени су за обраду од стране рачунара и слободни су за употребу у било којој сврхе, било комерцијалне, било некомерцијалне [57].

Уз нове технологије, отворени подаци су глобално постали кључни ресурс

који омогућава бројне користи за саме јавне институције, привреду, академски сектор, цивилно друштво, медије и све грађане.

Уз отворене податке, јавна управа постаје транспарентнија, при чему прима мање захтева за приступ информацијама од јавног значаја, па може да доноси боље и ефикасније одлуке, за квалитетнији живот грађана, може да понуди иновативне услуге и скрати време неопходно за размену информација међу институцијама.

Са друге стране, отворени подаци могу да буду ресурс за привреду, јер омогућавају стварање иновативних услуга, стварање нових компанија, изједначавање знања доступног малим и великим компанијама на тржишту, као и већу конкуренцију.

Користи отворених података очигледне су и у свакодневним примерима. Тако, нпр. професори географије могу користити отворене мапе, за потребе својих предавања. Новинари могу продубити своја истраживања и поткрепити своје текстове већим бројем података.

Република Србија је 2015. године имала пет скупова отворених података. Данас их има више од 1.800. Отворени су подаци о квалитету ваздуха, које се мере сензорима на различитим локацијама, наплати пореза у локалним самоуправама, саобраћајним незгодама на нивоу свих општина, нивоу електромагнетног поља на 26 локација широм Србије итд.

Закон о електронској управи из 2018. године увео је право грађана да захтевају податке у отвореној форми, како би могли да их употребе у друге сврхе. Отворени подаци су јавно добро које је доступно свима.

4.2. Пословна интелигенција

Складиштење података и пословна интелигенција су технике које обезбеђују пословним људима информације и алате који су им потребни за доношење одлука о оперативном и стратешком пословању [58].

Корисници су углавном пословни људи у компанијама. Међутим, немају сви пословни људи исти „значај“ при доношењу пословних одлука. Потребно је посебно се побринути за оне особе које доносе стратешке пословне одлуке. Добра пословна одлука може донети велике бенефите компанијама [59]. Зато су кључни корисници руководиоци, менаџери и аналитичари у компанијама. Складиште података и пословна интелигенција (*DW/BI*) су системи високог утицаја и значаја у организацији.

Интелигенција представља менталну особину коју чини неколико различитих способности: учење из искуства, прилагођавање новим ситуацијама и њихово разумевање, као и способност да се у новим околностима примене претходно стечена знања, те да сналажење у њима буде брзо и учинковито. Интелигенција подразумева и способност разумевања и коришћења апстрактних појмова и идеја, односно апстрактно мишљење, те могућност брзог и лаког учења и савладавања задатих проблема. По дефиницији *Речника српског језика* Матице Српске, интелигенција је “способност брзог схватања и разумевања, умна обдареност, бистрина, разум”. “Интелигенција је способност прилагођавања променама”, рекао је чувени енглески астрофизичар Стивен Хокинг, и сам један од најинтелигентнијих људи на свету, док немачки писац Бертолд Брехт сматра да “интелигенција није у томе да не правимо грешке, већ да што пре увидимо како да их претворимо у нешто добро [60]”.

Сама реч *интелигенција* у термину пословна интелигенција, изворно означава извештавање или обавештавање намењено подршци пословном одлучивању. Овако конципирана, интелигенција поспешује сналажење у новим ситуацијама и брзину и тачност решавања проблема, при чему се од психолошке употребе термина интелигенција разликује по томе што се не односи на способност личности којом се служи појединац, већ на технолошко решење које користи

одређени систем. Пословна интелигенција није ни производ, ни систем, већ архитектура и колекција интегрисаних оперативних апликација и апликација за потпору одлучивању [61]. Она представља коришћење свих потенцијала података, самим тим и информација и знања у одређеном систему ради доношења бољих и исправних пословних одлука и у складу с тим, идентификацију нових пословних могућности. Просто прикупљање података и недавање пажње анализи истих, доводи до нагомилавања података, који немају смисла. Компаније се гуше у подацима, а жедне су знања [62]. Информациона основа на којој почива пословна интелигенција то себи никако не сме да дозволи. Код ње је пресудан квалитет податка који је ушао у меморију без обзира да ли је она научна, техничка, социјална, економска, демографска, политичка, или нека друга сфера [63]. Информатички подржану пословну интелигенцију могуће је замислити као огромну машину за сортирање, анализирање, критиковање, оцењивање и синтезу необрађених података, који на крају тих процеса постају сама суштина људског знања [64].

Покушај примене трансакционих база података за доношење пословних одлука показао се као веома спор, обзиром да су сами упити веома сложени и споро се извршавају, тако да менаџери не могу добити тренутне информације које су потребне за доношење пословних одлука [65].

Полазећи од тумачења појма система као скупа објеката-ентитета и њихових међусобних веза усмерених ка остваривању заједничког циља, информациони систем (ИС) се може дефинисати као систем у коме се везе између објеката и везе система са околином остварују разменом информација [62]. У ширем смислу информациони систем представља уређени и интегрисани скуп података, процеса, интерфејса, мрежа, технологија и људи који су у међусобној корелацији у циљу подршке и побољшања свакодневних пословних операција и подршке менаџменту у решавању пословних проблема, планирања, управљања, предвиђања, координирања и доношења одлука [66].

Имплементација *BI* техника повећава употребну вредност постојећег информационог система организације, услед чега је интересовање за *BI* системе велико и још увек расте [67].

Обзиром да не постоји универзална дефиниција појма *Business Intelligence*, различити аутори га дефинишу на различите начине. Једна од најчешће коришћених и општијих дефиниција је следећа: “*Business Intelligence* је такво коришћење података које води ка доношењу бољих пословних одлука. Оно се односи на приступ, анализу и откривање нових могућности” [68].

Неки аутори *Business Intelligence* дефинишу на следећи начин [68]:

- У [69], аутори дефинишу: “*Business Intelligence* није ни производ нити систем. Она представља архитектуру и колекцију интегрисаних оперативних апликација и апликација за подршку одлучивању као и базе података које пословној организацији омогућава лак приступ подацима.”
- У [70], аутори дефинишу: “Претварање података у знање је задатак апликација познатих под називом *Business Intelligence*. *Business Intelligence* је скупина нових апликација обликованих тако да могу организовати и структуриране податке о пословним трансакцијама на начин који омогућава анализу корисну у подршци одлучивању и оперативним активностима организације.”
- Аутор у [71] наводи: “Када компанија користи податке из својих извршних система при доношењу одлука, то је *Business Intelligence*. *Business Intelligence* системи преузимају податке из извршних система и реструктурирају их тако да корисници могу откривати раније непознате информације о предузећу, пословању, клијентима, пословним активностима или производима.”
- Према аутору у [72]: “*Business Intelligence* је процес прикупљања расположивих интерних и релевантних екстерних података, и њихова обрада у корисне информације које могу помоћи пословним корисницима при доношењу одлука.”
- „Разматрано с макроспекта, пословна интелигенција је сложена, агрегирана категорија која се ствара систематичним, унапред циљаним прикупљањем података о макроекономским кретањима у одређеној геополитичкој средини, њиховим организовањем и структурираним

бележењем односно похрањивањем, претраживањем, као и логичком и рачунарском обрадом у циљу откривања макроекономских трендова или тенденција ради предвиђања и прогнозирања процеса и догађаја.“[73]

- Пословна интелигенција представља коришћење свих потенцијала података и информација у предузећу ради доношења бољих пословних одлука и у складу с тим идентификацију нових пословних могућности [74].
- Пословна интелигенција помаже корисницима да се одлуке које доносе више ослањају на информације и користи се, између осталог, за оптимизацију пословних планова и пословних процеса, као и за предвиђање потенцијалних проблема [75].

Најцелисходнија дефиниција пословне интелигенције и једна од данас најцитиранијих је дефиниција Хауарда Дреснера, аналитичара Gartner групе: „Пословна интелигенција је процес трансформације података у информације и кроз процес сазнања - трансформација тих информација у знање“ [76].

Дреснерова дефиниција пословну интелигенцију одређује користећи три кључна појма: податак, информација и знање.

Податак је било која чињеница, број или текст који може бити обрађен и ажуриран током пословне обраде. Под подацима се подразумевају детаљи о разним стварима, догађајима и трансакцијама, који су структурирани, класификовани и меморисани, али најчешће нису организовани на начин да имају неко конкретно значење. Подаци могу бити нумеричког или алфанумеричког типа, фигуре, звукови или слике. У овој дисертацији коришћени су нумерички и алфанумерички подаци меморисани кроз разне елементе пословног система организације.

Информације су подаци који су организовани тако да имају конкретно значење за оног ко их прима. Прималац информације интерпретира то значење и из њих изводи закључке. Да би подаци неко значење за одређену акцију или одлуку, они се прво морају обработити. Обрађени подаци могу имати значаја у процесу одлучивања.

Патерни, асоцијације или односи између свих тих података могу произвести информације. На пример, анализа података о продаји може произвести

информације на основу којих се може утврдити који производи се продају, под којим условима и када.

Информације су кључни фактор у доношењу одлука на било ком нивоу, почев од почетног преко оперативног и тактичког, до стратешког нивоа одлучивања. С обзиром на важност сваке одлуке, информације на бази којих се оне доносе морају бити потпуне, правовремене и прецизне. Међутим, пут од података до информација и до могућности предвиђања будућих пословних потеза (односно откривање знања у тим подацима) није нимало лак.

Знање се састоји од података организованих на начин да преносе искуство, акумулирано учење и експертско знање, а у вези проблема који треба разрешити. Знање представља скуп података процесираних са циљем да извуку критичне импликације и рефлектују искуства из прошлости.

Свака организација поседује податке о свом пословању најчешће у разним врстама информационих система, документима, пословним књигама и на другим носиоцима података. Процесирање тих података, проналажење скривених релација међу њима, идентификација нових могућности за боље пословање, односно трансформација података у информације и знање, представљају значајан развојни потенцијал сваке организације.

На Слици 4, приказан је концепт тумачења података, превођења у информације и добијање и долазак до знања.



Слика 4. Веза између контекста и разумевања знања

Количина података који настају одвијањем пословних догађаја и реализацијом пословних процеса у предузећу пропорционална је бројности тих догађаја и сложености процеса. Међутим, између количине података и њиховог квалитета не мора увек постојати директна пропорционалност. Велика количина података не мора уједно значити висок квалитет, вредност и употребљивост тих података у процесима одлучивања. Ово наводи на потребу за контролисањем квалитета података, односно управљање системом у циљу прикупљања што квалитетнијих података.

Обједињујући у себи концепте складишта података, *OLAP* и *Data mining*, *BI* концепт задовољава потребу менаџмента организација за информацијама на свим нивоима управљања: стратегијском, оперативном и тактичком нивоу. За ниво стратегијског управљања неопходне су информације више општости, у виду збирних прегледа и историјског карактера. Насупрот томе, за ниво тактичког управљања потребне су специфичне информације уског домена, претежно у виду „сирових“ података, врло актуелних, углавном квантитативне природе и велике прецизности и тачности [77].

Развој информационих система и методологија развоја *BI* система обухвата следеће фазе: дефинисање захтева, анализа, дизајн и имплементација [78], а централни артефакт је модел и моделовање. Управо те кораке препознају, и на њима инсистирају аутори у [69] и *ORACLE* [79].

Дефинисањем захтева треба сагледати систем и уочити потребе корисника. Дефинисањем захтева обавља се и идентификација система. У овом кораку ће се паралелно користити класичне методе и објектне применом *UML*-а [73].

Анализа треба да омогући, за претходно дефинисане пословне активности, израду предлога решења система и као крајњи циљ има опис појединих функционалности кроз израду дијаграма интеракције. У овом кораку *UML* ће имати примат [80].

Дизајн формира коначну структуру будућег система кроз дефинисање модела класа, базе података (применом *UML*-а) и припрему података за учитавање у будући систем.

Реализација представља последњи корак постављене методологије и круну целокупног процеса. Конкретне “последнице“ овог корака су складиште података, упити, *OLAP* коцке и *Data Mining* алгоритми реализовани кроз конкретне апликације корисничког интерфејса [81].

Анализа података омогућава приступ формирању мета података. Мета подаци су подаци о подацима [80]. Помоћу њих се идентификује садржај и локације података у складишту података. Мета подаци представљају и везу складишта података и апликације која складиште користи.

4.3. Складишта података

Централно место целе стратегије пословне интелигенције припада складишту података. Складиште података је посебно дизајнирана, аналитичка база података у којој се чувају подаци прикупљени из интерних или екстерних извора, и која омогућава сложене, унапред непредвиђене (*ad-hoc*, енгл.) приступе великом броју различитих података [82].

Складиште података је специфична база података намењена подршци одлучивању у одређеној организацији. Бил Инмон је дефинисао складиште података као субјектно оријентисану, интегрисану, временски зависну и непроменљиву колекцију података намењену подршци одлучивању у одређеној организацији [83]. Подаци у складиштима података су непроменљиви, тј. чим се неки податак упише у складиште података, могуће му је само приступити [84]. Ралф Кимбал дефинише складиште података као копију трансакционих података структурисану тако да омогућава задавање упита и анализу података [85].

Анализирање шаблона података и трендова захтева постојање великог броја историјских података. Из тог разлога аналитичке базе података не садрже ажурне податке, већ чувају информације из одређеног тренутка времена [86].

Циљ складиштења података није да се подаци искључиво складиште, већ и да пословни менаџери могу сами да врше анализе. Чак и у свету успешних менаџера постоје људи који нису довољно технички образовани, а имају потребу за информацијама, а да при томе не знају да програмирају. Такође, коришћењем складишта података компаније желе да се што више приближе клијентима, не би

ли боље разумели њихове захтеве и потребе. Ово су још неки од разлога за увођење складишта података [87]:

1. менаџерима су потребни брзи одговори на многа питања да би остали конкурентни на тржишту;
2. аналитичарима су потребне кориснички оријентисане (*user-friendly*, енгл.) апликације које ће им давати одговоре, тако да они не морају да траже помоћ од других људи;
3. пословни системи подршке одлучивању морају приступати подацима из различитих извора и интегрисати их у једну целину;
4. системи за подршку одлучивању морају бити оптимизовани за давање одговора, тако да се лако и брзо може спроводити интерактивна анализа и извештавање;
5. извршиоци и доносиоци одлука су покретни, тј. они имају потребу за постојањем мобилних система за подршку одлучивању;
6. мора постојати могућност временске анализе, анализе промена и анализе трендова.

Чињеница је фокус процеса доношења одлука. Она обично представља скуп догађаја који се појављују у компанији. Примери чињеница у домену саобраћајних незгода су: учесници незгоде, време догађаја, локација, итд. Од суштинске важности за чињеницу је да има одређене динамичке аспекте, односно, да се на неки начин развија кроз време. Концепти који се у извору података често ажурирају добри су кандидати за чињенице, док они који су приближно статички, нису. Избор чињеница треба да се заснива или на просечној учесталости промена или на специфичним интересима анализе [88].

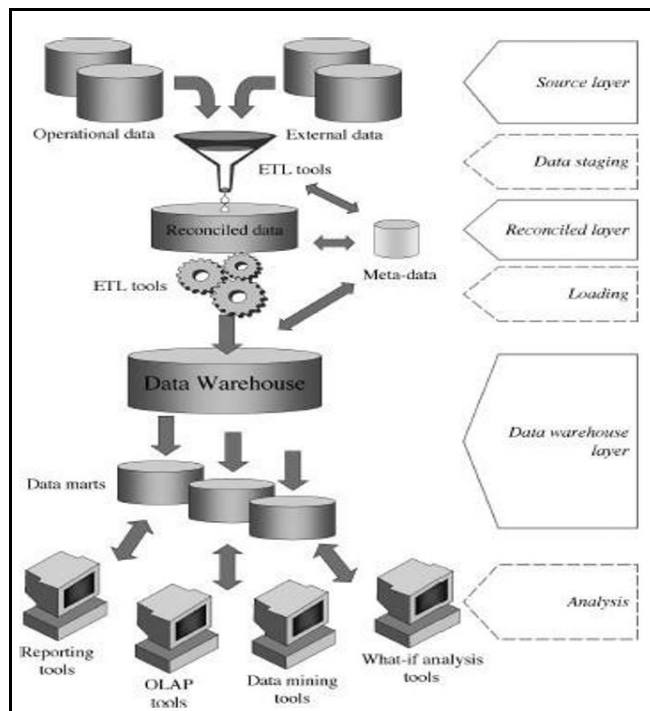
Мера је нумеричко својство чињенице и описује један од њених квантитативних аспеката који су од интереса за анализу.

Димензија је својство чињенице са коначним доменом и описује једну од њених координата анализе.

У раду је коришћена трослојна архитектура складишта података. У овој архитектури, трећи слој је тзв. „компромисни“ слој податка или оперативно складиште података. Овај слој материјализује оперативне податке добијене након

интегрисања из одређеног извора података и накнадног прочишћавања. Као резултат подаци су интегрисани, конзистентни, исправни, ажурни и детаљни. Слика 5 приказује складиште података које није попуњено из својих извора директно, већ из оперативног складишта података [89].

Главна предност “компромисног“ слоја података је то што ствара заједничку референцу модела података на нивоу целог предузећа. У неким случајевима, компромисни слој се директно користи да би били лакше обављени неки оперативни задаци, као што су креирање дневних извештаја који се не могу задовољавајуће припремити помоћи корпоративних апликација. Компромисни подаци доводе до веће редувантности оперативних извора података. Потребно је знати да двослојне архитектуре могу имати компромисни слој који није посебно материјализован, већ само виртуелан, јер се дефинише као конзистентно интегрисани поглед оперативних извора података.



Слика 5. Трослојна архитектура складишта података [90]

Чишћење је најзначајнија фаза у складиштењу података, обзиром да се на овај начин побољшава квалитет података који је прилично низак у изворима из којих се подаци прибављају [91].

Главне карактеристике чишћења података које се одвија у *ETL* алатима су исправљање грешака и хомогенизација. Ови алати користе специфичне речнике за исправљање грешака у куцању, као и за препознавање синонима.

Учитавање у складиште података је последњи корак. Оно се може обавити на два начина:

- *Refresh*, што подразумева комплетно „преписивање“ складишта података, што значи да су старији подаци замењени. Освежавање се обично користи у комбинацији са статичком екстракцијом приликом иницијалног попуњавања складишта података [92];
- *Update*, што подразумева само промене примењене на изворним подацима који су већ у складишту података. Измена се обично спроводи без брисања или модификације постојећих података. Ова техника се користи у комбинацији са инкременталном екстракцијом, за редовно ажурирање складишта података [93].

У раду је коришћен *Update*, где се подаци према потреби стално ажурирају са портала отворених података. Имплементирано складиште података је имплементирано у димензионом облику, који се састоји од табела чињеница и димензионих табела с којима су оне повезане, према [94].

5. Модификовани *Density-Based Spatial Clustering of Applications with Noise* алгоритам

Алгоритам представља јасно и децидно програмирани софтвер, који покреће листу прецизно дефинисаних инструкција које служе за прорачуне или решавање проблема. Алгоритми раде по стриктно предефинисаним и програмираним упутствима.

Најважнија карактеристика алгоритама је аутоматизација. Сврха алгорита је да извршава функције које, као активности, изводи човек, али на бржи, прецизнији и поузданији начин.

Кластеризација заснована на густини са шумом (*DBSCAN*) је метода машинског учења заснована на густини која групише податке према њиховој густини, проналази кластере неправилних облика и величина, и детектује податке који представљају шум. За разлику од традиционалних метода, *DBSCAN* не захтева од корисника да унапред одреди број кластера, што га чини погодним за анализу геопросторних података где структура података није унапред дефинисана. Међутим, традиционални *DBSCAN* третира све податке једнако, што ограничава његову ефикасност у ситуацијама где се тачкама може доделити и одређена вредност, тј. значај. Пондерисани алгоритми засновани на густини могу превазићи ова ограничења тако што податке вреднују на основу тежине и учесталости.

Аутори [95] у раду *Recent Advances in Traffic Accident Analysis and Prediction: A Comprehensive Review of ML Techniques* систематски анализирају примену различитих техника машинског учења у предикцији саобраћајних незгода, истичући да модели као што су *Random Forest*, *SVM* и неуронске мреже омогућавају боље разумевање образаца ризика и предикцију учесталости и тежине несрећа.

Аутори [96] у раду *AI models for road safety: prediction of crash frequency and severity* развијају моделе засноване на вештачкој интелигенцији који комбинују статистичке методе и дубоко учење ради предвиђања учесталости и тежине саобраћајних незгода, чиме се омогућава креирање превентивних стратегија за управљање безбедношћу на путевима.

Аутори [97] у раду *Road Crash Analysis and Modeling: A Systematic*

Review of Methods дају систематски преглед метода анализе саобраћајних незгода, наглашавајући примену машинског учења у моделирању ризика, класификацији фактора који доводе до несрећа и процени вероватноће њиховог настанка.

Аутори [98] у раду *Deep learning for traffic accident severity prediction* примењују дубоке неуронске мреже за предвиђање тежине саобраћајних незгода, користећи комплексне улазне податке као што су карактеристике пута, услови вожње и понашање возача, чиме се постиже висока тачност у процени исхода несрећа.

Аутори [99] у раду *Machine learning-based crash risk assessment using driver behavior data* фокусирају се на анализу понашања возача кроз податке о брзини, маневрима и реакцијама, користећи моделе машинског учења за процену ризика од несрећа у реалном времену и идентификацију критичних образаца који повећавају вероватноћу удеса.

Аутори [100] у раду *Comprehensive analysis of clustering algorithms: exploring limitations and innovative solutions* анализирају *DBSCAN* алгоритам, истичући његове предности у откривању кластера произвољног облика и отпорности на шум, али и ограничења која се односе на осетљивост параметара ϵ и *minPts*, као и проблеме при кластеровању података различите густине. Као иновативна решења представљене су варијанте попут *HDBSCAN*-а, који омогућава адаптивно откривање кластера различите густине, *MW-DBSCAN*-а који уводи тежине ради боље обраде хетерогених података, као и приступи који комбинују *DBSCAN* са методама дубоког учења ради примене на сложене, високо-димензионалне скупове података.

Аутори [101] у раду *A Rapid Review of Clustering Algorithms* анализирају *DBSCAN* алгоритам, истичући његову способност откривања кластера произвољног облика и отпорност на шум, али указују и на ограничења која се односе на осетљивост параметара ϵ и *minPts*, као и проблеме при кластеровању података различите густине. Као решења ових недостатака представљене су иновативне варијанте попут *HDBSCAN*-а, који омогућава адаптивно откривање кластера различите густине, и тежинских модификација *DBSCAN*-а.

Аутори [102] у раду *Recent innovations in deep embedded clustering and spectral clustering* усмеравају пажњу на савремене модификације класичних

метода кластеровања, са посебним акцентом на дубоко уграђено кластеровање и спектралне приступе. У раду се истиче да дубоко уграђено кластеровање комбинује репрезентационо учење са процесом кластеровања, чиме се постиже боља обрада високо-димензионалних података, док спектрални алгоритми користе графове и Лапласијане ради откривања сложених структура које традиционалне методе не могу да препознају. Аутори наглашавају да ове иновације ублажавају ограничења класичних алгоритама, као што су осетљивост на параметре, проблеми са различитим густинама и ограничена скалабилност, и отварају могућности за примену у областима попут анализе слика, биоинформатике и друштвених мрежа.

Аутори [103] у раду *Advances in hierarchical and graph-based clustering* разматрају *DBSCAN* алгоритам у контексту поређења са хијерархијским и граф-базираним методама кластеровања. Истичу да *DBSCAN* има предност у откривању кластера произвољног облика и отпорности на шум, али указују на ограничења која се односе на осетљивост параметара ϵ и *minPts*, као и на проблеме при кластеровању података различите густине. У раду се наглашава да новије хијерархијске и граф-базираних технике могу да ублаже ова ограничења кроз адаптивно одређивање густине и боље моделирање односа између тачака, чиме се *DBSCAN* позиционира као полазна тачка за развој хибридних приступа кластеровању.

Аутори [104] у раду *Hybrid clustering approaches for large-scale data mining* разматрају *DBSCAN* алгоритам као део хибридних техника кластеровања које се примењују на велике скупове података. Истичу да *DBSCAN* има предност у откривању кластера произвољног облика и отпорности на шум, али наглашавају и његова ограничења, пре свега осетљивост на параметре ϵ и *minPts* и проблеме при кластеровању података различите густине. У раду се показује да интеграција *DBSCAN*-а са хијерархијским и граф-базираним методама, као и са техникама дубоког учења, може да ублажи ова ограничења и омогући ефикасније кластеровање у условима велике димензионалности и комплексности података.

У овом раду представљен је модификовани пондерисани *DBSCAN* алгоритам (*MW-DBSCAN*), осмишљен да побољша тачност кластеризације. Традиционални *DBSCAN* све тачке третира равномерно, док *MW-DBSCAN* у

процес кластеризације укључује тежину и учесталост тачака додељивањем пондера појединачним записима. Коришћењем нових шема проширења радијуса када се кластери не преклапају и интеграцијом пондера у процес кластеризације, *MW-DBSCAN* побољшава способност алгоритма да прецизније детектује кластере. Експериментални резултати показују да предложено решење постиже већу тачност у односу на традиционални *DBSCAN* алгоритам, нарочито у хетерогеним окружењима.

Аутори у [105] представили су *DBSCAN* алгоритам кластерованја који се примењује на просторне податке. Он представља алгоритам кластерованја који кластере одређује на основу густине распореда елемената скупа података у простору. *DBSCAN* алгоритам идентификује кластере груписањем тачака унутар одређеног радијуса [105]. Алгоритам лоцира регионе високе густине и детектује издвојене тачке. Алгоритам се ослања на два параметра: растојање између суседних тачака ϵ и минималан број тачака потребан да би се формирао густ регион, *MinPts*. На основу ϵ и *MinPts*, тачке се могу класификовати као језгро (*core*), граница (*boundary*) или шум (*outliers*). Језгро је тачка података која има најмање *MinPts* у свом суседству. Гранична тачка је дефинисана као она која има мање од *MinPts* у свом суседству, али је ипак достижна из језгра. Тачка која није језгро и није достижна из било ког језгра сматра се издвојеном тачком (*outlier*). *DBSCAN* је поступак корак по корак који подразумева избор почетне тачке, испитивање њеног окружења, проширење кластера ако је тачка језгро, и итерацију све док се све тачке не обраде. Ако је почетна тачка језгро, алгоритам рекурзивно процењује све тачке достижне густином како би формирао кластер, док граничне тачке не омогућавају даље проширење. Аутори [105] су предложили да се ϵ најбоље одређује емпиријски када је *MinPts* једнак четири.

Главне предности *DBSCAN* алгоритма су робусност у односу на шум, флексибилност у препознавању различитих облика кластера и независност од унапред одређеног броја кластера. Међутим, због високе осетљивости на избор параметара, *DBSCAN* може захтевати значајно време за обраду великих скупова података. *DBSCAN* се показао ефикасним у детекцији аномалија, кластеризацији радарских рефлексија, здравству, сегментацији купаца, геопросторној анализи и другим применама.

Аутори [106] у раду *Faster DBSCAN via subsampled similarity queries* представљају варијанту *DBSCAN* алгоритма која користи подузорковање упита сличности ради убрзања процеса кластеровања. Овај приступ омогућава значајно смањење рачунарских трошкова и времена извршавања, а да се при томе задржи висока тачност резултата. Аутори показују да техника подузорковања омогућава ефикасније руковање великим и високо-димензионалним скуповима података, чиме се *DBSCAN* чини применљивијим у условима где класична верзија алгоритма постаје прескупа за употребу.

Аутори [107] у раду *DBSCAN Clustering Algorithm Based on Density* представљају варијанту *DBSCAN*-а која се заснива на прецизнијем дефинисању густине, са циљем да се побољша тачност кластеровања и смањи осетљивост на параметре ϵ и $minPts$.

Аутори [108] у раду *K-DBSCAN: An improved DBSCAN algorithm for big data* предлажу побољшану верзију *DBSCAN*-а прилагођену великим скуповима података, која користи *K-means* иницијализацију ради ефикаснијег руковања високо-димензионалним и масивним подацима.

Аутори [109] у раду *DBSCAN++: Efficient density-based clustering for streaming data* развијају *DBSCAN++* алгоритам који омогућава кластеровање података у реалном времену, прилагођено стриминг окружењима, уз значајно убрзање процеса и боље управљање динамичким променама у подацима.

Аутори [110] у раду *Weighted DBSCAN for imbalanced datasets* уводе тежинску модификацију *DBSCAN*-а, којом се различитим тачкама додељује различита важност, чиме се побољшава кластеровање у условима неуравнотежених скупова података и омогућава боља идентификација мањинских класа.

Аутори [111] у раду *Hybrid clustering for transportation safety data* примењују *DBSCAN* у комбинацији са хијерархијским техникама кластеровања, како би идентификовали ризичне зоне у саобраћају и побољшали прецизност откривања „црних тачака“ кроз интеграцију различитих густина података.

Аутори [112] у раду *Real-time accident hotspot detection using DBSCAN variants (ACM)* развијају варијанте *DBSCAN*-а прилагођене стриминг окружењима, омогућавајући откривање саобраћајних „црних тачака“ у реалном времену кроз анализу долазних података о незгодама и динамичко ажурирање

ризичних локација.

Аутори [113] у раду *Integration of GIS and ML for black spot identification* комбинују *DBSCAN* са ГИС алатима и техникама машинског учења, чиме се црне тачке идентификују кроз просторне кластере незгода и предиктивне моделе који узимају у обзир инфраструктурне и демографске факторе.

Аутори [114] у раду *Predictive modeling of accident severity using ensemble learning* користе *DBSCAN* као иницијалну фазу просторног кластерованја, након чега се примењују ансамбл модели ради предикције тежине незгода у идентификованим ризичним зонама.

Аутори [115] у раду *Comparative study of clustering methods in traffic safety analysis* упоређују *DBSCAN* са другим методама кластерованја, показујући да је *DBSCAN* ефикасан у откривању „црних тачака“ саобраћаја, нарочито у условима где подаци садрже шум и кластере различитих облика.

Аутори [116] у раду *Adaptive DBSCAN for spatio-temporal accident data* развијају адаптивну верзију *DBSCAN*-а која узима у обзир просторне и временске димензије података, омогућавајући идентификацију динамичких „црних тачака“ које се мењају током дана или сезона.

Аутори [117] у раду *Graph-based clustering for traffic incident prediction* комбинују *DBSCAN* са граф-базираним техникама, чиме се црне тачке идентификују кроз анализу односа између локација незгода и њихових повезаности у транспортној мрежи.

Аутори [118] у раду *Empirical Bayesian models for black spot classification* интегришу резултате добијене *DBSCAN* кластерованјем са бајесовским моделима, чиме се црне тачке класификују на основу вероватноће ризика и историјских података о незгодама.

Аутори [119] у раду *Crash frequency prediction using random forest and gradient boosting* користе *DBSCAN* за иницијално просторно груписање незгода, а затим примену ансамбл техника ради предвиђања учесталости несрећа у идентификованим кластерима.

Аутори [120] у раду *Multi-weight DBSCAN applied to traffic accident datasets* примењују *MW-DBSCAN*, варијанту која уводи тежине ради боље обраде хетерогених података, омогућавајући прецизнију идентификацију „црних тачака“ у саобраћајним скуповима података са неуравнотеженим

дистрибуцијама.

Предложени модификовани алгоритам кластеризације засновани на пондерисаној густини формира кластере који боље одговарају условима из стварног света када се појединачним подацима додељују тежине или оцене значаја. Предложени алгоритам проширује традиционалне приступе засноване на густини увођењем пондера у процес кластеризације.

За разлику од *DBSCAN* алгоритма, који групише тачке искључиво на основу густине, предложени модификовани *DBSCAN* алгоритам додељује тежине које одређују значај сваке тачке. Уместо да броји колико се тачака налази у њиховом окружењу, алгоритам сабира њихове пондере. Сходно томе, мањи број важнијих тачака може формирати гушћи кластер него велики број мање значајних тачака.

Модификовани алгоритам обухвата поступак који се изводи корак по корак и функционише на следећи начин:

Корак 1: Тачкама се додељују пондери који одражавају њихов значај и утицај пре кластеризације,

Корак 2: Иницијализују се ϵ и *MinPts*,

Корак 3: Проверава се суседство сваке тачке (црта се круг са радијусом ϵ и траже се све тачке унутар њега),

Корак 4: Пондери се сабирају (уместо бројања суседа),

Корак 5: Одређује се тип тачке (језгро, граница, шум),

Корак 6: Идентификују се кластери,

Корак 7: Проширења се рекурзивно понављају све док се не пронађу сви квалификовани суседи.

Пондери додељени тачкама у саобраћају могу бити засновани на факторима као што су тежина незгоде, обим саобраћаја, учесталост незгода итд. Ово обезбеђује да кластери одражавају и обрасце густине и релативни значај појединачних података. На тај начин зоне високог ризика могу бити идентификоване прецизније. Треба нагласити да се предложени модификовани алгоритам може применити у бројним областима, као што су сајбер-безбедност, праћење животне средине, здравство, аналитика у малопродаји и многе друге.

Постоји низ потреба за анализом просторних података, тј. података који

су везани за простор. Све већи број података се добија и сателитских снимака, кристалографије, рендгенских зрака или из других аутоматских суистема. Потребе за аутоматским откривањем знања у просторним подацима постају све веће.

Кластер алгоритми су ефикасни у задацима откривања класа. Њихова примена на великим скуповима просторних података захтева одређене предуслове [105]:

1. Основно знање о области над којом се врше истраживања да би се одредили улазни параметри,
2. Откривање кластера одређеног облика, обзиром да кластери у просторним скуповима могу бити сферични, линеарни, издужени и сл.,
3. Добра ефикасност на великим скуповима података, тј. скалабилност.

До појаве алгоритма *DBSCAN* други кластер алгоритми нису могли задовољити сва три услова [105]. *DBSCAN* алгоритам захтева један улазни параметар, а помаже кориснику при одређивању датог параметра. Може да открива кластере различитих облика и ефикасан је чак и на великим просторним скуповима података.

У модификованом *DBSCAN* алгоритму, пондерисање се односи на процес додељивања различитих нивоа значаја тачкама података у складу са тежином и учесталости. Алгоритам користи просторне или геопросторне координате за идентификацију кластера. Он функционише под претпоставком да се, поред просторне идентификације, тачки која се смешта у кластер може доделити пондер, на следећи начин:

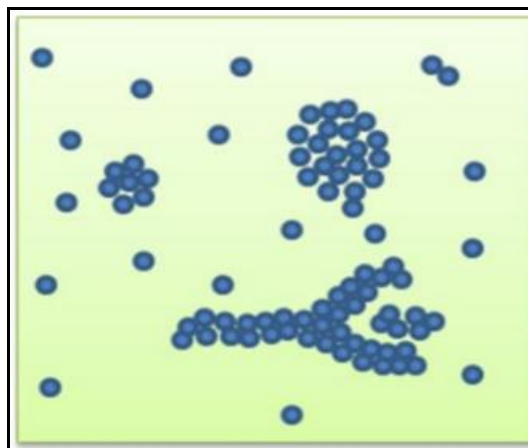
$$w = \alpha S + \beta F$$

где w представља пондер, S и F представљају нормализовану оцену тежине тачке и локалну учесталост појаве тачака, док параметри α и β контролишу утицај ових фактора.

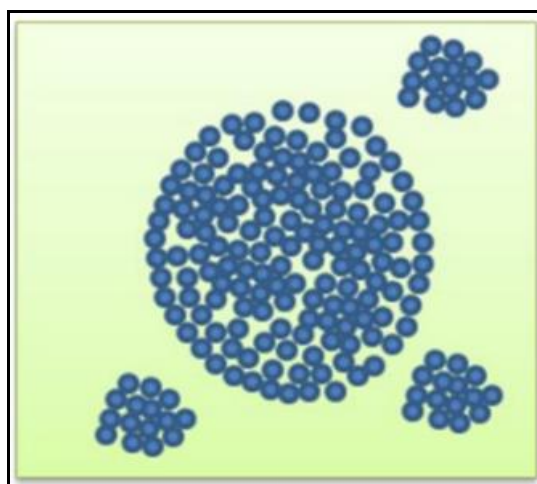
5.1. Појам кластера заснованих на густини

Посматрајући скупове тачака на сликама 6 и 7, лако се уочавају кластери тачака као и тачке које не припадају ни једном кластеру. Ови кластери су лако уочљиви због тога што у сваком кластеру преовладава

одређена густина распореда тачака, која је значајно већа од густине тачака ван кластера. Такође, густина тачака које не припадају кластерима је знатно мања од густине тачака у било ком од кластера.



Слика 6. Пример просторних скупова података



Слика 7. Пример просторних скупова података

Формализација овог запажања је приказана на скупу података D у k -димензионалном простору S . Идеја кластеровања код *DBSCAN* алгоритма је да се свака тачка налази у оквиру одређеног радијуса који садржи одређени минимални број тачака, тј. да у оквиру одређеног радијуса постоји одређена густина (број тачака у одређеном простору).

Одстојање између две тачке може се мерити разним метрикама. Метрике утичу на облик кластера који се проналази.

Алгоритам може функционисати са било којом функцијом одстојања (метриком), тако да се за различите потребе могу изабрати различите функције. Зарад боље визуелизације, у примеру који следи, подаци су представљени у

дводимензионалном простору и коришћењем Еуклидског одстојање.

5.2. Основни појмови

Следи објашњавање основних појмова који су потребни за разумевање алгоритма *DBSCAN*.

Дефиниција 1 (*Eps* околина)

Eps околина неке тачке p се означава као $N_{EPS}(p)$ и дефинише се са (1):

$$N_{EPS}(p) = \{q \in D \mid \text{dist}(p, q) \leq EPS\} \quad (1)$$

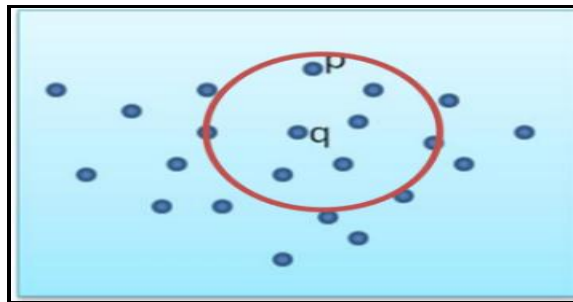
Да би тачка (случај) припадала кластеру потребно је да у њеној *Eps* околини постоји најмање *MinPts* тачака. Овакав приступ има недостатак јер постоје две врсте тачака у кластеру: тачке унутар кластера (централне тачке) и тачке које се налазе на ивици кластера (граничне тачке). Обзиром да је густина око централних тачака већа него око граничних тачака, потребно је то уважити када се траже кластери.

Због различите густине око централних и граничних тачака параметар *MinPts* може да се постави на релативно малу вредност да би биле укључене све тачке које припадају истом кластеру. Параметри *Eps* и *MinPts* се код *DBSCAN* алгоритма могу разликовати од кластера до кластера.

Дефиниција 2 (директна доступност):

Тачка p је директно доступна кроз густину од тачке q ако важи (2).

$$p \in N_{EPS}(q) \wedge |N_{EPS}(q)| \geq MinPts(2)$$



Слика 8. Тачка p је директно доступна кроз густину од q , кружница представља *Eps* околину од q и *MinPts* је 4, што показује да је q централна тачка у кластеру

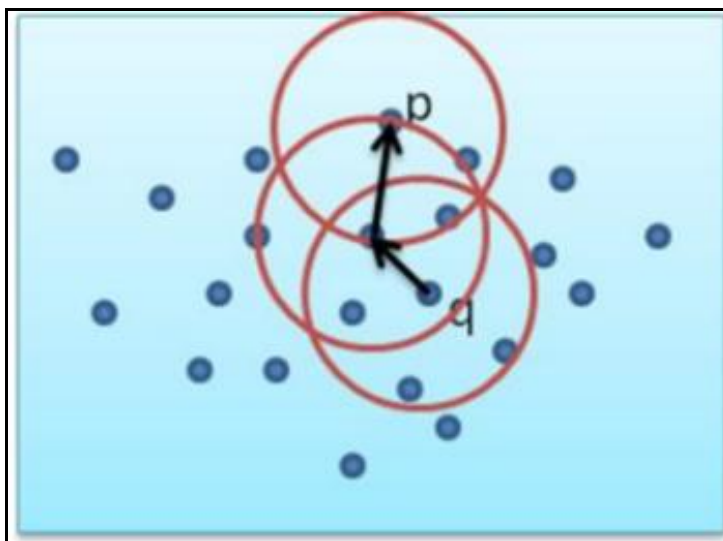
Дефиниција 3 (доступност кроз густину):

Тачка p је доступна кроз густину од тачке q ако постоји низ тачака p_1, p_2, \dots, p_n при чему је $p_1=q$, $p_n=p$ при чему је p_{i+1} тачка која је директно доступна кроз густину од p_i . Доступност кроз густину је проширење директне

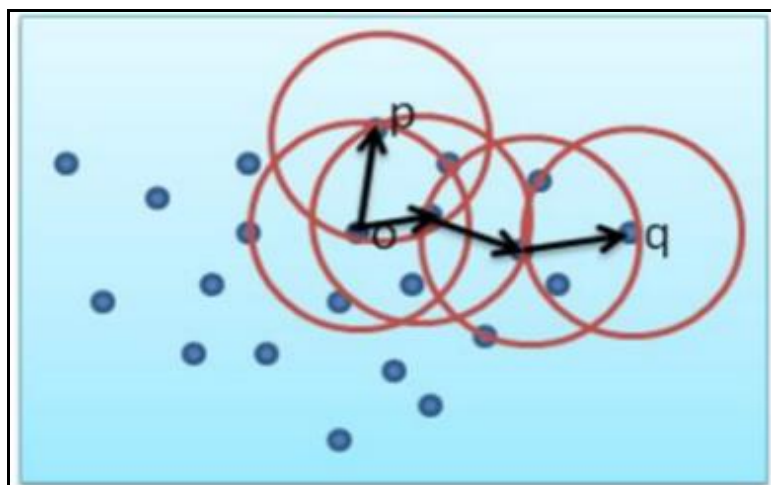
повезаности кроз густину. Ова релација је транзитивна, али није симетрична, што значи да две граничне тачке не морају бити доступне кроз густину, али у сваком кластеру C мора постојати централна тачка од које је било која гранична тачка доступна кроз густину.

Дефиниција 4 (повезаност кроз густину):

Тачка p је повезана кроз густину са тачком q ако постоји тачака o таква да су тачке p и q доступне кроз густину од тачке o .



Слика 9. Тачка p је повезана кроз густину од тачке q



Слика 10. Граничне тачке p и q су повезане кроз густину од централне тачке o

Дефиниција 5 (кластер):

Нека је D скуп тачака. Кластер C је непразан подскуп скупа D који задовољава следеће услове:

1. $\forall p, q: p \in C$ и q је доступно кроз густину од p при условима Eps и $MinPts$, онда и $q \in C$.
2. $\forall p, q: p \in C, p$ је повезано кроз густину са q при условима Eps и $MinPts$.

“Шум” се дефинише у односу на добијене кластере. Представљаја скуп тачака које не припадају ни једном од кластера.

Дефиниција 6 (шум):

Нека су C_1, \dots, C_k кластери на скупу D при условима Eps и $MinPts$ и $i=1, 2, \dots, k$. Тада се “шум” дефинише као скуп тачака из скупа D који не припадају ни једном од кластера C_i .

Следећа лема је битна за проверу тачности кластер алгоритма. Узевши параметре Eps и $MinPts$ кластер се може открити у два корака. Прво се бира тачка из скупа података која задовољава услов да буде централна тачка и она је почетна тачка у односу на коју креће претрага по подацима. Централна тачка је она тачка у чијој се Eps околини налази најмање $MinPts$ тачака. Након тога се одређују све тачке које су доступне кроз густину од почетне тачке чиме се добија кластер који садржи почетну тачку.

Лема 1.

Нека је p тачка скупа D и $|N_{EPS}(p)| \geq MinPts$. Тада је скуп $O = \{o \mid o \in D\}$ при чему су тачке o доступне кроз густину од тачке p при условима Eps и $MinPts$, кластер при условима Eps и $MinPts$.

Кластер C при условима Eps и $MinPts$ може да има различите централне тачке као почетне. Међутим, свака тачка скупа C је доступна кроз густину од било које од централних тачака скупа C , те кластер C садржи све оне тачке које су доступне кроз густину од почетне централне тачке скупа C .

Лема 2.

Нека је C кластер при условима Eps и $MinPts$ и нека је p било која тачка унутар C која задовољава $|N_{EPS}(p)| \geq MinPts$. Тада је скуп C једнак скупу O .

5.3. Алгоритам

DBSCAN алгоритам је осмишљен да открије кластере и шум у просторном скупу података ослањајући се на дефиниције 5 и 6 (дефиниције кластера и шума). У идеалном случају, ослањајући се на Eps и $MinPts$ су

познати за сваки кластер као и барем по једна тачка за сваки кластер. Тада би се могле открити све тачке које су доступне кроз густину од тих репрезентативних тачака користећи дате параметре.

Проблем је што не постоји једноставан начин да се ова информација има унапред за све кластере. Међутим, постоји једноставна и ефективна хеуристика која одређује параметре Eps и $MinPts$ на основу кластера са најмањом густином. Због тога *DBSCAN* користи вредности Eps и $MinPts$ за све кластере. Параметри кластера са најмањом густином су добри кандидати за вредности глобалних параметара, одређујући најмању густину која се не сматра “шумом”.

5.4. Откривање параметара Eps и $MinPts$

У овом делу је описан хеуристички метод за откривање параметара Eps и $MinPts$. Предложена хеуристика у [105] полази од тога да се Eps емпиријски најбоље одређује када је $MinPts = 4$.

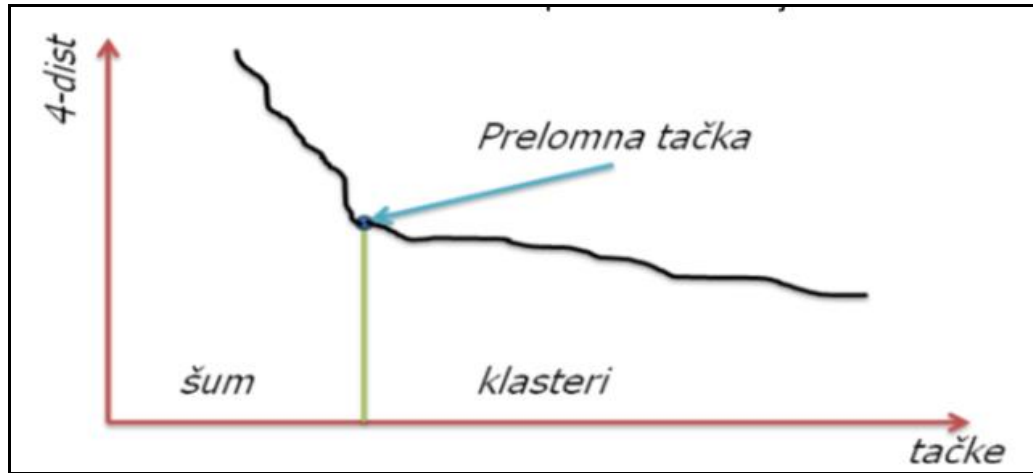
Нека је d одстојање од тачке p до њених k најближих суседа. Тада d околина тачке p садржи тачно $k+1$ тачку. За задато k дефинише се функција $k-dist$ на следећи начин: за сваку тачку из скупа се одреди k -то најмање одстојање (нпр. за $k=4$ бира се четврта по реду најмања вредност одстојања). Када се сортирају тачке у нерастући поредак по својим $k-dist$ вредностима може да се направи граф где су на апциси тачке које треба да се кластерују, а на ординати $k-dist$ вредности. Овај график се назива “сортирани $k-dist$ граф”.

Са графика приказаног на Слици 11 је потребно наћи тзв. преломну тачку. Обзиром да није лако аналитички одредити ову тачку, кориснику се препушта да изабере тачку где се график “ломи”. Избором те тачке, корисник одређује параметар Eps . Преломна тачка је прва тачка у којој се график “ломи” на сортираном $k-dist$ графу.

DBSCAN тражи два параметра Eps и $MinPts$. Међутим, емпиријски је доказано да се $k-dist$ графови за $k > 4$ не разликују значајно од $4-dist$, а захтевају више времена за израчунавање. Овим се елиминише параметар $MinPts$ тако што му се даје вредност 4 за све скупове.

Да би пронашао кластер, *DBSCAN* алгоритам рад започиње тако што одређује почетну тачку p и одређује све тачке q које су доступне кроз густину од p при условима Eps и $MinPts$. Ако је p централна тачка овај поступак даје кластер при условима Eps и $MinPts$ (по Леми 2). Ако је p гранична тачка, тада

нема тачака које су доступне кроз густину ρ и *DBSCAN* алгоритам узима следећу тачку из скупа на разматрање.



Слика 11. Сортиран 4-dist граф и преломна тачка

Како се користе глобалне вредности Eps и $MinPts$, *DBSCAN* може да споји два кластера по дефиницији 5 у један кластер, ако су два кластера различите густине “близу” један другом. Нека се удаљеност два скупа тачака S_1 и S_2 дефинише као (3):

$$dist(S_1, S_2) = \min\{dist(p, q) \mid p \in S_1, q \in S_2\} \quad (3)$$

Алгоритам *DBSCAN* се одвија по следећој процедури:

- 1) Израчунају се сва одстојања у скупу тачака
 - 1.1. Од свих одстојања се креира матрица одстојања (колоне и врсте су тачке, а у пресецима је вредност одстојања између тих тачака)
2. Формира се сортирани 4-dist граф.
 - 2.1. За сваку тачку одредити четврто по реду најмање одстојање;
 - 2.2. Сортирати тачке у опадајућем поредку по овим вредностима;
 - 2.3. Нацртати график за овако сортиране тачке.
3. На сортираном 4-dist графу се проналази преломна тачка (почетна тачка прве “равнице”).
4. Вредност 4 се узима за $MinPts$, а вредност преломне тачке на графику се узима за параметар Eps . Ако конкретан случај захтева другачију вредност за $MinPts$ онда се од корака 2 узима та вредност за креирање графа: 3-dist, 5-dist, и сл.
5. Користећи матрицу одстојања траже се оне тачке које имају најмање $MinPts$ тачака од којих су одаљене не више од параметара Eps .

- 5.1. Узима се тачка која задовољава услове *Eps* и *MinPts*. Она постаје центар кластера коме припадају све тачке које су од ње удаљене за више од *Eps*;
- 5.2. За сваку тачку која припада том кластеру се поново проверава да ли је и она центар кластера, ако јесте онда се та два кластера спајају (односно први кластер се даље проширује). Ако пак та тачка не задовољава услове, она постаје гранична тачка кластера;
- 5.3. Када су услови проверени за све тачке у кластеру, одређује се једна од преосталих тачака из скупа и цео поступак се понавља и од ње настаје следећи кластер;
- 5.4. Ако су у скупу преостале тачке које не припадају кластерима, нити се од њих може направити нови кластер, оне постају шум.

5.5. Пример

Пример функционисања *DBSCAN* алгоритма приказан је на скупу података који приказује државе и очекиване животне векове мушкараца и жена у тим државама.

Табела 5. Скуп података са државама и очекиваним животним веком [121]

| Држава | Очекивани животни век | Становника по телевизору | Становника по лекару | Очекивани животни век жена | Очекивани животни век мушкараца |
|----------------|-----------------------|--------------------------|----------------------|----------------------------|---------------------------------|
| Аргентина | 70.5 | 4 | 370 | 74 | 67 |
| Бангладеш | 53.5 | 315 | 6166 | 53 | 54 |
| Бразил | 65 | 4 | 684 | 68 | 62 |
| Канада | 76.5 | 1.7 | 449 | 80 | 73 |
| Кина | 70 | 8 | 643 | 72 | 68 |
| Колумбија | 71 | 5.6 | 1551 | 74 | 68 |
| Египат | 60.5 | 15 | 616 | 61 | 60 |
| Етиопија | 51.5 | 503 | 36660 | 53 | 50 |
| Француска | 78 | 2.6 | 403 | 82 | 74 |
| Немачка | 76 | 2.6 | 346 | 79 | 73 |
| Индија | 57.5 | 44 | 2471 | 58 | 57 |
| Индонезија | 61 | 24 | 7427 | 63 | 59 |
| Иран | 64.5 | 23 | 2992 | 65 | 64 |
| Италија | 78.5 | 3.8 | 233 | 82 | 75 |
| Јапан | 79 | 1.82 | 609 | 82 | 76 |
| Кенија | 61 | 96 | 7615 | 63 | 59 |
| Северна Кореја | 70 | 90 | 370 | 73 | 67 |
| Јужна Кореја | 70 | 4.9 | 1066 | 73 | 67 |
| Мексико | 72 | 6.6 | 600 | 76 | 68 |
| Мароко | 64.5 | 21 | 4873 | 66 | 63 |
| Мјанмар | 54.5 | 592 | 3485 | 56 | 53 |
| Пакистан | 56.5 | 73 | 2364 | 57 | 56 |
| Перу | 64.5 | 14 | 1016 | 67 | 62 |
| Филипини | 64.5 | 8.8 | 1062 | 67 | 62 |
| Пољска | 73 | 3.9 | | | |
| Румунија | 72 | 6 | 559 | 75 | 69 |
| Русија | 69 | 3.2 | 259 | 74 | 64 |
| Јужна Африка | 64 | 11 | 1340 | 67 | 61 |
| Шпанија | 78.5 | 2.6 | 275 | 82 | 75 |
| Судан | 53 | 23 | 12550 | 54 | 52 |

| | | | | | |
|------------------|------|-----|------|----|----|
| Тајван | 75 | 3.2 | 965 | 78 | 72 |
| Тајланд | 65.5 | 11 | 4883 | 71 | 66 |
| Турска | 70.5 | 3 | 226 | 75 | 66 |
| Украјина | 70.5 | 3 | 226 | 75 | 66 |
| Велика Британија | 76 | 3 | 611 | 79 | 72 |
| САД | 75.5 | 1.3 | 404 | 79 | 72 |
| Венецуела | 74.5 | 5.6 | 576 | 78 | 71 |
| Вијетнам | 65 | 29 | 3096 | 67 | 63 |

Да би се подаци из Табеле 5 могли искористити, најпре се приступа поступку нормализације. Сви подаци се свде на опсег од 0 до 1 при чему ће подаци у свакој колони задржати своје оригиналне односе вредности. На примеру колоне “Очекивани животни век” уочава се да је највећа вредност 79 за државу Јапан. Нова вредност (V_n) се рачуна тако што се тренутна вредност (V) подели са највећом вредношћу колоне (V_{max}), односно:

$$V_{n=\frac{v}{V_{max}}}$$

На овај начин, добија се нова табела са подацима приказаних у Табели 6.

Табела 6. Нормализовани подаци у опсегу вредности од 0 до 1

| Држава | Очекивани животни век | Становника по телевизору | Становника по лекару | Очекивани животни век жена | Очекивани животни век мушкараца |
|------------------|-----------------------|--------------------------|----------------------|----------------------------|---------------------------------|
| Аргентина | 0.89 | 0.01 | 0.01 | 0.90 | 0.88 |
| Бангладеш | 0.68 | 0.53 | 0.17 | 0.65 | 0.71 |
| Бразил | 0.82 | 0.01 | 0.02 | 0.83 | 0.82 |
| Канада | 0.97 | 0.00 | 0.01 | 0.98 | 0.96 |
| Кина | 0.89 | 0.01 | 0.02 | 0.88 | 0.89 |
| Колумбија | 0.90 | 0.01 | 0.04 | 0.90 | 0.89 |
| Египат | 0.77 | 0.03 | 0.02 | 0.74 | 0.79 |
| Етиопија | 0.65 | 0.85 | 1.00 | 0.65 | 0.66 |
| Француска | 0.99 | 0.00 | 0.01 | 1.00 | 0.97 |
| Немачка | 0.96 | 0.00 | 0.01 | 0.96 | 0.96 |
| Индија | 0.73 | 0.07 | 0.07 | 0.71 | 0.75 |
| Индонезија | 0.77 | 0.04 | 0.20 | 0.77 | 0.78 |
| Иран | 0.82 | 0.04 | 0.08 | 0.79 | 0.84 |
| Италија | 0.99 | 0.01 | 0.01 | 1.00 | 0.99 |
| Јапан | 1.00 | 0.00 | 0.02 | 1.00 | 1.00 |
| Кенија | 0.77 | 0.16 | 0.21 | 0.77 | 0.78 |
| Северна Кореја | 0.89 | 0.15 | 0.01 | 0.89 | 0.88 |
| Јужна Кореја | 0.89 | 0.01 | 0.03 | 0.89 | 0.88 |
| Мексико | 0.91 | 0.01 | 0.02 | 0.93 | 0.89 |
| Мароко | 0.82 | 0.04 | 0.13 | 0.80 | 0.83 |
| Мјанмар | 0.69 | 1.00 | 0.10 | 0.68 | 0.70 |
| Пакистан | 0.72 | 0.12 | 0.06 | 0.70 | 0.74 |
| Перу | 0.82 | 0.02 | 0.03 | 0.82 | 0.82 |
| Филипини | 0.82 | 0.01 | 0.03 | 0.82 | 0.82 |
| Пољска | 0.92 | 0.01 | 0.01 | 0.94 | 0.91 |
| Румунија | 0.91 | 0.01 | 0.02 | 0.91 | 0.91 |
| Русија | 0.87 | 0.01 | 0.01 | 0.90 | 0.84 |
| Јужна Африка | 0.81 | 0.02 | 0.04 | 0.82 | 0.80 |
| Шпанија | 0.99 | 0.00 | 0.01 | 1.00 | 0.99 |
| Судан | 0.67 | 0.04 | 0.34 | 0.66 | 0.68 |
| Тајван | 0.95 | 0.01 | 0.03 | 0.95 | 0.95 |
| Тајланд | 0.83 | 0.02 | 0.13 | 0.87 | 0.87 |
| Турска | 0.89 | 0.01 | 0.01 | 0.91 | 0.87 |
| Украјина | 0.89 | 0.01 | 0.01 | 0.91 | 0.87 |
| Велика Британија | 0.96 | 0.01 | 0.02 | 0.96 | 0.95 |

| | | | | | |
|-----------|------|------|------|------|------|
| САД | 0.96 | 0.00 | 0.01 | 0.96 | 0.95 |
| Венецуела | 0.94 | 0.01 | 0.02 | 0.95 | 0.93 |
| Вијетнам | 0.82 | 0.05 | 0.08 | 0.82 | 0.83 |

За државу Аргентину, одстојање Аргентине од Бангладеша износи 0.66. На исти начин врши се и прорачун свих осталих одстојања. Наравно, како је одстојање Аргентине од Бангладеша једнако као и одстојање од Бангладеша до Аргентине, биће потребно израчунати $38 \cdot 37 / 2 = 703$ одстојања.

Од израчунатих одстојања креира се матрица одстојања за све државе, како је приказано у табели која следи у наставку, да би се видела одстојања. Матрица одстојања представља матрицу којој су редови и колоне државе, а на пресецима се налазе њихова одстојања. Јасно је да ће матрица на дијагонали имати вредност 0 (одстојање државе са самом собом), као и да је матрица симетрична по главној дијагонали (одстојање државе А ка држави Б је исто као и одстојање од државе Б ка држави А).

Одређивање почетних вредности

По хеуристици описаној у претходном делу, за *MinPts* се узима вредност 4, а *Eps* се одређује на основу сортираног *4-dist* графа. Овај график се формира на следећи начин:

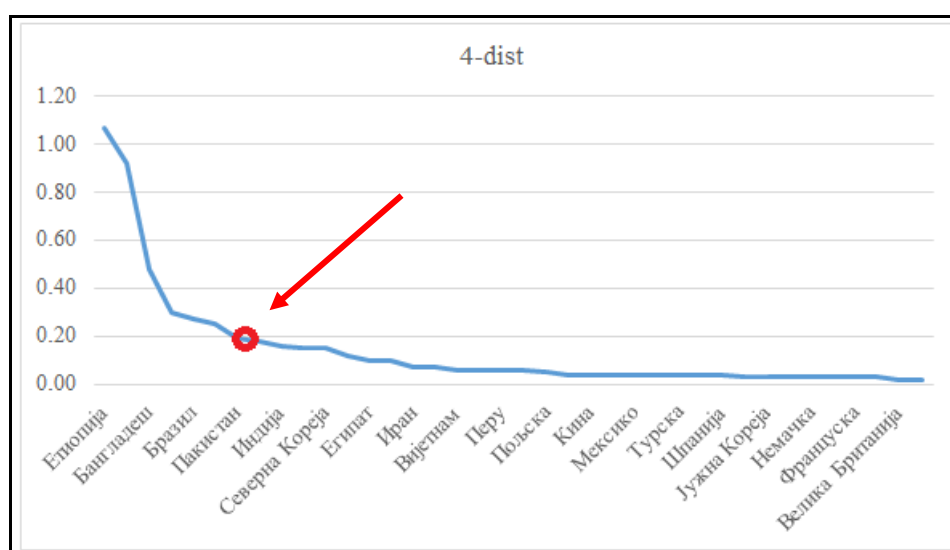
1. За сваку државу се бира четврта по реду најмања вредност одстојања,
2. График се формира за добијене вредности сортиране у опадајућем редоследу.

У Табели 7 приказане су четврте по реду најмање вредности одстојања земаља сортиране у опадајућем редоследу:

Табела 7. Сортиране вредности четвртог најмањег одстојања за сваку државу

| Држава | <i>4-dist</i> |
|----------------|---------------|
| Етиопија | 1.07 |
| Мјанмар | 0.92 |
| Бангладеш | 0.48 |
| Судан | 0.30 |
| Бразил | 0.27 |
| Венецуела | 0.25 |
| Пакистан | 0.19 |
| Кенија | 0.18 |
| Индија | 0.16 |
| Индонезија | 0.15 |
| Северна Кореја | 0.15 |
| Тајланд | 0.12 |
| Египат | 0.10 |
| Мароко | 0.10 |
| Иран | 0.07 |
| Филипини | 0.07 |
| Вијетнам | 0.06 |
| Јужна Африка | 0.06 |

| | |
|------------------|------|
| Перу | 0.06 |
| Русија | 0.06 |
| Пољска | 0.05 |
| Италија | 0.04 |
| Кина | 0.04 |
| Колумбија | 0.04 |
| Мексико | 0.04 |
| Румунија | 0.04 |
| Турска | 0.04 |
| Украјина | 0.04 |
| Шпанија | 0.04 |
| Аргентина | 0.03 |
| Јужна Кореја | 0.03 |
| Канада | 0.03 |
| Немачка | 0.03 |
| Тајван | 0.03 |
| Француска | 0.03 |
| Јапан | 0.03 |
| Велика Британија | 0.02 |
| САД | 0.02 |



Слика 12. Сортирани 4-dist график и преломна тачка

Преломна тачка има вредност одстојања 0.19 (вредност за државу Пакистан).

Креирање кластера

Сада је потребно посматрати матрицу одстојања на којој су обележене оне вредности које су мање или једнаке 0.19. Бира се свака држава и посматра се које државе се налазе у њеној околини. Ако има мање од 4 државе у тој околини онда се те државе налазе у кластеру.

У случају Пакистана, уочава се да се у околини налазе тачно четири државе (Египат, Индија, Индонезија и Кенија). Ово значи да се формира кластер коме је центар Пакистан. Даље се посматрају државе које припадају овом кластеру

- У околини Египта: Бразил, Индија, Индонезија, Иран, Кенија, Мароко, Пакистан, Перу, Филипини, Јужна Африка и Вијетнам.
- У околини Индије: Бразил, Египат, Индонезија, Иран, Кенија, Мароко, Пакистан, Перу, Филипини, Јужна Африка и Вијетнам.
- У околини Индонезије: Египат, Индија, Иран, Кенија, Мароко, Пакистан, Јужна Африка, Тајланд и Вијетнам.
- У околини Кеније: Индија, Индонезија, Мароко, Пакистан и Вијетнам.

Све тачке које су у почетном кластеру са Пакистаном такође су центри својих кластера, што значи да се сви ови кластери спајају у један кластер кога чине: Пакистан, Бразил, Египат, Индонезија, Иран, Мароко, Кенија, Индија, Перу, Филипини, Јужна Африка, Тајланд и Вијетнам. На примеру Египат је директно повезан кроз густину са Пакистаном, али је и Бразил повезан кроз густину са Пакистаном, тако што је директно повезан са Египтом.

Сада се посматра Бразил као тачка у кластеру коме је центар Египат и у његовој околини се налазе: Аргентина, Кина, Колумбија, Египат, Индија, Индонезија, Иран, Јужна Кореја, Северна Кореја, Мексико, Мароко, Перу, Филипини, Пољска, Румунија, Русија, Јужна Африка, Шпанија, Тајван, Тајланд, Турска, Украјина, Велика Британија, САД, Венецуела и Вијетнам.

Како се и овај кластер спаја са претходним, ситуација је таква да се скоро све државе налазе у једном кластеру. Преостале четири државе које се налазе у овом кластеру су: Бангладеш, Етиопија, Мијанмар и Судан. Ове четири државе немају ни једну државу у својој *Eps* околини. Дакле, резултат кластерованја је:

- Један кластер који броји 34 државе;
- Шум кога чине шест кластера: Бангладеш, Етиопија, Мијанмар и Судан.

Анализом ове четири државе у односу на остале, може се приметити да је ту веома низак очекивани животни век (51 – 54 године), такође велики број становника по телевизору и лекару. Евидентно је да овај шум чине оне државе које су мање развијене.

Пример употребе модификованог алгоритма

Пример примене наведеног алгоритма су подаци о саобраћајним

незгодама.

Аутори у [122] наводе да су “црне тачке” подручја у мрежи јавних путева на којима је број незгода знатно већи од очекиваног. Постоји неколико метода које могу бити коришћене за проналажење ових “црних тачака” и оне обично користе систем позиционирања заснован на броју пута и деоница пута, што је корисно у случајевима када желимо да пронађемо “црне тачке” само на једном путу.

Међутим, ове методе имају и недостатке, посебно у случају раскрсница, међуградских подручја итд. Било би боље користити *GPS* координате за њихову идентификацију [123]. У свом раду, аутори за идентификацију “црних тачака” користе методу кластерована засновану на *DBSCAN*-у. Наводе да *DBSCAN* представља општи алгоритам кластерована, па је прилагођен за рад са саобраћајним незгодама идентификованим помоћу *GPS* координата. Такође, алгоритам представља брзу и тачну технику која израчунава густину незгода у датом подручју.

Аутори [124] у свом раду предлажу двостепени модел идентификације “црних тачака” који је заснован на *DBSCAN* алгоритму са циљем примене на густини просторног кластерована како би били идентификовани и лоцирани пловни путеви са вишим стопама несрећа.

У свом раду [125], аутори наводе да последњих година, питања безбедности градског саобраћаја све више привлаче пажњу, обзиром да саобраћајне незгоде представљају значајне претње по безбедност живота и имовине људи. Нестандардно понашање возача у вожњи често се дешава истовремено са саобраћајним незгодама, а количина података о нестандартном понашању у вожњи је знатно већа од количине података о незгодама, што га чини свеобухватнијим одразом услова вожње и заслужује детаљнију анализу. Постојећа истраживања ретко повезују податке о нестандартном понашању у вожњи са координатама на мапи, а употреба метода визуелизације за приказивање дистрибуције тачака података је ограничена [126]. У свом раду предлажу побољшани алгоритам кластерована *DBSCAN* заснован на подацима о нестандартном понашању у вожњи и користи алгоритам минималног обухватног стабла за израчунавање густине сваког кластера кластерована. Сматрају да се израчунавањем густине кластера кластерована, може боље

разумети просторна дистрибуција нестандардног понашања у возњи, даље анализирати степен кластеровања ових кластера и проучити дистрибуцију нестандардног понашања у возњи.

Скуп тачака у простору које представљају саобраћајне незгоде могуће је сматрати репрезентативним, обзиром да се ради о просторним подацима, на основу географске дужине и географске ширине, где се саобраћајна незгода догодила. Део података за једну општину, за временски период од 3 године, дат је у наставку текста.

| id | identifikator | grad | opsina | vreme | duzina | sirina | visina | tip | opis | oznaka | |
|------|---------------|-------|--------|------------------|------------|-----------|--------|-------------|---|--|-----------|
| 8801 | 149390 | UŽICE | ARLJE | 14.01.2024.19:55 | 20.10467 | 43.74033 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkudeno sa kolovoz na pravcu | cvcv_8801 |
| 8802 | 149410 | UŽICE | ARLJE | 15.01.2024.07:25 | 20.10974 | 43.75589 | 5a | mat stacion | SNŠA NAJMANJE DVA VOZILA – SKRETANJE IЛИ PRELAZAK | Najmanje dva vozila koja se leku istim putem u suprotnim smerovima u sletanje ulavo ispred drugog vozila | cvcv_8802 |
| 8803 | 149033 | UŽICE | ARLJE | 25.01.2024.07:30 | 20.15118 | 43.70698 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkasa kolovoz u letrini | cvcv_8803 |
| 8804 | 149089 | UŽICE | ARLJE | 26.01.2024.01:00 | 20.09788 | 43.74592 | 5a | mat stacion | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkasa kolovoz u letrini | cvcv_8804 |
| 8805 | 149211 | UŽICE | ARLJE | 09.02.2024.19:30 | 20.144152 | 43.70523 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom i prevrtanje | cvcv_8805 |
| 8806 | 149735 | UŽICE | ARLJE | 14.02.2024.18:45 | 20.089594 | 43.74933 | 5a | mat stacion | SNŠA NAJMANJE DVA VOZILA – BEZ SKRETANJA | Ostale nezgode sa najmanje dva vozila – suprotni smerovi bez sletanja | cvcv_8806 |
| 8807 | 149218 | UŽICE | ARLJE | 15.02.2024.23:15 | 20.115547 | 43.74225 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkudeno sa kolovoz na pravcu | cvcv_8807 |
| 8808 | 149420 | UŽICE | ARLJE | 16.02.2024.15:50 | 20.100155 | 43.75428 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom bez prepreka na kolovozu na nepoznat, ne specifičan način | cvcv_8808 |
| 8809 | 149791 | UŽICE | ARLJE | 10.03.2024.14:45 | 20.072314 | 43.74959 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa učesnik, jednog vozila i biciklista | cvcv_8809 |
| 8810 | 149750 | UŽICE | ARLJE | 16.03.2024.09:30 | 20.1119489 | 43.625885 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkudeno sa kolovoz na pravcu | cvcv_8810 |
| 8811 | 149877 | UŽICE | ARLJE | 20.03.2024.20:30 | 20.124394 | 43.68587 | 5a | mat stacion | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkasa kolovoz u letrini | cvcv_8811 |
| 8812 | 149130 | UŽICE | ARLJE | 27.03.2024.01:25 | 20.124394 | 43.68584 | 5a | mat stacion | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkasa kolovoz u letrini | cvcv_8812 |
| 8813 | 149898 | UŽICE | ARLJE | 01.04.2024.14:10 | 20.1197222 | 43.700833 | 5a | mat stacion | SNŠA PRELAZAK | Bele se leku duž kolovoza u smeru kretanja vozila | cvcv_8813 |
| 8814 | 149996 | UŽICE | ARLJE | 02.04.2024.02:45 | 20.081103 | 43.78184 | 5a | поведајен | SNŠA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa desne strane kolovoza | cvcv_8814 |
| 8815 | 147041 | UŽICE | ARLJE | 03.04.2024.12:50 | 20.11001 | 43.73744 | 5a | поведајен | SNŠA NAJMANJE DVA VOZILA – BEZ SKRETANJA | Ostale nezgode sa najmanje dva vozila – suprotni smerovi bez sletanja | cvcv_8815 |
| 8816 | 147146 | UŽICE | ARLJE | 12.04.2024.06:55 | 20.089104 | 43.73389 | 5a | поведајен | SNŠA NAJMANJE DVA VOZILA – SKRETANJE IЛИ PRELAZAK | Najmanje dva vozila koja se leku različitim putevima uz prelazak izvoza raskrsnicu. Svi ođ koji je jedno preklazi preko kolovoza, bez sletanja | cvcv_8816 |
| 8817 | 147420 | UŽICE | ARLJE | 24.04.2024.10:50 | 20.118496 | 43.70544 | 5a | поведајен | SNŠA NAJMANJE DVA VOZILA – SKRETANJE IЛИ PRELAZAK | Najmanje dva vozila koja se leku istim putem u istom smeru u sletanje, sletanje ulavo ispred drugog vozila | cvcv_8817 |
| 8818 | 147434 | UŽICE | ARLJE | 06.05.2024.02:30 | 20.106489 | 43.671712 | 5a | mat stacion | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkudeno sa kolovoz na pravcu | cvcv_8818 |
| 8819 | 147866 | UŽICE | ARLJE | 08.06.2024.16:50 | 20.101321 | 43.74934 | 5a | mat stacion | SNŠA NAJMANJE DVA VOZILA – SKRETANJE IЛИ PRELAZAK | Najmanje dva vozila koja se leku istim putem u istom smeru u sletanje, sletanje ulavo ispred drugog vozila | cvcv_8819 |
| 8820 | 147917 | UŽICE | ARLJE | 13.06.2024.21:20 | 20.1055837 | 43.748279 | 5a | поведајен | SNŠA NAJMANJE DVA VOZILA – SKRETANJE IЛИ PRELAZAK | Najmanje dva vozila koja se leku istim putem u suprotnim smerovima u sletanje ulavo ispred drugog vozila | cvcv_8820 |
| 8821 | 147928 | UŽICE | ARLJE | 16.06.2024.00:05 | 20.106285 | 43.748205 | 5a | mat stacion | SNŠA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa leve strane kolovoza | cvcv_8821 |
| 8822 | 148059 | UŽICE | ARLJE | 26.06.2024.21:30 | 20.126472 | 43.746329 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkudeno sa kolovoz na pravcu | cvcv_8822 |
| 8823 | 148083 | UŽICE | ARLJE | 28.06.2024.09:30 | 20.107983 | 43.790291 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkasa kolovoz u letrini | cvcv_8823 |
| 8824 | 148079 | UŽICE | ARLJE | 28.06.2024.20:25 | 20.118539 | 43.705321 | 5a | поведајен | SNŠA NAJMANJE DVA VOZILA – SKRETANJE IЛИ PRELAZAK | Najmanje dva vozila koja se leku istim putem u suprotnim smerovima u sletanje ulavo ispred drugog vozila | cvcv_8824 |
| 8825 | 148134 | UŽICE | ARLJE | 03.07.2024.15:30 | 20.105585 | 43.748305 | 5a | поведајен | SNŠA NAJMANJE DVA VOZILA – BEZ SKRETANJA | Najmanje dva vozila koja se leku u istom smeru – sustizanje | cvcv_8825 |
| 8826 | 148158 | UŽICE | ARLJE | 04.07.2024.22:20 | 20.096261 | 43.73235 | 5a | mat stacion | SNŠA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa desne strane kolovoza | cvcv_8826 |
| 8827 | 1482571 | UŽICE | ARLJE | 13.07.2024.16:45 | 20.094215 | 43.747979 | 5a | mat stacion | SNŠA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa desne strane kolovoza | cvcv_8827 |
| 8828 | 148282 | UŽICE | ARLJE | 14.07.2024.10:45 | 20.112217 | 43.790289 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkudeno sa kolovoz na pravcu | cvcv_8828 |
| 8829 | 149357 | UŽICE | ARLJE | 21.07.2024.21:00 | 20.095329 | 43.747488 | 5a | mat stacion | SNŠA NAJMANJE DVA VOZILA – BEZ SKRETANJA | Ostale nezgode sa najmanje dva vozila – suprotni smerovi bez sletanja | cvcv_8829 |
| 8830 | 148626 | UŽICE | ARLJE | 12.08.2024.19:10 | 20.085019 | 43.750197 | 5a | mat stacion | SNŠA NAJMANJE DVA VOZILA – BEZ SKRETANJA | Ostale nezgode sa najmanje dva vozila – suprotni smerovi bez sletanja | cvcv_8830 |
| 8831 | 148628 | UŽICE | ARLJE | 12.08.2024.14:10 | 20.1001173 | 43.778533 | 5a | поведајен | SNŠA NAJMANJE DVA VOZILA – BEZ SKRETANJA | Ostale nezgode sa najmanje dva vozila – suprotni smerovi bez sletanja | cvcv_8831 |
| 8832 | 148680 | UŽICE | ARLJE | 15.08.2024.13:45 | 20.094989 | 43.784806 | 5a | mat stacion | SNŠA NAJMANJE DVA VOZILA – SKRETANJE IЛИ PRELAZAK | Najmanje dva vozila koja se leku različitim putevima u sletanje ulavo i uz malik vozila zdesna | cvcv_8832 |
| 8833 | 148694 | UŽICE | ARLJE | 17.08.2024.14:55 | 20.09911 | 43.781131 | 5a | поведајен | SNŠA NAJMANJE DVA VOZILA – SKRETANJE IЛИ PRELAZAK | Najmanje dva vozila koja se leku istim putem u suprotnim smerovima u sletanje ulavo ispred drugog vozila | cvcv_8833 |
| 8834 | 148788 | UŽICE | ARLJE | 19.08.2024.06:45 | 20.094994 | 43.751188 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkudeno sa kolovoz na pravcu | cvcv_8834 |
| 8835 | 149124 | UŽICE | ARLJE | 04.09.2024.12:40 | 20.089321 | 43.767954 | 5a | поведајен | SNŠA JEDNIM VOZILOM | Negoda sa jednim vozilom – štamkudeno sa kolovoz na pravcu | cvcv_8835 |
| 8836 | 148486 | UŽICE | ARLJE | 05.09.2024.17:04 | 20.094775 | 43.752494 | 5a | mat stacion | SNŠA NAJMANJE DVA VOZILA – SKRETANJE IЛИ PRELAZAK | Najmanje dva vozila koja se leku istim putem u istom smeru u sletanje, sudar u sustizanje | cvcv_8836 |
| 8837 | 1490380 | UŽICE | ARLJE | 13.09.2024.12:05 | 20.117047 | 43.688529 | 5a | поведајен | SNŠA NAJMANJE DVA VOZILA – BEZ SKRETANJA | Ostale nezgode sa najmanje dva vozila – suprotni smerovi bez sletanja | cvcv_8837 |

Слика 13. Део података саобраћајних незгода у општини Ариље за период од 2021. до 2024. године

У 2008. години, Међународна асоцијација путних радника (PIARC) обезбедила је систематски оквир за одређивање тежине саобраћајних незгода [49]. У оквиру овог модела, записима о незгодама додељују се пондери на основу нивоа повреда: незгоде са мањим повредама (огреботине, мање посекотине, повреде унутар усне дупље) добијају пондер 1; незгоде са тешким повредама (животно угрожавајуће повреде које, ако се не третирају одмах, могу довести до озбиљних телесних оштећења или смрти) добијају пондер 10;

незгоде са смртним исходом добијају пондер 85. Овај начин пондерисања тежине незгода испитиван је и у студијама Европске комисије.

Модификовани *DBSCAN* алгоритам идентификује кластере око „дрних тачака“ на основу пондерисаног броја индикатора саобраћајних незгода (WNI_i), према формули:

$$WNI_i = 85n_{fi} + 10n_{si} + n_{mi}, i = 1, \dots, n_{tot}$$

где n_{fi} , n_{si} и n_{mi} представљају број саобраћајних незгода са смртним исходом, тешким и мањим повредама за сваки запис, док n_{tot} представља укупан број записа.

Модификовани *DBSCAN* алгоритам рангира кластере на основу збира свих WNI вредности унутар сваког кластера. Ако је тај збир већи или једнак 85, кластери се задржавају за даљу анализу и сортирају се по опадајућем редоследу значаја. Сви остали кластери се не узимају у обзир. Ако се кластери преклапају, а центар једног кластера је ван радијуса другог, кластер са мањим збиром се укључује у већи. У свакој наредној итерацији нови радијус се проширује на просечан радијус израчунат за све кластере. Нови збир се истовремено поново израчунава, што омогућава моделу да прецизније процени значај кластера. Адаптивно прилагођавање обезбеђује прецизније дефинисање просторних граница, омогућавајући тачнију детекцију зона високог ризика.

Нумерички подаци за обраду су географска дужина и географска ширина. Поред њих подаци који се из текстуалних могу конвертовати у нумеричке, а на основу њихових пондерисаних вредности су следећи подаци о типу саобраћајне незгоде:

1. Саобраћајне незгоде са погинулим лицима - 85
2. Саобраћајне незгоде са повређеним лицима - 10
3. Саобраћајне незгоде са материјалном штетом - 1

Део података припремљених за прорачун растојања, дат је на Слици 14.

| | id | id_jedan | duzina_jedan | sirina_jedan | oznaka_jedan | id_dva | duzina_dva | sirina_dva | oznaka_dva | vrsta | rastojanje | vrednost |
|----|----|----------|--------------|--------------|--------------|--------|------------|------------|------------|----------------|------------------|----------|
| 1 | 1 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8802 | 20.109074 | 43.736896 | cvor_8802 | Sa mat.stetom | 672.031174613497 | 1 |
| 2 | 2 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8803 | 20.151193 | 43.706668 | cvor_8803 | Sa povredjenim | 5430.63924120468 | 10 |
| 3 | 3 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8804 | 20.097938 | 43.745592 | cvor_8804 | Sa mat.stetom | 671.131747566719 | 1 |
| 4 | 4 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8805 | 20.144152 | 43.705323 | cvor_8805 | Sa povredjenim | 5172.72802578424 | 10 |
| 5 | 5 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8806 | 20.089954 | 43.749813 | cvor_8806 | Sa mat.stetom | 1467.05235267432 | 1 |
| 6 | 6 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8807 | 20.115547 | 43.742251 | cvor_8807 | Sa povredjenim | 876.506924690533 | 10 |
| 7 | 7 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8808 | 20.100155 | 43.754129 | cvor_8808 | Sa povredjenim | 1392.28634302808 | 10 |
| 8 | 8 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8809 | 20.072314 | 43.749596 | cvor_8809 | Sa povredjenim | 2738.3238171266 | 10 |
| 9 | 9 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8810 | 20.1118489 | 43.6262885 | cvor_8810 | Sa povredjenim | 12872.9377650001 | 10 |
| 10 | 10 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8811 | 20.124384 | 43.68567 | cvor_8811 | Sa mat.stetom | 6460.69878678659 | 1 |
| 11 | 11 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8812 | 20.124369 | 43.68584 | cvor_8812 | Sa mat.stetom | 6442.09387018556 | 1 |
| 12 | 12 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8813 | 20.1197222 | 43.7008333 | cvor_8813 | Sa mat.stetom | 4735.53392946749 | 1 |
| 13 | 13 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8814 | 20.091103 | 43.758184 | cvor_8814 | Sa povredjenim | 2101.00849496199 | 10 |
| 14 | 14 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8815 | 20.11001 | 43.733744 | cvor_8815 | Sa povredjenim | 1016.48255951569 | 10 |
| 15 | 15 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8816 | 20.099104 | 43.753399 | cvor_8816 | Sa povredjenim | 1340.0618493184 | 10 |
| 16 | 16 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8817 | 20.118456 | 43.705414 | cvor_8817 | Sa povredjenim | 4217.54079765836 | 10 |
| 17 | 17 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8818 | 20.106489 | 43.671712 | cvor_8818 | Sa mat.stetom | 7814.49705256214 | 1 |
| 18 | 18 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8819 | 20.101321 | 43.774514 | cvor_8819 | Sa mat.stetom | 3618.94707836844 | 1 |
| 19 | 19 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8820 | 20.1055937 | 43.748279 | cvor_8820 | Sa povredjenim | 697.953050238974 | 10 |
| 20 | 20 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8821 | 20.106295 | 43.748205 | cvor_8821 | Sa mat.stetom | 698.134127834698 | 1 |
| 21 | 21 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8822 | 20.125472 | 43.746329 | cvor_8822 | Sa povredjenim | 1742.26031897981 | 10 |
| 22 | 22 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8823 | 20.107983 | 43.730291 | cvor_8823 | Sa povredjenim | 1331.64049795002 | 10 |
| 23 | 23 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8824 | 20.118539 | 43.705321 | cvor_8824 | Sa povredjenim | 4229.27223216578 | 10 |
| 24 | 24 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8825 | 20.105595 | 43.748306 | cvor_8825 | Sa povredjenim | 700.947048177591 | 10 |
| 25 | 25 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8826 | 20.096261 | 43.753235 | cvor_8826 | Sa mat.stetom | 1416.98018276051 | 1 |
| 26 | 26 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8827 | 20.064215 | 43.747979 | cvor_8827 | Sa mat.stetom | 3324.90042978005 | 1 |
| 27 | 27 | 8801 | 20.10467 | 43.742033 | cvor_8801 | 8828 | 20.113217 | 43.750289 | cvor_8828 | Sa poginulim | 1146.90441565791 | 85 |

Слика 14. Преглед израчунатих података

Након одређивања вредности растојања, нормализације података и издвајања четврте најмање вредности за сваку локацију, добија се следеће (Табела 8):

Табела 8. Сортиране вредности четвртог најмањег одстојања за сваку тачку

| локација | вредност |
|------------|----------|
| cvor_36834 | 1 |
| cvor_8810 | 0.647551 |
| cvor_66372 | 0.57621 |
| cvor_36835 | 0.548826 |
| cvor_36816 | 0.520058 |
| cvor_36817 | 0.368375 |
| cvor_66346 | 0.216342 |
| cvor_8803 | 0.194802 |
| cvor_66338 | 0.177685 |
| cvor_36848 | 0.177325 |
| cvor_8805 | 0.174305 |
| cvor_36833 | 0.1715 |
| cvor_36851 | 0.165566 |
| cvor_66365 | 0.164982 |
| cvor_66339 | 0.161833 |
| cvor_36809 | 0.157563 |
| cvor_36803 | 0.151684 |
| cvor_66356 | 0.147963 |
| cvor_66357 | 0.140428 |
| cvor_66345 | 0.13738 |

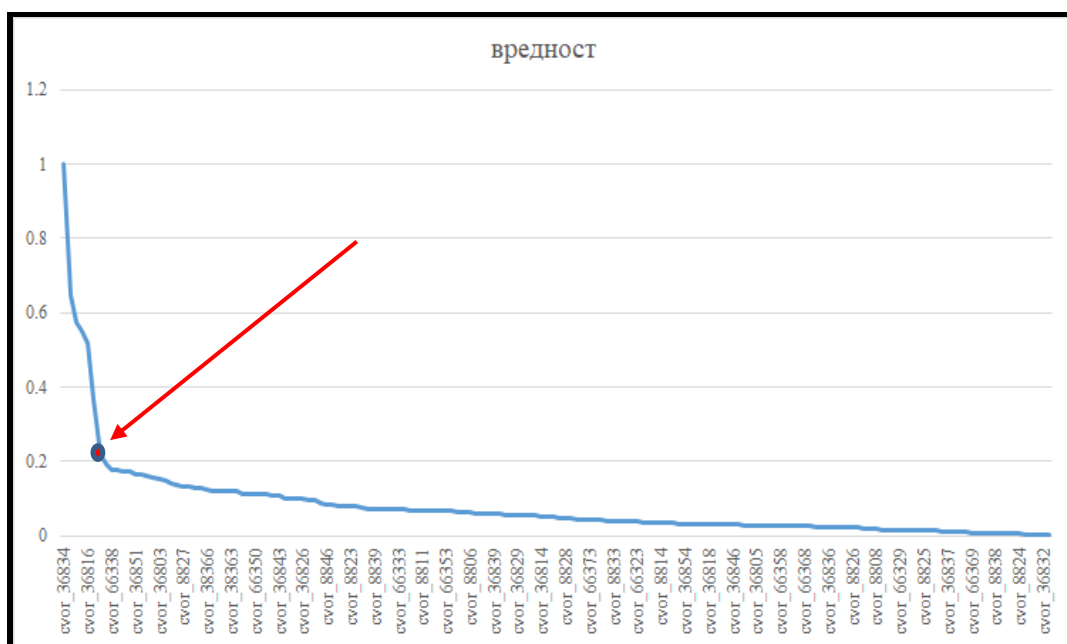
| | |
|------------|----------|
| cvor_8827 | 0.132506 |
| cvor_8822 | 0.131381 |
| cvor_36827 | 0.127331 |
| cvor_8818 | 0.126967 |
| cvor_38366 | 0.124052 |
| cvor_8829 | 0.121546 |
| cvor_36819 | 0.11996 |
| cvor_36828 | 0.119773 |
| cvor_38363 | 0.119478 |
| cvor_66362 | 0.119478 |
| cvor_8840 | 0.114509 |
| cvor_8844 | 0.113757 |
| cvor_66350 | 0.112639 |
| cvor_8809 | 0.112106 |
| cvor_36852 | 0.111589 |
| cvor_36830 | 0.108462 |
| cvor_36843 | 0.10734 |
| cvor_8837 | 0.101938 |
| cvor_66371 | 0.10129 |
| cvor_36845 | 0.100498 |
| cvor_36826 | 0.098501 |
| cvor_8802 | 0.094979 |
| cvor_8807 | 0.094669 |
| cvor_66359 | 0.089604 |
| cvor_8846 | 0.086247 |
| cvor_66330 | 0.084943 |
| cvor_36815 | 0.081738 |
| cvor_66325 | 0.08146 |
| cvor_8823 | 0.080069 |
| cvor_36820 | 0.078843 |
| cvor_66360 | 0.075736 |
| cvor_36811 | 0.073944 |
| cvor_8839 | 0.073724 |
| cvor_8813 | 0.07259 |
| cvor_36841 | 0.072131 |
| cvor_36838 | 0.071368 |
| cvor_66333 | 0.071164 |
| cvor_36822 | 0.071086 |
| cvor_8801 | 0.069698 |
| cvor_36801 | 0.069515 |
| cvor_8811 | 0.069515 |
| cvor_36821 | 0.068516 |
| cvor_36849 | 0.068516 |
| cvor_36853 | 0.068016 |
| cvor_66353 | 0.06749 |
| cvor_8812 | 0.066854 |

| | |
|------------|----------|
| cvor_66344 | 0.064022 |
| cvor_66321 | 0.063927 |
| cvor_8806 | 0.063475 |
| cvor_66322 | 0.061126 |
| cvor_36825 | 0.061033 |
| cvor_8815 | 0.059241 |
| cvor_36839 | 0.059024 |
| cvor_8819 | 0.059024 |
| cvor_66340 | 0.057849 |
| cvor_66343 | 0.05642 |
| cvor_36829 | 0.056153 |
| cvor_8842 | 0.055621 |
| cvor_36813 | 0.055531 |
| cvor_8804 | 0.054755 |
| cvor_36814 | 0.052573 |
| cvor_66334 | 0.050668 |
| cvor_66334 | 0.050668 |
| cvor_66367 | 0.04947 |
| cvor_8828 | 0.048002 |
| cvor_36807 | 0.047952 |
| cvor_66352 | 0.045284 |
| cvor_66341 | 0.043865 |
| cvor_66373 | 0.043267 |
| cvor_8835 | 0.043267 |
| cvor_8843 | 0.043108 |
| cvor_8831 | 0.041281 |
| cvor_8833 | 0.04113 |
| cvor_8847 | 0.039417 |
| cvor_36840 | 0.038887 |
| cvor_66332 | 0.038193 |
| cvor_66323 | 0.037765 |
| cvor_66366 | 0.036705 |
| cvor_36824 | 0.035797 |
| cvor_8834 | 0.034571 |
| cvor_8814 | 0.034234 |
| cvor_36804 | 0.034102 |
| cvor_36847 | 0.034102 |
| cvor_66319 | 0.033457 |
| cvor_36854 | 0.033337 |
| cvor_36812 | 0.033088 |
| cvor_8841 | 0.032546 |
| cvor_36802 | 0.03241 |
| cvor_36818 | 0.031514 |
| cvor_66335 | 0.031514 |
| cvor_66337 | 0.03076 |
| cvor_36806 | 0.030013 |

| | |
|------------|----------|
| cvor_36846 | 0.029867 |
| cvor_66363 | 0.029727 |
| cvor_66326 | 0.029245 |
| cvor_66342 | 0.029067 |
| cvor_36805 | 0.028694 |
| cvor_8845 | 0.028656 |
| cvor_36808 | 0.028452 |
| cvor_66347 | 0.028399 |
| cvor_66358 | 0.027501 |
| cvor_66354 | 0.027285 |
| cvor_66364 | 0.026582 |
| cvor_8836 | 0.026326 |
| cvor_66368 | 0.02604 |
| cvor_66318 | 0.025858 |
| cvor_66370 | 0.025219 |
| cvor_8830 | 0.025187 |
| cvor_36836 | 0.024463 |
| cvor_66348 | 0.024403 |
| cvor_66351 | 0.0242 |
| cvor_66349 | 0.02326 |
| cvor_8826 | 0.02244 |
| cvor_36844 | 0.022131 |
| cvor_66361 | 0.02101 |
| cvor_66327 | 0.020804 |
| cvor_8808 | 0.017386 |
| cvor_36842 | 0.017165 |
| cvor_66324 | 0.016983 |
| cvor_66374 | 0.016983 |
| cvor_66329 | 0.01697 |
| cvor_66331 | 0.016916 |
| cvor_66336 | 0.016875 |
| cvor_8820 | 0.016875 |
| cvor_8825 | 0.016454 |
| cvor_8821 | 0.016206 |
| cvor_36810 | 0.013856 |
| cvor_66328 | 0.013092 |
| cvor_36837 | 0.013073 |
| cvor_66320 | 0.01163 |
| cvor_38364 | 0.009855 |
| cvor_36798 | 0.009731 |
| cvor_66369 | 0.00899 |
| cvor_36823 | 0.00877 |
| cvor_38365 | 0.00877 |
| cvor_8817 | 0.008395 |
| cvor_8838 | 0.008143 |
| cvor_8832 | 0.00806 |

| | |
|------------|----------|
| cvor_36850 | 0.007556 |
| cvor_66355 | 0.007236 |
| cvor_8824 | 0.006821 |
| cvor_36799 | 0.001744 |
| cvor_8816 | 0.001744 |
| cvor_36831 | 0.001717 |
| cvor_36832 | 0.001431 |
| cvor_36800 | 0.001268 |

Графички приказ дат је на Слици 15.



Слика 15. Сортирани 4-dist график и преломна тачка

Преломна тачка је *cvor_66338* са вредношћу 0.177685, $PT(cvor_66338, 0.177685)$.

У наведеном случају, као и у претходном, може се одредити један кластер према претходно дефинисаном поступку у претходном примеру.

6. Детаљан опис математичког модела за одређивање „црних тачака“ (односно деоница, зона)

Математички модел за одређивање просторних кластера може се дефинисати као скуп релација које описују или дефинишу везе између појединих величина у процесу идентификације кластера.

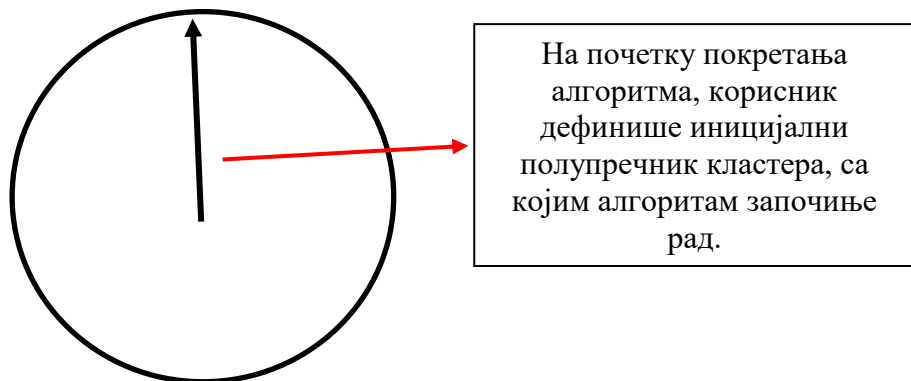
6.1. Формулација математичког модела

Корак 1.

На основу захтева *DBSCAN* алгоритма, његова примена на великим скуповима просторних података захтева одређене предуслове по [105]:

1. Основно знање о области над којом се врше истраживања да би били одређени улазни параметри,
2. Откривање кластера одређеног облика, обзиром да кластери у просторним скуповима могу бити сферични, линеарни, издужени и сл,
3. Добра ефикасност на великим скуповима података, тј. скалабилност.

На основу тачке 1, у модификованом алгоритму познавалац области, тј. корисник дефинише величину кластера који жели да идентификује. У наредним корацима се за сваки од кластера прорачунава сума вредности свих тачака у том кластеру. Сваки пут алгоритам проверава идентификовање и постојање самосталних кластера, који се одмах издвајају. Уколико дође до преклапања кластера, иде се ка њиховом сажимању, док се не добије јединствени кластер или више њих, што је описано у наставку.



Слика 16. Облик кластера који користи модификовани алгоритам

Ево формуле која одговара описаном делу иницијализације у модификованом алгоритму:

$$W(C_i) = \sum_{p \in C_i} w(p)$$

где је:

- $W(C_i)$ — укупна вредност (пондерисана сума) за кластер C_i ,
- $p \in C_i$ — тачке које припадају кластеру,
- $w(p)$ — пондер додељен тачки p .

Правило за валидност кластера:

$$C_i \text{ је самосталан кластер } \Leftrightarrow W(C_i) \geq \theta$$

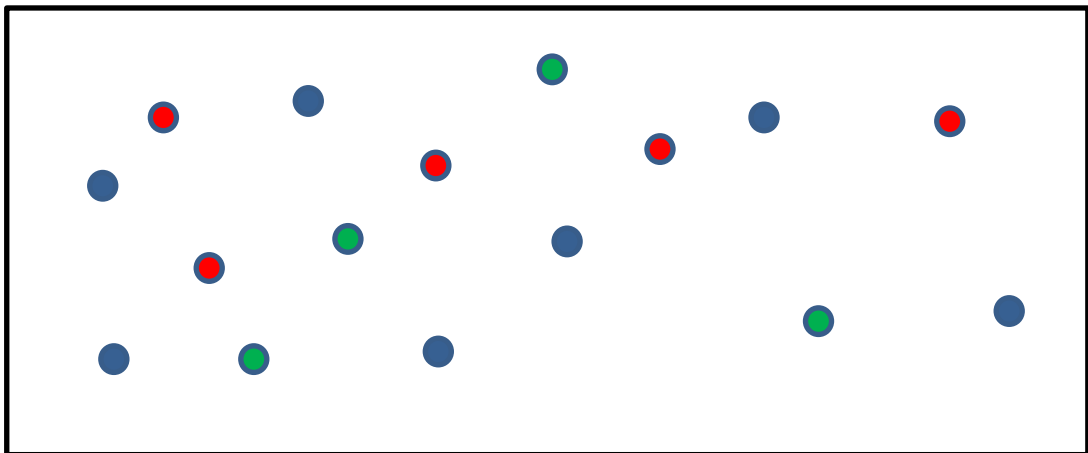
где је θ праг значаја (нпр. 85 за кластер са смртним исходом).

Правило за преклапање кластера:

$$C_{new} = \begin{cases} C_i & \text{ако } W(C_i) > W(C_j) \\ C_i \cup C_j & \text{ако се } C_i \text{ и } C_j \text{ преклапају} \\ & \text{иначе} \end{cases}$$

Корак 2.

За сваку тачку у простору могуће је дефинисати један или више параметара који одређују њену вредност, тј. “тежину”. Одређене текстуалне вредности могу бити конвертоване у нумеричке податке према признатим принципима и конвенцијама.



Слика 17. Преглед примера скупа тачака на које се примењује алгоритам

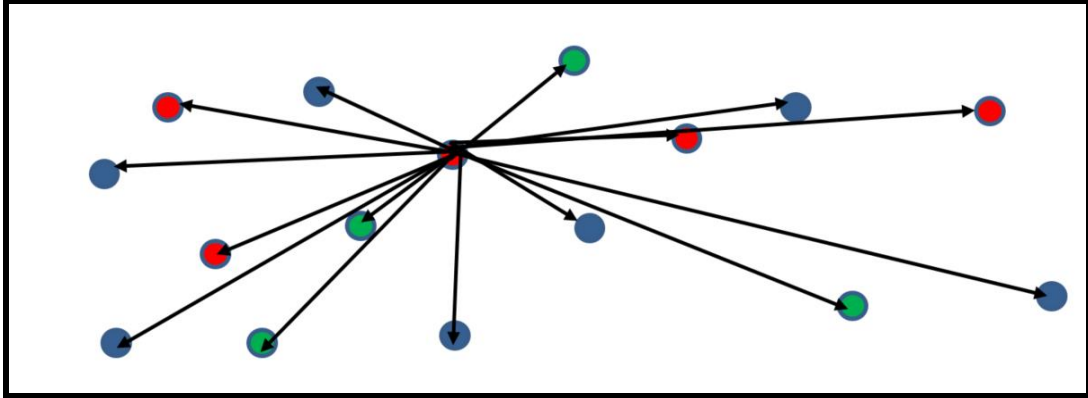
$$w(p) = \sum_{k=1}^m \gamma_k \cdot x_k(p)$$

где је:

- $w(p)$ — тежина тачке p ,
- $x_k(p)$ — вредност параметра k за тачку p (нумерички податак добијен конверзијом из текстуалне вредности),
- γ_k — коефицијент који одређује значај параметра k ,
- m — укупан број параметара који се узимају у обзир.

Корак 3.

За сваку тачку у кластеру рачуна се растојање са свим осталим тачкама у посматраном скупу. Пример за једну тачку приказан је на Слици 18.



Слика 18. Преглед растојања за једну тачку из примера скупа тачака на које се примењује алгоритам

За овај део можемо формално записати поступак рачунања растојања сваке тачке p са свим осталим тачкама q у посматраном скупу D :

$$dist(p,q)=\sqrt{\sum_{k=1}^m(x_k(p)-x_k(q))^2}$$

где је:

- $dist(p,q)$ — Евклидово растојање између тачака p и q ,
- $x_k(p)$ — вредност параметра k за тачку p ,
- m — број параметара (димензија) који описују тачку.

Затим се за сваку тачку p дефинише скуп њених суседа:

$$N(p)=\{q \in D \mid dist(p,q) \leq \varepsilon\}$$

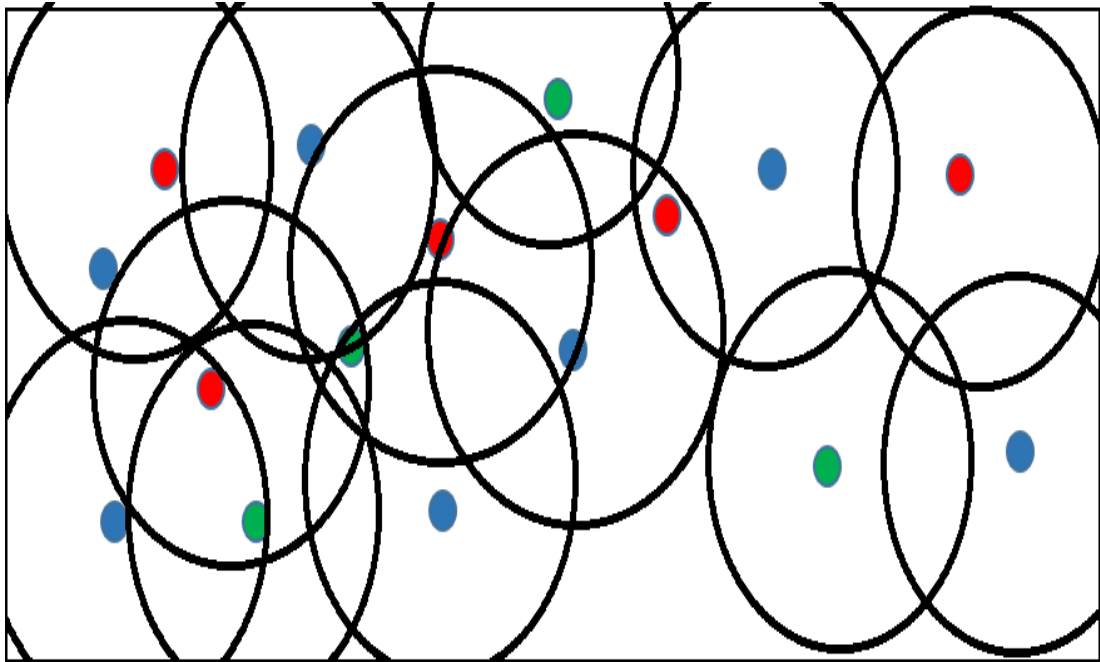
Корак број 4.

За сваку од тачки врши се прорачун суме вредности унутар иницијално дефинисаног кластера.

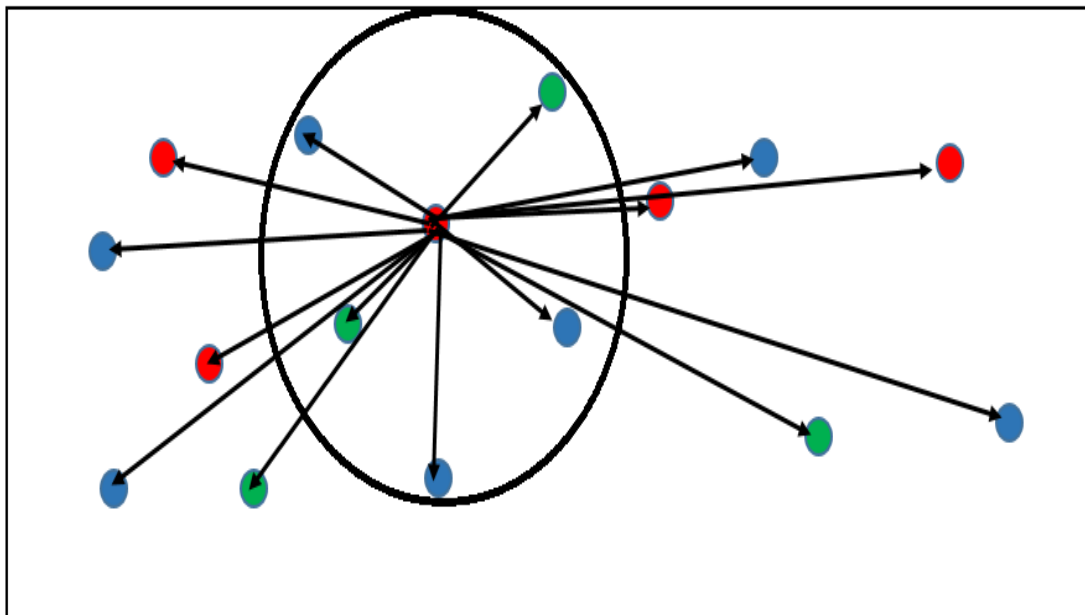
$$W(C_i)=\sum_{p \in C_i} w(p)$$

где је:

- $W(C_i)$ — укупна пондерисана вредност кластера C_i ,
- $p \in C_i$ — све тачке које припадају кластеру,
- $w(p)$ — тежина (пондер) додељена тачки p .



Слика 19. Преглед кластера за сваку од тачки из примера скупа тачака на које се примењује алгоритам



Слика 20. Преглед тачки обухваћених једним од кластера из примера скупа тачака на које се примењује алгоритам

Након прорачуна свих вредности, врши се издвајање оних кластера који задовољавају услов да је сума вредности унутар њих већа од задате граничне вредности k .

$$C_i \text{ је валидан кластер} \Leftrightarrow W(C_i) = \sum_{p \in C_i} w(p) \geq k$$

где је:

- C_i — кластер,

- $W(C_i)$ — укупна пондерисана вредност кластера,
- $w(p)$ — тежина тачке p ,
- k — задата гранична вредност (праг значаја).

Корак 5.

Наставак алгоритма подразумева рад са издвојеним кластерима. Вредности издвојених кластера су сортиране у нерастућем редоследу, тако да алгоритам наставља рад са кластерима који имају највећу израчунату вредност, па са мањом вредношћу, док не заврши рад са свим кластерима.

C_1, C_2, \dots, C_n тако да важи $W(C_1) \geq W(C_2) \geq \dots \geq W(C_n)$

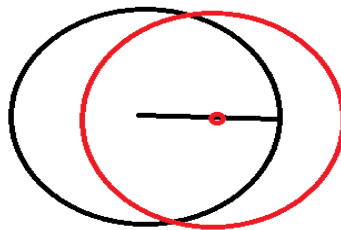
У овом кораку следи издвајање и одбацивање одређених кластера.

Услов за задржавање кластера:

$$C_i \text{ се задржава} \Leftrightarrow W(C_i) \geq k$$

У великом броју случајева добијени кластери неће бити јединствени и самостални. Чест случај је да се два кластера преклапају, тако да је потребно одредити који је кластер самосталнији и њега изабрати, а други одбацити. У предложеном алгоритму предвиђена су два сценарија.

- а) *Центар једног издвојеног кластера налази се у опсегу полупречника другог издвојеног кластера.*

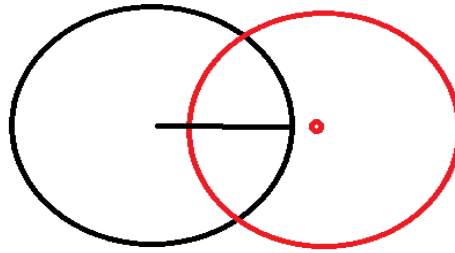


Слика 21. Приказ сценарија када се два кластера преклапају, а растојање њихових центроида је мање од полупречника кластера

У овом случају кластер са већом вредношћу наставља са радом, а кластер са мањом вредношћу се одбацује. Број прорачунатих “унутрашњих кластера” се инкрементира за 1. Овај број се узима у следећем кораку као један од критеријума за сортирање кластера по вредности.

$$\text{dist}(\text{center}(C_i), \text{center}(C_j)) \leq r_j \Rightarrow \begin{cases} C_i & \text{ако } W(C_i) > W(C_j) \\ C_j & \text{иначе} \end{cases}$$

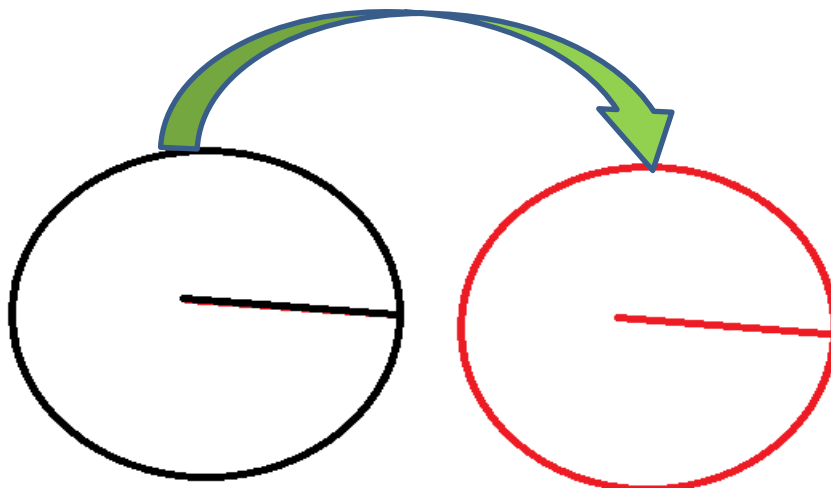
b) Два кластера се преклапају, али је центар једног издвојеног кластера ван опсега полупречника другог издвојеног кластера.



Слика 22. Приказ сценарија када се два кластера преклапају, а растојање њихових центроида је веће од полупречника кластера

У овом случају кластер са већом вредношћу наставља са радом, а кластер са мањом вредношћу се одбацује. Број прорачунатих “спољних кластера” се инкрементира за 1. Овај број се узима у следећем кораку као један од критеријума за сортирање кластера по вредности. Такође разлика између центара кластера умањена за вредност полупречника се додаје на декларисану суму. Овај параметар се на крају узима за рад у наредном кораку алгоритма, где се иницијално задати полупречник инкрементује за прорачунату вредност. Циљ је да се на крају добију највећи могући јединствени кластери.

c) Уколико се два кластера не преклапају, престаје се са радом до тог тренутка коришћеног кластера. Ово је тренутак када алгоритам започиње са радом са следећим у низу самосталних кластера и поступак се понавља док се не дође до следећег самосталног кластера.



Слика 23. Приказ сценарија када се два кластера не преклапају

Када алгоритам заврши рад са свим кластерима, врши се сортирање добијених вредности у неоппадајућем редоследу. Као критеријуми се узимају вредности пондерисаних вредности, број унутрашњих кластера и број спољашњих кластера, просечна вредност растојања описаног у тачки b приказане процедуре.

C_1, C_2, \dots, C_n тако да важи $(W(C_1), I(C_1), O(C_1), d(C_1)) \geq (W(C_2), I(C_2), O(C_2), d(C_2)) \geq \dots$

где је:

- $W(C_i)$ — пондерисана вредност кластера,
- $I(C_i)$ — број унутрашњих кластера,
- $O(C_i)$ — број спољашњих кластера,
- $d(C_i)$ — просечна вредност растојања из сценарија (b).

Након тога се полупречник кластера увећава за прорачунату вредност првог кластера у низу из претходног корака.

$$r_{new}(C_1) = r(C_1) + \Delta r(C_1)$$

Корак 2 се понавља све док се не добију јединствени кластери највећег полупречника.

Математички доказ побољшаног рада *DBSCAN* алгоритма модификованим алгоритмом креће од полазне дефиниције да се стандардни *DBSCAN* кластер дефинише само на основу броја тачака:

$$|N(p)| \geq MinPts$$

Модификовани *DBSCAN* кластер дефинише се на основу пондерисане суме вредности:

$$W(C_i) = \sum_{p \in C_i} w(p) \geq k$$

где је тежина тачке:

$$w(p) = \alpha \cdot S(p) + \beta \cdot F(p)$$

Нека је информативност кластера:

$$I(C) = \frac{W(C)}{|C|}$$

- За *DBSCAN*:

$$I_{DBSCAN}(C) = I$$

- За модификовани *DBSCAN*:

$$I_{MW}(C) = \frac{\sum_{p \in C} w(p)}{|C|}$$

Доказ: Ако постоји бар једна тачка са тежином $w(p) > I$, добија се:

$$I_{MW}(C) \geq I_{DBSCAN}(C)$$

што значи да модификовани *DBSCAN* кластер садржи више информација о значају тачака него *DBSCAN*.

По питању дискриминисаности у *DBSCAN*-у све тачке имају једнаку тежину, па важи:

$$W_{DBSCAN}(C) = |C|$$

У модификованом *DBSCAN*-у:

$$W_{MW}(C) = \sum_{p \in C} w(p)$$

Доказ: Ако постоје две тачке p_1, p_2 са различитим тежинама $w(p_1) \neq w(p_2)$, онда:

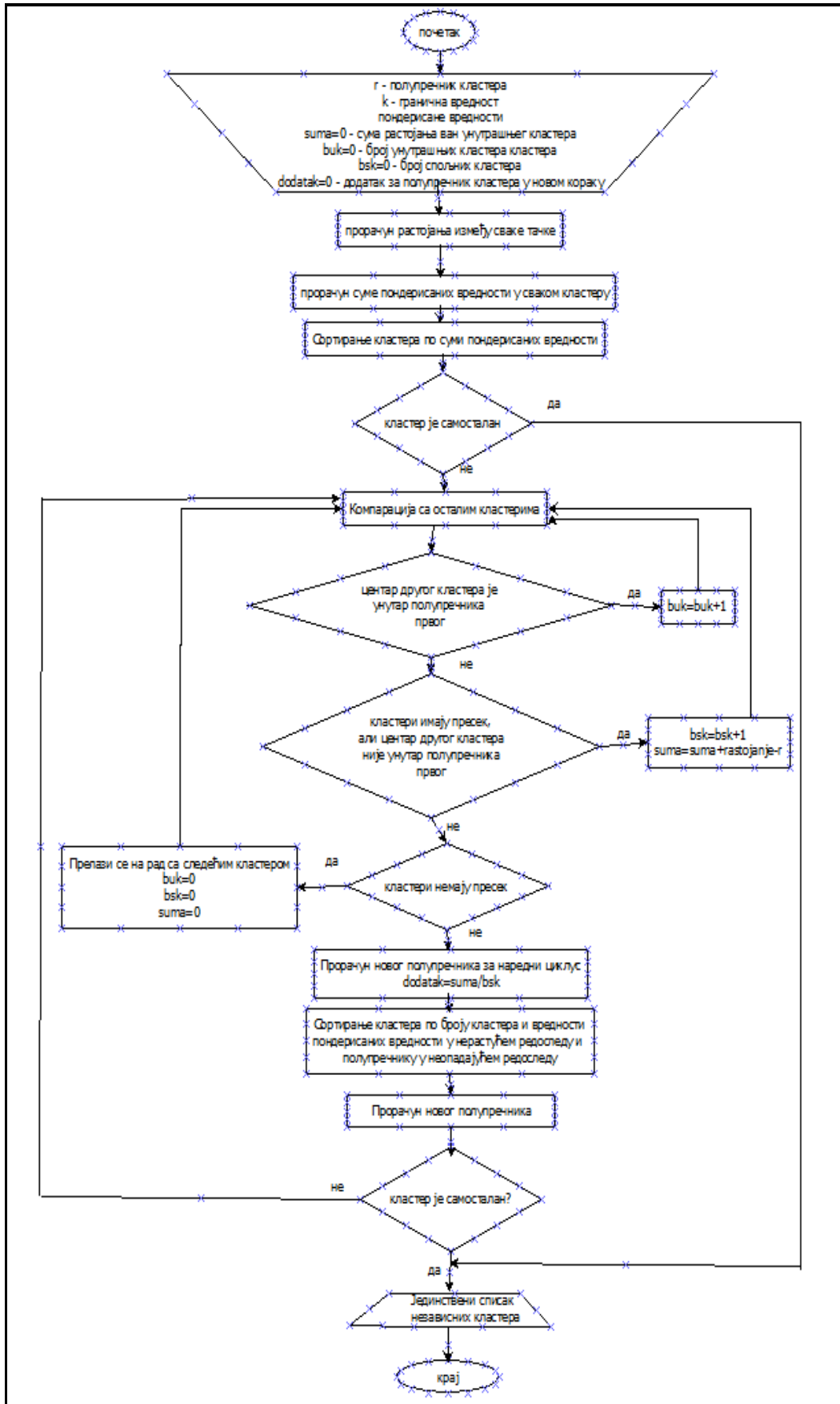
$$W_{MW}(C) \neq W_{DBSCAN}(C)$$

што показује да модификовани *DBSCAN* може да разликује кластере по значају, а не само по броју тачака.

Практични ефекат

- *DBSCAN* добро ради на великим скуповима података, али не разликује тежину појединачних тачака.
- Модификовани *DBSCAN* додаје пондерисање, па кластери одражавају стварни ризик (нпр. тежина незгода у садејству са учесталашћу).
- То значи да добијени кластери нису само статистички, већ и **значајнији** за анализу и употребу.

Листа корака рада модификованог алгоритма приказана је на Слици 24.



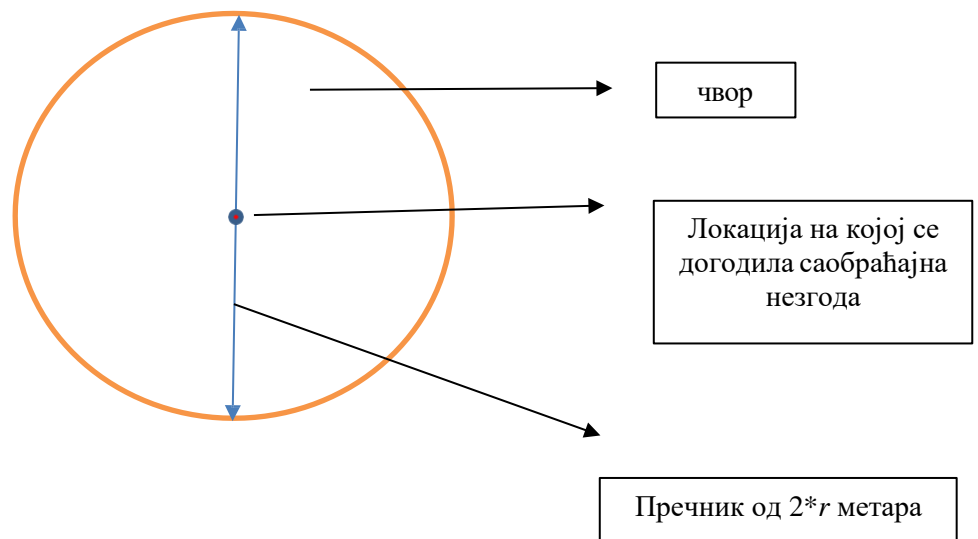
Слика 24. Модификовани алгоритам

6.2. Примена модификованог алгоритма на примеру са саобраћајним незгодама

У случају саобраћајних незгода и идентификацији “црних тачака” почетна вредност полупречника је 300 метара, а гранична вредност суме пондерисаних вредности 85. То је случај да се у кластеру догодила минимум једна саобраћајна незгода са погинулим или девет саобраћајних незгода са повређеним лицима.

Вредност ПБСН у чвору се израчунава тако да се број најтежих саобраћајних незгода (Бр. СН ПОГ) у чвору множи са коефицијентом (пондером) 85, број саобраћајних незгода са повређеним лицима (Бр. СН ПЛ) се множи са коефицијентом (пондером) 10, док се број саобраћајних незгода са материјалном штетом (Бр. СН МШ) множи са пондером 1.

Наведене вредности пондера су укупни социо-економски трошкови саобраћајних незгода добијени на основу истраживања спроведених у Великој Британији.

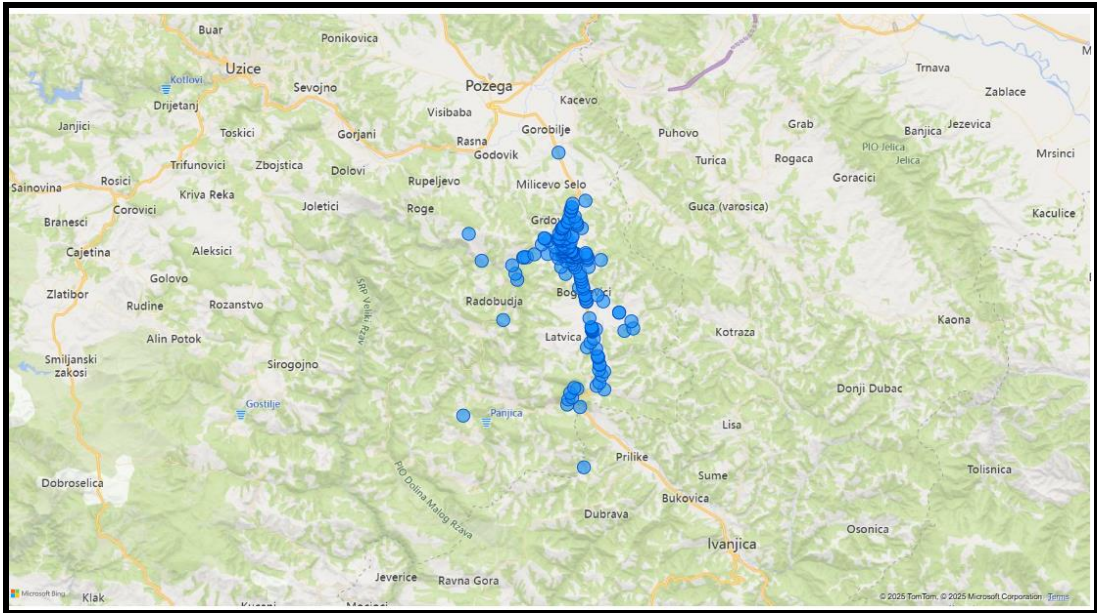


Слика 25. Визуелна презентација изгледа једног чвора

Затим следи провера да ли је израчуната вредност ПБСН у чвору већа или једнака од граничне вредности k .

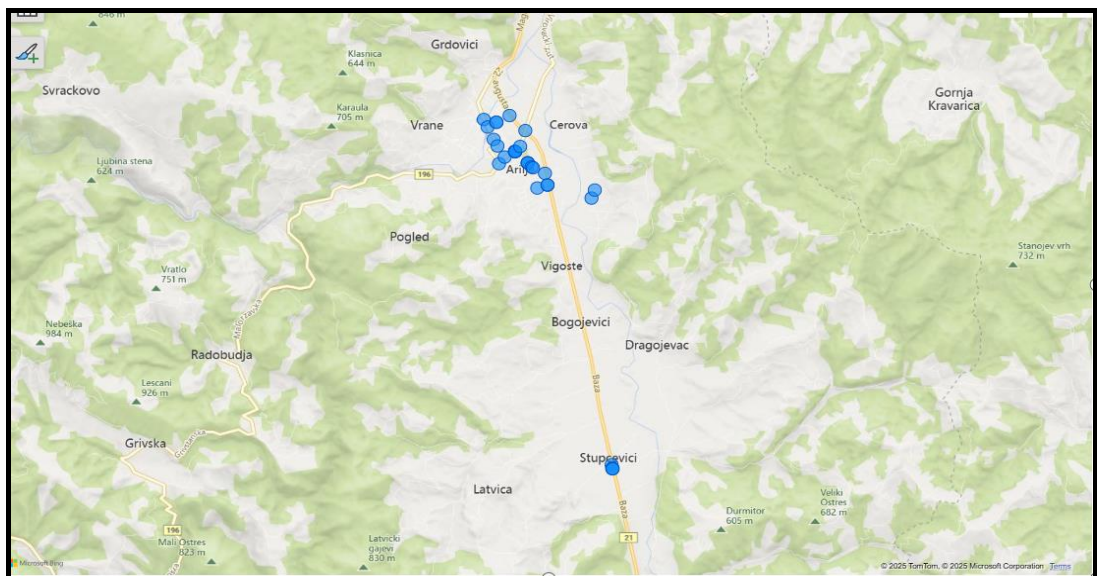
Гранична вредност k је предефинисана пре пуштања алгоритма у рад. У случају да је вредност ПБСН већа или једнака k , алат даље наставља процес издвајања потенцијалних опасних места, у супротном, чвор се одбацује за даљу

анализу у процесу преидентификације „црних тачака“. У предложеном алгоритму, за вредност k је примењена вредност већа од 85, што значи да се у чвору налази најмање једна саобраћајна незгода са настрадалим лицем, или 9 саобраћајних незгода са повређеним лицима. На Слици 26 приказан је скуп података о саобраћајним незгодама са којима алгоритам стартује рад.



Слика 26. Скуп података са којим алгоритам започиње рад

Након прорачуна свих пондерисаних вредности у кластеру задатог иницијалног полупречника, издвајају се центроиди само оних кластера чија је сума пондерисаних вредности већа од задатог k .



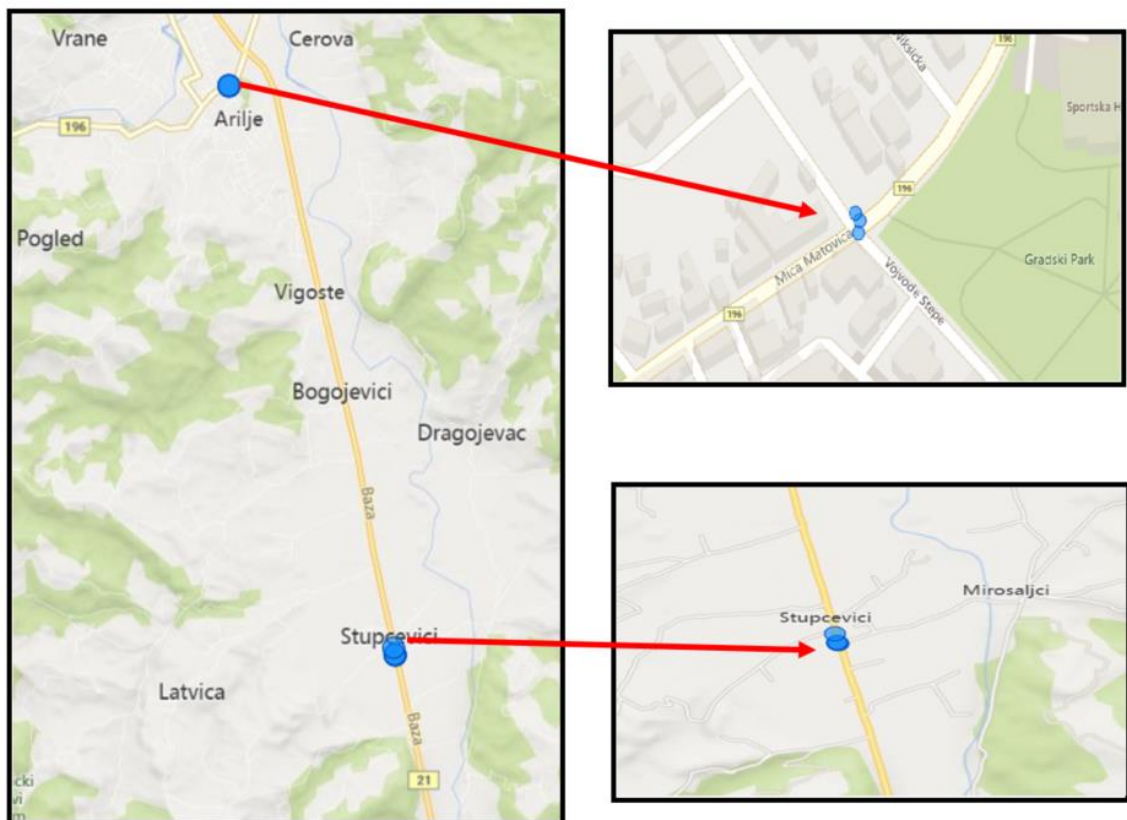
Слика 27. Центроиди издвојених кластера који задовољавају иницијалне критеријуме

Након тога следи провера да ли се сваки од издвојених чворова преклапа

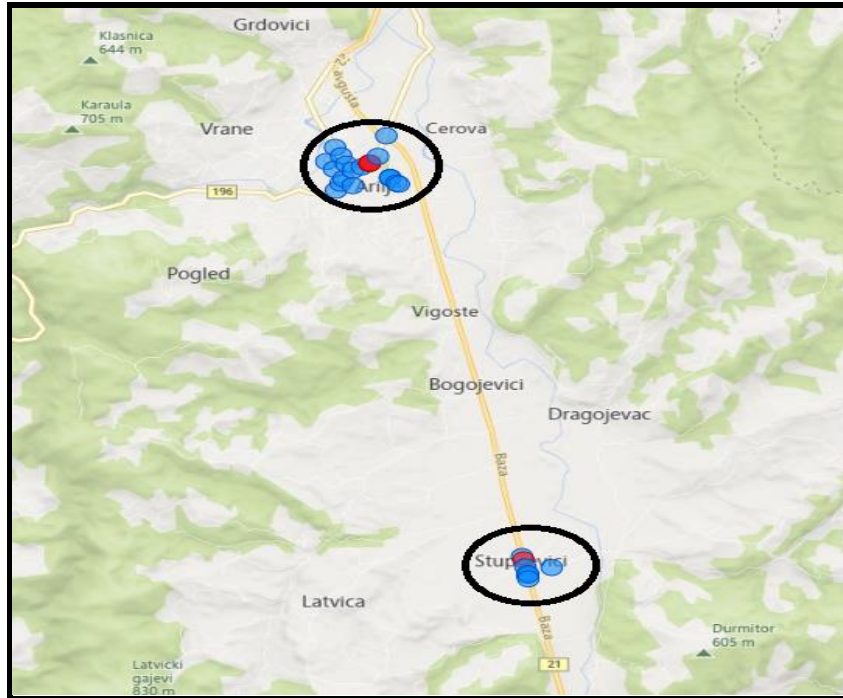
или додирује са једним или више чворова код којих је вредност ПБСН већа или једнака од k .

У првом дефинисаном кораку, као почетне локације од којих се полази узимају се локације на којима се догодила саобраћајна незгода са погинулим или повређеним лицима. Није редак случај да се чворови „описани“ око саобраћајне незгоде преклапају или додирују. Из тог разлога је важно да алат (алгоритам) препозна такве могућности и практично одреди који је чвор од веће вредности. Чворови инфериорније вредности се одбацују, а наставља се рад са супериорнијим. На тај начин спречено је понављање током процеса филтрирања, а чворови се концентришу са циљем идентификације изолованих јединствених, самосталних и највећих кластера.

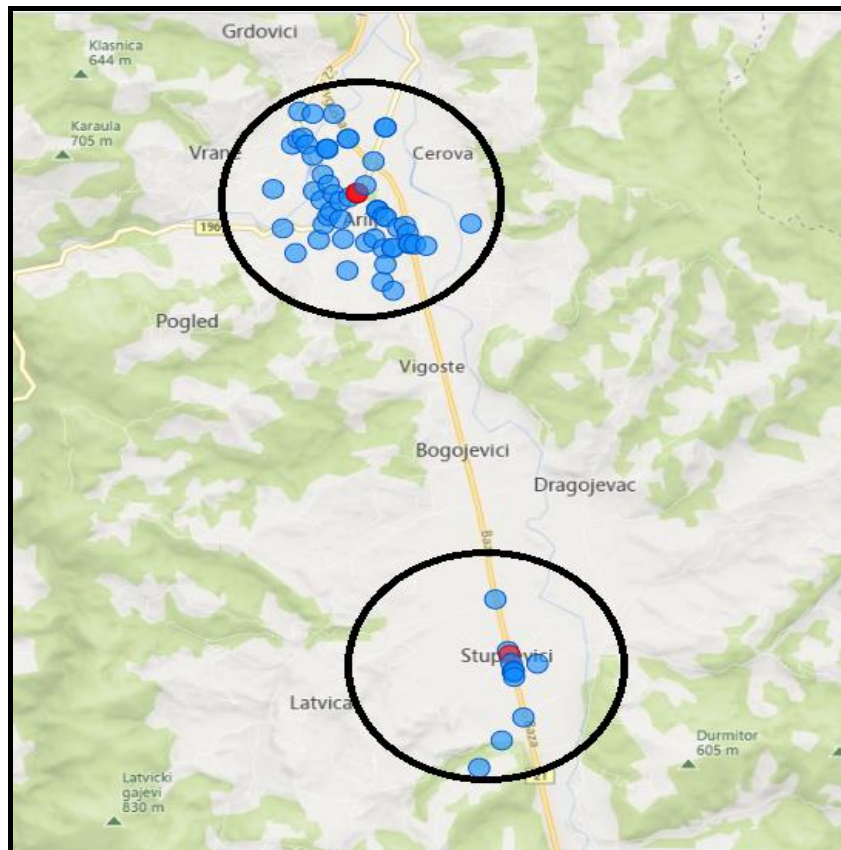
Резултати рада модификованог алгоритма, након компарације и прорачуна издвојених кластера, приказани су на сликама 28-33. Може се констатовати да је број потенцијалних центроида смањен у односу на претходни корак, а њихов полупречник повећан, што и јесте циљ рада алгоритма.



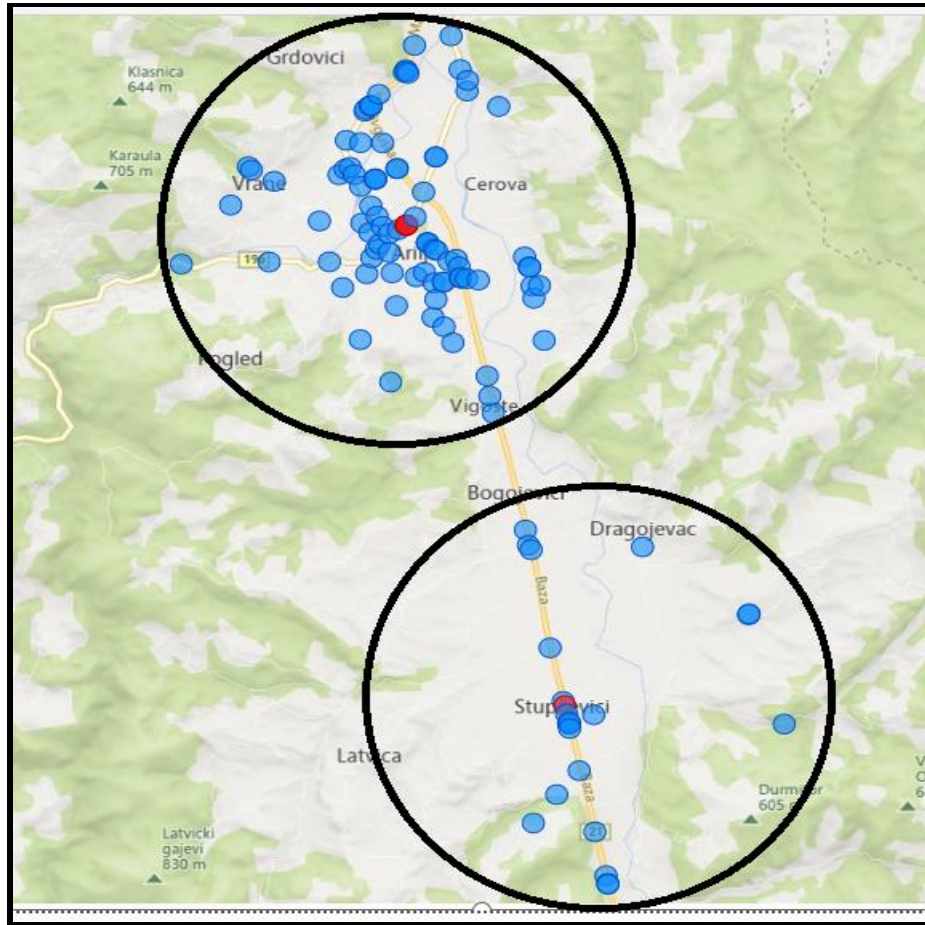
Слика 28. Резултати добијени модификованим алгоритмом након сажимања издвојених кластера



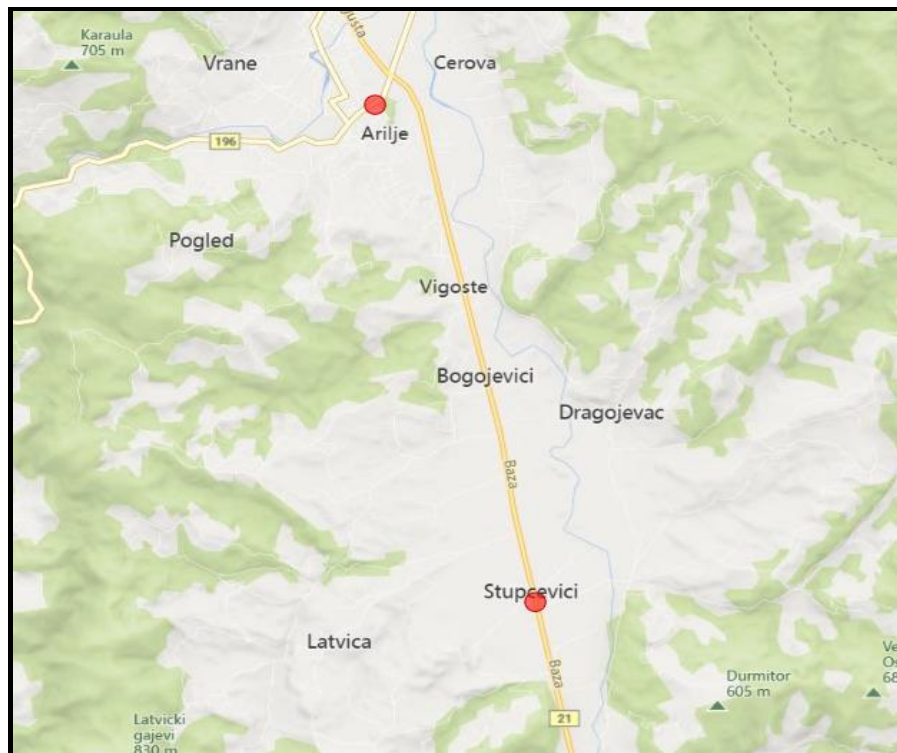
Слика 29. Први издвојени независни кластери



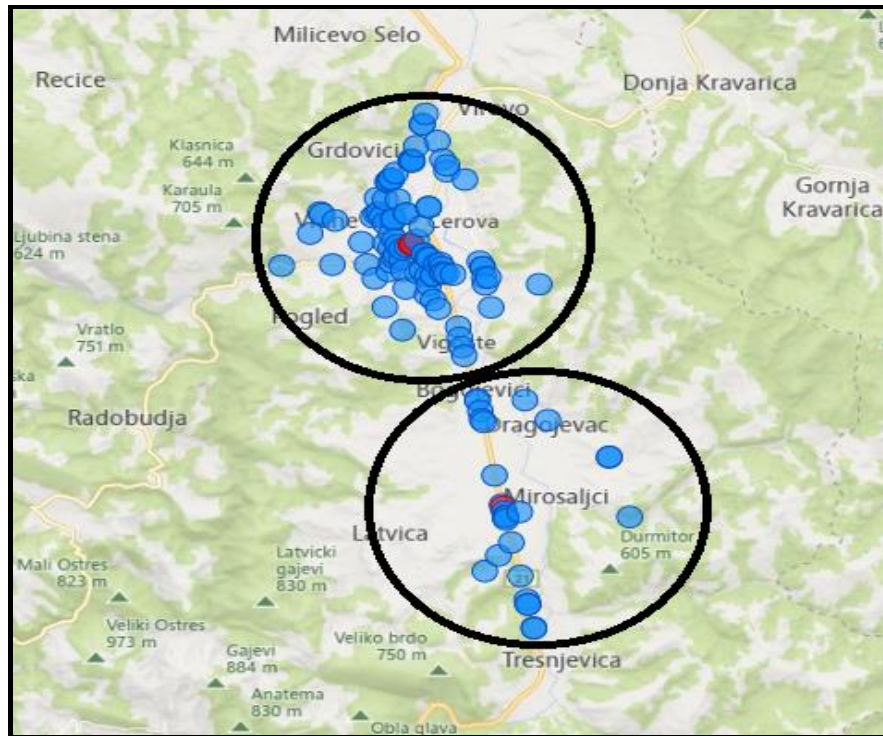
Слика 30. Кластери добијени у следећој итерацији са проширеним полупречником



Слика 31. Кластери добијени у следећој итерацији са проширеним полупречником



Слика 32. Коначни центроиди добијени у последњем кораку рада модификованог алгоритма



Слика 33. Преглед добијених кластера радом модификованог алгоритма

6.3. Компаративна анализа добијених резултата

Употребом предложеног модификованог алгоритма добијају се два кластера, за разлику од стандардног, који генерише само један кластер. На тај начин могу се добити праведније одлуке, а све наведено може се користити у идентификацији црних тачака у саобраћају.

7. Архитектура и конструкција окружења за идентификацију “црних тачака” на путевима у Републици Србије коришћењем модификованог алгоритма и система пословне интелигенције

У савременом окружењу, пословни процеси и научни инструменти могу лако генерисати количине података које се мере у терабајтима. Брзо се шири јаз између могућности прикупљања података и способности анализе података како у бизнису, науци тако и у подацима који се односе на живот и рад сваког појединца, укључујући и саобраћајне незгоде.

На најнижем концептуалном нивоу, налази се база података, у којој се бележе сви релевантни подаци о природи проблема који се решава. Нарастањем пословних података из једног временског интервала у други, стварају се предуслови за развој и пројектовање адекватног складишта података. Добро пројектовано складиште података служи као основа од које све полази и подпомаже процесе доношења одлука, а не замењује доносиоца одлуке. Прикупљање података кроз фазе развоја неког система, карактерише велика количина података [127]. Тако база података пословног система нараста по свом обиму, и у циљу анализе података, ствара се потреба за креирањем складишта података, које прати пословање истог. Како се за стратегијско одлучивање, углавном не користе свакодневни оперативни резултати пословања система, већ збирни и изведени из предходног временског периода, складиште података представља добру основу за пословно одлучивање. Обзиром да ниво стратегијског одлучивања најчешће прате неструктурирани проблеми, иде на руку да складиште података постане добра основа за савремено одлучивање. Складиште података као новији технолошки концепт, има управо улогу да обједини сродне податке виталних функција система у облик погодан за спровођење разних анализа. Тако се насупрот свакодневной оперативной обради података јавља потреба за аналитичком обрадом података.

Једна од неизоставних техника која прати све етапе претраживања података од чишћења, преко препроцесирања и претраживања података је визуелизација података. Визуелизација података може на једноставан и ефикасан начин да укаже на основне смернице за даљу анализу у било којој етапи истраживања. Појам визуелизације се најчешће објашњава као “ментална слика”. У техничком смислу она се бави приказивањем понашања односно

стања, посебно при комплексним условима, на начин разумљив људском оку.

Визуализацију података могуће је посматрати са неколико аспеката:

- визуализација сегмента неочишћених података,
- визуализација очишћених података,
- визуализација препроцесираних података.

За потврду оправданости увођења и примене предложеног и модификованог алгоритма, у циљу унапређења процеса идентификације опасних места на путевима у Републици Србији, користе се јавно доступни подаци, обзиром да се ти догађаји свакодневно дешавају, база података стално расте, а подаци се могу лако трансформисати и користити у оквиру система заснованих на *SQL* бази података. Поред уобичајених захтева за анализу самих саобраћајних незгода, у комбинацији и са другим подацима који су такође јавно доступни, нпр. подаци о временским условима, могу се постављати и други, сложенији захтеви.

Саобраћајне незгоде, нажалост, нису тако ретки догађаји и дешавају се свакодневно. Поред тога, могуће су на целокупној путној инфраструктури, како просторно, тако и временски, што отежава предвиђање где ће се десити. Предложена методологија као циљ има идентификацију и визуелни приказ опасних локација, а са циљем да доносиоцима одлука омогући да у реалном времену доносе одлуке. На пример, на основу овако добијених информација, могу праведније и боље распоредити људство на терену, побољшати доступност екипа и спасилачких капацитета.

7.1. Прикупљање захтева

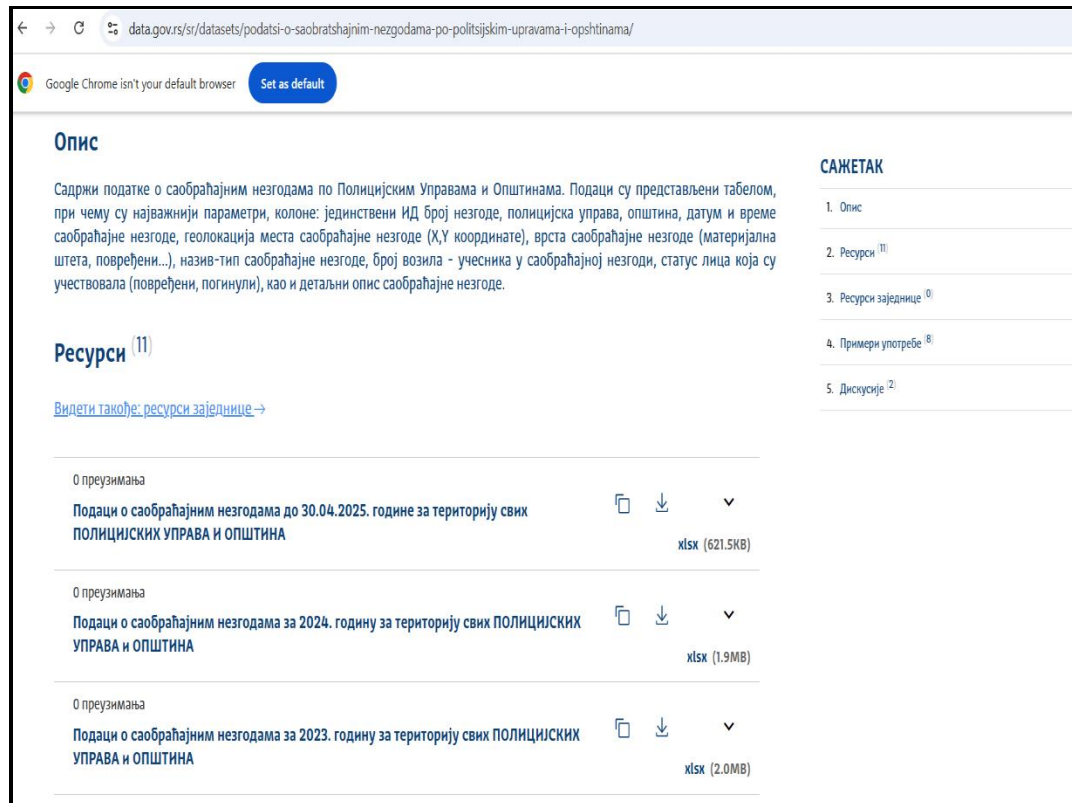
Прикупљање захтева представља битан моменат у развоју система пословне интелигенције. У овој дисертацији коришћена је техника истраживања.

Истраживање је вршено тако што су прикупљени сви јавно доступни подаци који се односе на податке о саобраћајним незгодама. У тексту који следи су дати прегледи прикупљених података.

7.2. Прикупљање података

Први корак при изради складишта података, које служи као полазна

основа за добијање одговора у систему пословне интелигенције, је прикупљање података [128]. Подаци су преузети са портала отворених података Републике Србије. Тема са које су преузети подаци везана је за Јавну безбедност, одељак Подаци о Саобраћајним незгодама по полицијским управама и општинама. На Слици 34 приказан је изглед *web* странице портала Отворених података Републике Србије, са кога су преузети подаци.



Слика 34. Изглед веб стране Портала отворених Података Републике Србије

Скупови података на порталу отворених података дати су у *.xlsx* формату, који је изузетно погодан за трансформацију, сажимање и анализу, а доступан је и за конверзију у многе друге формате, погодне за анализу података.

На порталу се могу пронаћи фајлови који садрже податке о саобраћајним незгодама на територији Републике Србије, по полицијским управама и општинама за временски период од 2015. године до 2024. године. Подаци су представљени табеларно, при чему су најважнији параметри: јединствени ИД број незгоде, датум и време саобраћајне незгоде, геолокација места саобраћајне незгоде, врста саобраћајне незгоде (материјална штета, повређени...), назив/тип саобраћајне незгоде, број возила - учесника у

Архитектура и конструкција окружења за идентификацију "црних тачака" на путевима у Републици Србије коришћењем модификованог алгоритма и система пословне интелигенције

саобраћајној незгоди, статус лица која су учествовала (повређени, погинули и др.), као и детаљни опис саобраћајне незгоде.

Изглед једне табеле, приказан је на Слици 35.

| A | B | C | D | E | F | G | H | I |
|---------------|---------|----------|------------------|-----------|------------|---------------|--|--|
| identifikator | grad | opstina | vreme | duzina | sirina | vrsta | tip | opis |
| 1351381 | BEograd | BARAJEVO | 01.01.2022.02:30 | 0040508 | 44.58907 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa jednim vozilom - sližak uleno sa kolovoza na pravcu |
| 1351384 | BEograd | BARAJEVO | 01.01.2022.16:10 | 0034629 | 44.60308 | sa pornejanom | SI SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa desne strane kolovoza |
| 1352261 | BEograd | BARAJEVO | 06.01.2022.12:10 | 00414231 | 44.57167 | sa mat.stomom | SI SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa desne strane kolovoza |
| 1352171 | BEograd | BARAJEVO | 06.01.2022.11:00 | 0039882 | 44.60251 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa jednim vozilom - sližak uleno sa kolovoza na pravcu |
| 1351871 | BEograd | BARAJEVO | 06.01.2022.20:15 | 00344523 | 44.602351 | sa mat.stomom | SI SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa desne strane kolovoza |
| 1352208 | BEograd | BARAJEVO | 06.01.2022.17:30 | 00392748 | 44.633833 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa jednim vozilom - sližak uleno sa kolovoza na pravcu |
| 1353199 | BEograd | BARAJEVO | 16.01.2022.22:00 | 0033927 | 44.59985 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa jednim vozilom - sližak uleno sa kolovoza na pravcu |
| 1353448 | BEograd | BARAJEVO | 17.01.2022.11:00 | 0037807 | 44.59178 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila - suprotni smerovi bez skretanja - krivina udes |
| 1354597 | BEograd | BARAJEVO | 22.01.2022.12:15 | 00334319 | 44.583252 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - SKRETANJE LI PRAZAM | Najmanje dva vozila koja se kreću u istom smeru u suprotnim smerovima u skretanje uleno ispred drugog vozila |
| 1354978 | BEograd | BARAJEVO | 22.01.2022.15:20 | 0030762 | 44.581877 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - uključivanje u saobraćaj |
| 1356697 | BEograd | BARAJEVO | 13.02.2022.14:10 | 0030219 | 44.590683 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - sustizanje |
| 1356844 | BEograd | BARAJEVO | 15.02.2022.18:20 | 0041353 | 44.59901 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - SKRETANJE LI PRAZAM | Najmanje dva vozila koja se kreću u istom smeru u suprotnim smerovima u skretanje ispred drugog vozila |
| 1359224 | BEograd | BARAJEVO | 17.02.2022.13:15 | 0041261 | 44.637693 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - sustizanje |
| 1357344 | BEograd | BARAJEVO | 17.02.2022.11:30 | 00373661 | 44.623001 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - SKRETANJE LI PRAZAM | Najmanje dva vozila koja se kreću u istom smeru u istom smeru u skretanje, skretanje uleno ispred drugog vozila |
| 1357393 | BEograd | BARAJEVO | 17.02.2022.22:40 | 00375311 | 44.615009 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa učešćem jednog vozila i životinja |
| 1357290 | BEograd | BARAJEVO | 19.02.2022.07:35 | 0041167 | 44.57714 | sa mat.stomom | SI SA PARKIRANIM VOZILIMA | Ostali sudar sa parkiranim vozilom |
| 1359703 | BEograd | BARAJEVO | 19.02.2022.18:50 | 0040030 | 44.49628 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Ostale negode sa najmanje dva vozila - suprotni smerovi bez skretanja |
| 1358440 | BEograd | BARAJEVO | 23.02.2022.07:10 | 0041411 | 44.57688 | sa mat.stomom | SI SA PARKIRANIM VOZILIMA | Ostali sudar sa parkiranim vozilom |
| 1359393 | BEograd | BARAJEVO | 28.02.2022.20:35 | 0048901 | 44.623056 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - SKRETANJE LI PRAZAM | Najmanje dva vozila koja se kreću različitim putevima u prolazak kroz raskrsnicu, ili od kojih jedno prelazi preko kolovoza, bez skretanja |
| 1361888 | BEograd | BARAJEVO | 01.03.2022.13:50 | 0041703 | 44.55149 | sa pornejanom | SI SA JEDNIM VOZILOM | Negodna sa jednim vozilom - sližak uleno sa kolovoza na pravcu |
| 1360102 | BEograd | BARAJEVO | 03.03.2022.16:00 | 0041207 | 44.63793 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - sustizanje |
| 1359799 | BEograd | BARAJEVO | 04.03.2022.23:10 | 0041487 | 44.51706 | sa mat.stomom | SI SA PARKIRANIM VOZILIMA | Ostali sudar sa parkiranim vozilom |
| 1360831 | BEograd | BARAJEVO | 08.03.2022.13:25 | 0041499 | 44.58226 | sa mat.stomom | SI SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa desne strane kolovoza |
| 1361160 | BEograd | BARAJEVO | 11.03.2022.19:50 | 004268 | 44.57323 | sa mat.stomom | SI SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa leve strane kolovoza |
| 1362751 | BEograd | BARAJEVO | 13.03.2022.13:50 | 0041703 | 44.55149 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa učešćem jednog vozila i životinja |
| 1362744 | BEograd | BARAJEVO | 18.03.2022.09:00 | 00420296 | 44.626733 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa učešćem jednog vozila i životinja |
| 1363905 | BEograd | BARAJEVO | 20.03.2022.10:40 | 00415851 | 44.552871 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Ostale negode sa najmanje dva vozila - suprotni smerovi bez skretanja |
| 1361873 | BEograd | BARAJEVO | 20.03.2022.14:45 | 00354485 | 44.553356 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Ostale negode sa najmanje dva vozila - suprotni smerovi bez skretanja |
| 1363808 | BEograd | BARAJEVO | 20.03.2022.10:20 | 00415851 | 44.552871 | sa pornejanom | SI SA PRAZAM | Prelazak preko kolovoza suprotno od smeru kretanja vozila |
| 1364938 | BEograd | BARAJEVO | 21.03.2022.11:10 | 00429314 | 44.62313 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Ostale negode sa najmanje dva vozila - suprotni smerovi bez skretanja |
| 1365402 | BEograd | BARAJEVO | 24.03.2022.13:15 | 0041309 | 44.644936 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - SKRETANJE LI PRAZAM | Najmanje dva vozila koja se kreću u istom smeru u istom smeru u skretanje uleno ispred drugog vozila |
| 1363571 | BEograd | BARAJEVO | 25.03.2022.08:45 | 00388517 | 44.633132 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - SKRETANJE LI PRAZAM | Najmanje dva vozila koja se kreću u istom smeru u suprotnim smerovima u skretanje uleno ispred drugog vozila |
| 1362194 | BEograd | BARAJEVO | 27.03.2022.07:10 | 004182 | 44.5768 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - uključivanje u saobraćaj |
| 1365011 | BEograd | BARAJEVO | 28.03.2022.13:15 | 0041824 | 44.57819 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa jednim vozilom bez prepreka na kolovozu na neposred, neposredan način |
| 1365501 | BEograd | BARAJEVO | 06.04.2022.19:15 | 004122441 | 44.63719 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa jednim vozilom na kolovozu |
| 1365112 | BEograd | BARAJEVO | 10.04.2022.15:00 | 0035421 | 44.60092 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - sustizanje |
| 1366145 | BEograd | BARAJEVO | 16.04.2022.13:40 | 0041118 | 44.638444 | sa mat.stomom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - sustizanje |
| 1366198 | BEograd | BARAJEVO | 16.04.2022.15:00 | 00405908 | 44.6304145 | sa mat.stomom | SI SA JEDNIM VOZILOM | Negodna sa jednim vozilom - sližak uleno sa kolovoza u krivini |
| 1376960 | BEograd | BARAJEVO | 17.04.2022.14:51 | 0035586 | 44.600993 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - SKRETANJE LI PRAZAM | Najmanje dva vozila koja se kreću različitim putevima u skretanje uleno i uz nalik vozila sleva |
| 1367287 | BEograd | BARAJEVO | 22.04.2022.19:50 | 0036706 | 44.626908 | sa pornejanom | SI SA NAMAMJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila - desni sudar |

Слика 35. Изглед табеле преузете са портала отворених података [129]

У преузетом документу о саобраћајним незгодама, може се уочити седам променљивих које су представљене и одвојене по колонама. Називи и објашњења колоне приказани су Табелом 9.

Табела 9. Називи о објашњења променљивих које се користе из Скупа отворених података

| Редни број | Назив променљиве | Назив колоне | Објашњење |
|------------|------------------|---------------|--|
| 1. | NEZG_ID | identifikator | Идентификациони број саобраћајне незгоде |
| 2. | VREME_NEZ | vreme | Датум и време саобраћајне незгоде |
| 3. | VRSTA_NEZ | vrsta | Врста саобраћајне незгоде |
| 4. | NAZIV_TIP | tip | Назив и тип саобраћајне незгоде |
| 5. | NAZIV_DET | opis | Детаљни опис саобраћајне незгоде |
| 6. | WGS_X | duzina | X координата WGS (engl. World Geodetic System) или Географска дужина |
| 7. | WGS_Y | sirina | Y координата WGS (engl. World Geodetic System) или Географска ширина |

На тржишту база података данас су убедљиво најзаступљеније SQL базе података [130]. Под термином SQL базе података у пракси се подразумевају релационе базе података. С обзиром на то да DBMS и модел података представљају два најважнија аспекта базе података, за оне базе података које се заснивају на релационом моделу (самим тим које користе и релациони DBMS)

употребљава се термин релациона база података [131]. У основи релационих база података, које се заснивају на релационом моделу података, подаци се уносе у форми записа (тј. записа базе података, односно слога базе података), али је присутна и подршка за рад са објектним типовима података.

Главне карактеристике *SQL* база података су, поред заснованости на релационом моделу, да захтевају испуњеност *ACID* особина трансакција, да подржавају *SQL* упитни језик и да су погодне за вертикалну скалабилност. *SQL* представља стандардни упитни језик за рад са *SQL* базама података (релационим и уз одређена проширења и за рад са објектно-релационим базама података). Овај стандард је усвојен од стране *American National Standards Institute (ANSI, енгл.)* и постао је стандард који подржавају сви водећи произвођачи *SQL* база података. Свака трансакција над *SQL* базом података мора да испуни *ACID* особине [132]:

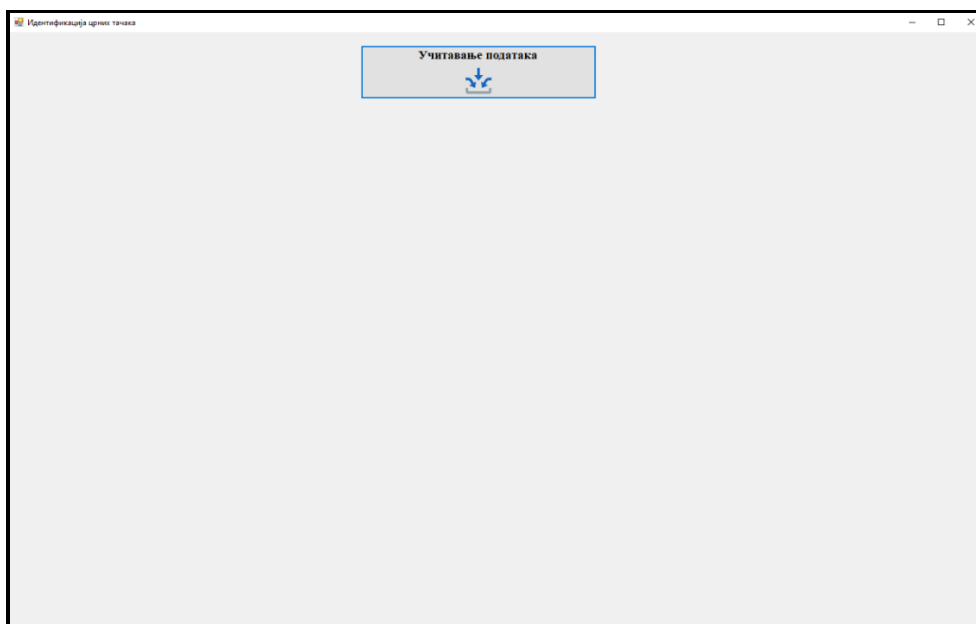
- Атомност (*Atomicity, енгл.*) – све операције трансакције над базом података се морају успешно обавити или ниједна неће бити реализована;
- Конзистентност (*Consistency, енгл.*) – пре почетка и након завршетка трансакције база података мора бити у конзистентном стању;
- Изолација (*Isolation, енгл.*) – ефекти две или више трансакција које се извршавају истовремено морају бити изоловани и
- Трајност (*Durability, енгл.*) – ефекти завршене трансакције не могу бити изгубљени.

Крајем прошлог века, у употреби су осим *SQL* база података, које су биле доминантне по заступљености [133], били присутни и други типови база података, али у знатно мањој мери. Временом, постављани су нови типови задатака пред *SQL* базе података, а очекивања од примене база података су се усложњавала. Неке од новонасталих потреба су решаване проширењем функционалности *SQL* база података, попут додавања кориснички дефинисаних типова, у циљу подршке објектних концепата, или додавања „Звезда“ шеме (*Star Schema, енгл.*), у циљу адекватнијег пројектовања складишта података (*Datawarehouse, енгл.*). Тако на пример *Microsoft SQL Server Management Studio* СУБП подржава рад са тзв. „Звезда“

трансформацијом [134]. Експанзија велике количине података различитог степена структурираности, доступних из различитих извора (не само традиционалних извора, већ и новијих попут друштвених мрежа), додатно је нагласила потребу брже обраде података променљиве структуре доступних у различитим форматима.

За потребе рада креирана је апликација у оквиру система пословне интелигенције, са циљем да корисник може самостално да је користи и да идентификује опасна места на одређеним локацијама, за одређени временски период, са сопственим параметрима дужине пута на којој се идентификују опасна места.

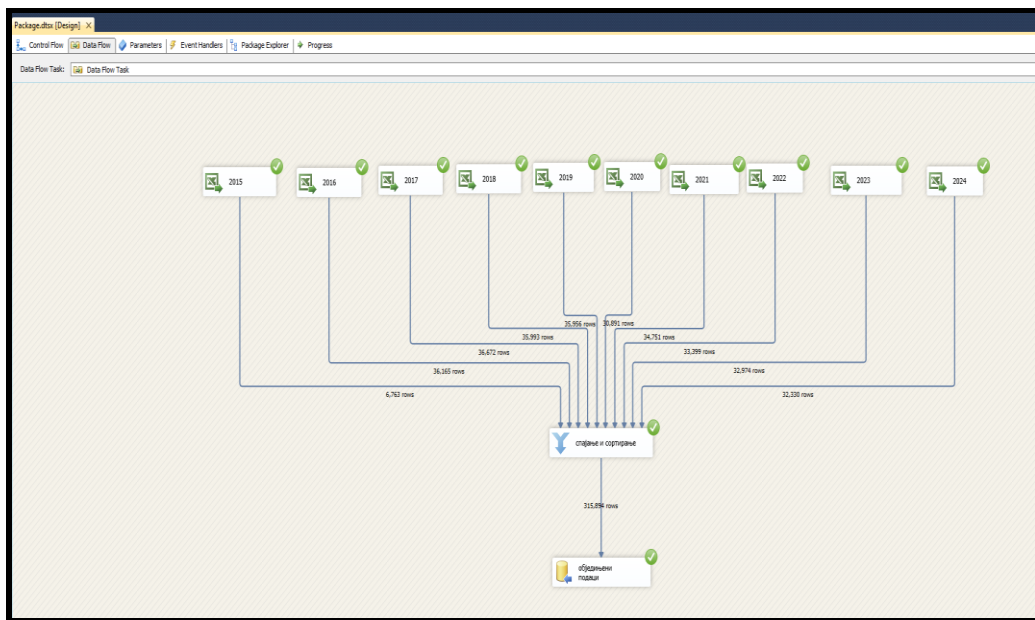
При покретању апликације кориснику се приказује форма као на Слици 36.



Слика 36. Изглед форме при покретању апликације

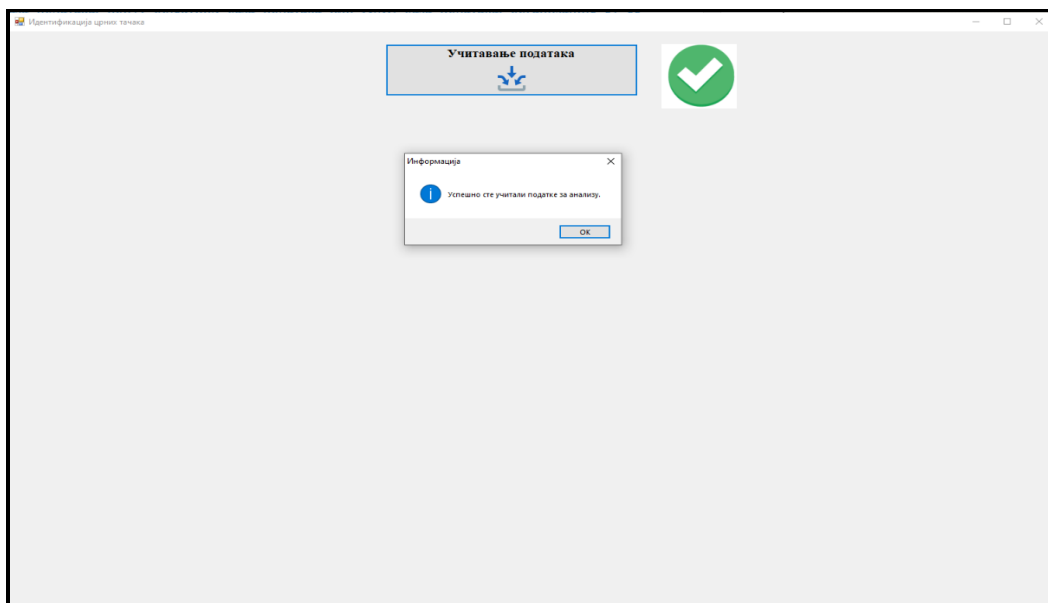
Кликом на дугме “Учитавање података” започиње се први корак у обради наведених података, а то је трансформација података из формата у коме су преузети и њихово уписивање у одговарајућу базу података.

Поступак учитавања података приказан је на Слици 37.



Слика 37. Учитавање података из изворних фајлова у складиште података

Уколико је учитавање података успешно извршено корисник ће добити обавештење као што је приказано на Слици 38.



Слика 38. Изглед форме након успешног учитавања података за анализу

Након учитавања података у базу података, подаци нам се налазе у саставу једне табеле. Изглед извода из табеле са подацима у *SQL Server Management Studio* приказан је на Слици 39. За сваку локацију, додата је колона, као обележје чвора, чија је та локација центар.

Архитектура и конструкција окружења за идентификацију “црних тачака” на путевима у Републици Србије коришћењем модификованог алгорита и система пословне интелигенције

| id | identifikator | grad | opstina | vreme | datum | brisa | vrsta | lo | opis | oznaka |
|-------|---------------|----------|----------|------------------|-----------|-----------|---------------|---|---|-----------|
| 52602 | 1420287 | NOVI SAD | SRBOPRAN | 28.05.2023.17.10 | 19.843887 | 45.538892 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću različitim putevima uz skretanje ulazno - čovni sudar sa vozilom koje nalazi zdesna | ovr_52602 |
| 52603 | 1420323 | NOVI SAD | SRBOPRAN | 06.06.2023.19.32 | 19.799573 | 45.544284 | Sa mat.stetom | SN SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa leve strane kolovoza | ovr_52603 |
| 52604 | 1420723 | NOVI SAD | SRBOPRAN | 18.06.2023.19.10 | 19.802034 | 45.554593 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje ulazno ispred drugog vozila | ovr_52604 |
| 52605 | 1430677 | NOVI SAD | SRBOPRAN | 10.06.2023.09.55 | 19.806993 | 45.556576 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću istim putem u istom smeru uz skretanje, sudar u saustizanju | ovr_52605 |
| 52606 | 1430730 | NOVI SAD | SRBOPRAN | 11.06.2023.02.00 | 19.800804 | 45.549514 | Sa poveljenim | SN SA JEDINIM VOZILOM | Nezgodе sa jednim vozilom u raskrscu ili na putji | ovr_52606 |
| 52607 | 1431588 | NOVI SAD | SRBOPRAN | 18.06.2023.16.30 | 19.788374 | 45.543084 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Ostale nezgodе sa napiranjem diva vozila koja se kreću različitim putevima uz skretanje | ovr_52607 |
| 52608 | 1432228 | NOVI SAD | SRBOPRAN | 21.06.2023.18.45 | 19.822602 | 45.503544 | Sa mat.stetom | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću različitim putevima uz prolazak kroz raskrscu, li od kojih jedno prelaz preko kolovoza, bez skretanja | ovr_52608 |
| 52609 | 1432229 | NOVI SAD | SRBOPRAN | 24.06.2023.13.15 | 19.810767 | 45.562712 | Sa poveljenim | SN SA JEDINIM VOZILOM | Nezgodа sa jednim vozilom - alizak sa kolovoza u krlini | ovr_52609 |
| 52610 | 1435365 | NOVI SAD | SRBOPRAN | 26.06.2023.06.55 | 19.795249 | 45.549506 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću različitim putevima uz skretanje ulazno u ulazizak vozila sleva | ovr_52610 |
| 52611 | 1434039 | NOVI SAD | SRBOPRAN | 07.07.2023.21.30 | 19.785996 | 45.497676 | Sa poveljenim | SN SA PARKIRANIM VOZILIMA | Ostali sudar sa parkiranim vozilom | ovr_52611 |
| 52612 | 1434348 | NOVI SAD | SRBOPRAN | 10.07.2023.18.00 | 19.793863 | 45.548225 | Sa poveljenim | SN SA JEDINIM VOZILOM | Nezgodа sa jednim vozilom na kolovozu | ovr_52612 |
| 52613 | 1435007 | NOVI SAD | SRBOPRAN | 14.07.2023.15.30 | 19.790009 | 45.517184 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću istim putem u istom smeru uz skretanje, sudar u saustizanju | ovr_52613 |
| 52614 | 1447838 | NOVI SAD | SRBOPRAN | 16.07.2023.03.20 | 19.849344 | 45.539435 | Sa poveljenim | SN SA JEDINIM VOZILOM | Nezgodа sa jednim vozilom - alizak sa kolovoza u krlini | ovr_52614 |
| 52615 | 1435978 | NOVI SAD | SRBOPRAN | 21.07.2023.07.45 | 19.841322 | 45.575119 | Sa poveljenim | SN SA JEDINIM VOZILOM | Nezgodа sa jednim vozilom - alizak ulazno na pravcu | ovr_52615 |
| 52616 | 1438042 | NOVI SAD | SRBOPRAN | 21.07.2023.13.50 | 19.917639 | 45.472347 | Sa mat.stetom | SN SA NALJANJE DIVA VOZILA - BEZ SKRETANJA | Ostale nezgodе sa napiranjem diva vozila - suprotni smerovi bez skretanja | ovr_52616 |
| 52617 | 1436283 | NOVI SAD | SRBOPRAN | 22.07.2023.14.40 | 19.843829 | 45.572158 | Sa poveljenim | SN SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa desne strane kolovoza | ovr_52617 |
| 52618 | 1437320 | NOVI SAD | SRBOPRAN | 26.07.2023.16.05 | 19.840584 | 45.575182 | Sa poveljenim | SN SA JEDINIM VOZILOM | Nezgodа sa jednim vozilom - alizak ulazno sa kolovoza na pravcu | ovr_52618 |
| 52619 | 1437802 | NOVI SAD | SRBOPRAN | 01.08.2023.16.50 | 19.802799 | 45.551866 | Sa mat.stetom | SN SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa leve strane kolovoza | ovr_52619 |
| 52620 | 1439653 | NOVI SAD | SRBOPRAN | 14.08.2023.17.50 | 19.785858 | 45.547772 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću različitim putevima uz prolazak kroz raskrscu, li od kojih jedno prelaz preko kolovoza, bez skretanja | ovr_52620 |
| 52621 | 1440324 | NOVI SAD | SRBOPRAN | 18.08.2023.16.30 | 19.749241 | 45.551991 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću istim putem u suprotnim smerovima uz skretanje ulazno ispred drugog vozila | ovr_52621 |
| 52622 | 1441699 | NOVI SAD | SRBOPRAN | 26.08.2023.01.45 | 19.772961 | 45.549809 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodа sa jednim vozilom - alizak ulazno sa kolovoza na pravcu | ovr_52622 |
| 52623 | 1442474 | NOVI SAD | SRBOPRAN | 31.08.2023.20.10 | 19.807861 | 45.549823 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću različitim putevima uz prolazak kroz raskrscu, li od kojih jedno prelaz preko kolovoza, bez skretanja | ovr_52623 |
| 52624 | 1443672 | NOVI SAD | SRBOPRAN | 08.09.2023.11.00 | 19.789896 | 45.532977 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - BEZ SKRETANJA | Napiranje diva vozila koja se kreću u istom smeru - saustizanje | ovr_52624 |
| 52625 | 1444446 | NOVI SAD | SRBOPRAN | 13.09.2023.16.00 | 19.825524 | 45.576087 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - BEZ SKRETANJA | Napiranje diva vozila koja se kreću u istom smeru - pretilanje | ovr_52625 |
| 52626 | 1453002 | NOVI SAD | SRBOPRAN | 19.09.2023.13.40 | 19.80569 | 45.559191 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodа sa jednim vozilom - alizak ulazno sa kolovoza na pravcu | ovr_52626 |
| 52627 | 1449770 | NOVI SAD | SRBOPRAN | 20.09.2023.20.50 | 19.761243 | 45.550842 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - BEZ SKRETANJA | Napiranje diva vozila koja se kreću u istom smeru - saustizanje | ovr_52627 |
| 52628 | 1445426 | NOVI SAD | SRBOPRAN | 22.09.2023.18.00 | 19.805092 | 45.556994 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću istim putem u suprotnim smerovima uz skretanje ulazno ispred drugog vozila | ovr_52628 |
| 52629 | 1446849 | NOVI SAD | SRBOPRAN | 05.10.2023.15.15 | 19.795912 | 45.549918 | Sa poveljenim | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću istim putem u suprotnim smerovima uz skretanje ulazno ispred drugog vozila | ovr_52629 |
| 52630 | 1447219 | NOVI SAD | SRBOPRAN | 07.10.2023.13.25 | 19.789476 | 45.548889 | Sa mat.stetom | SN SA NALJANJE DIVA VOZILA - SKRETANJE ILI PRELAZAK | Napiranje diva vozila koja se kreću istim putem u istom smeru uz skretanje, sudar u saustizanju | ovr_52630 |

Слика 39. Изглед дела табеле са подацима о саобраћајним незгодама на територији Републике Србије, након уписивања преузетих података са Портала отворених података Републике Србије

Сагледавањем наведених података и формата у којима су уписани, могуће је донети закључке који су примењени у остатку рада и били су од користи при креирању складишта података. Обзиром да се за идентификацију и одређивање опасних места на путевима користе подаци који представљају трогодишњи период, коришћен је скуп података о саобраћајним незгодама на територији Републике Србије за период од 01.01.2015. до 31.12.2024. године, а исти принцип може бити примењен на све остале скупове података по годинама.

- Поље у коме се воде подаци о датуму и времену саобраћајне незгоде, дато је у виду низа карактера. Обзиром да Систем за управљање базом података *SQL Server Management Studio* подржава широк спектар типова података који служе за анализу података по димензији време, над наведеном колоном биће извршена конверзија и трансформација тренутно коришћеног типа податка у нове типове података који ће бити сврсисходнији. Једно од решења је да се, за почетак, наведено поље модификује, тако да се посебно воде подаци о датуму, а посебно подаци о времену. При креирању складишта података, и димензије време, која је обавезна за складиште података, могу се на основу наведених вредности добити и вредности нових колона, нпр. дан у недељи, дан у месецу, дан у години и сл., што је детаљније

описано у делу при трансформацији података. Након реализације предходно описаног, добијамо нове две вредности за детаљнији опис временских карактеристика, када се догодила саобраћајна незгода, што је приказано на Слици 40.

| id | identifikator | grad | opstina | datum | vreme | duzina | asina | vrsta | tip | opis | oznaka |
|-----|---------------|---------|---------|-------------|-------|-----------|------------|----------------|--|---|----------|
| 285 | 1475633 | BEOGRAD | GROCKA | 02.05.2024. | 14.30 | 20.627888 | 44.6514846 | Sa mat.stetom | SN SA NAJMANJE DVA VOZILA - SKRETANJE ILI PRELAZAK | Najmanje dva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje udesno ispred drugog vozila | cvor_285 |
| 286 | 1474243 | BEOGRAD | GROCKA | 02.05.2024. | 21.00 | 20.58122 | 44.72864 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa jednim vozilom na kolovozu | cvor_286 |
| 287 | 1474961 | BEOGRAD | GROCKA | 05.05.2024. | 15.50 | 20.586319 | 44.758465 | Sa mat.stetom | SN SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom sa desne strane kolovoza | cvor_287 |
| 288 | 1475627 | BEOGRAD | GROCKA | 06.05.2024. | 15.50 | 20.6405 | 44.6180309 | Sa mat.stetom | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - susretanje | cvor_288 |
| 289 | 1476030 | BEOGRAD | GROCKA | 06.05.2024. | 20.00 | 20.597008 | 44.674072 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - susretanje | cvor_289 |
| 290 | 1474746 | BEOGRAD | GROCKA | 06.05.2024. | 18.45 | 20.613274 | 44.663376 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa učešćem jednog vozila i preprekama na ili iznad kolovoza | cvor_290 |
| 291 | 1476120 | BEOGRAD | GROCKA | 07.05.2024. | 05.30 | 20.640493 | 44.606377 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa jednim vozilom - slazak udesno sa kolovoza na pravcu | cvor_291 |
| 292 | 1478254 | BEOGRAD | GROCKA | 09.05.2024. | 10.50 | 20.618121 | 44.731169 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Ostale nezgode sa najmanje dva vozila - suprotni smerovi bez skretanja | cvor_292 |
| 293 | 1476540 | BEOGRAD | GROCKA | 10.05.2024. | 15.00 | 20.648165 | 44.737742 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - uključivanje u saobraćaj | cvor_293 |
| 294 | 1476548 | BEOGRAD | GROCKA | 14.05.2024. | 13.30 | 20.583231 | 44.693593 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa jednim vozilom na kolovozu | cvor_294 |
| 295 | 1475461 | BEOGRAD | GROCKA | 14.05.2024. | 18.30 | 20.586014 | 44.762572 | Sa mat.stetom | SN SA NAJMANJE DVA VOZILA - SKRETANJE ILI PRELAZAK | Najmanje dva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje ulevo ispred drugog vozila | cvor_295 |
| 296 | 1476086 | BEOGRAD | GROCKA | 15.05.2024. | 16.20 | 20.598488 | 44.743831 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - susretanje | cvor_296 |
| 297 | 1476549 | BEOGRAD | GROCKA | 15.05.2024. | 05.15 | 20.637673 | 44.628393 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa jednim vozilom - slazak ulevo sa kolovoza na pravcu | cvor_297 |
| 298 | 1476701 | BEOGRAD | GROCKA | 16.05.2024. | 20.00 | 20.637847 | 44.630241 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa jednim vozilom - slazak udesno sa kolovoza na pravcu | cvor_298 |
| 299 | 1476692 | BEOGRAD | GROCKA | 16.05.2024. | 13.30 | 20.608135 | 44.739686 | Sa mat.stetom | SN SA NAJMANJE DVA VOZILA - SKRETANJE ILI PRELAZAK | Najmanje dva vozila koja se kreću različitim putevima uz skretanje ulevo i uz nalazak vozila sleva | cvor_299 |
| 300 | 1478425 | BEOGRAD | GROCKA | 16.05.2024. | 19.30 | 20.642403 | 44.611379 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Najmanje dva vozila koja se kreću u istom smeru - susretanje | cvor_300 |
| 301 | 1475721 | BEOGRAD | GROCKA | 16.05.2024. | 19.00 | 20.642021 | 44.610018 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa učešćem jednog vozila i preprekama na ili iznad kolovoza | cvor_301 |
| 302 | 1475882 | BEOGRAD | GROCKA | 17.05.2024. | 16.30 | 20.668067 | 44.721418 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Ostale nezgode sa najmanje dva vozila - suprotni smerovi bez skretanja | cvor_302 |
| 303 | 1478795 | BEOGRAD | GROCKA | 17.05.2024. | 09.45 | 20.551987 | 44.763627 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Ostali sudar sa najmanje dva vozila koja se kreću u istom smeru - ostalo | cvor_303 |
| 304 | 1476386 | BEOGRAD | GROCKA | 18.05.2024. | 20.20 | 20.58169 | 44.66088 | Sa mat.stetom | SN SA PARKIRANIM VOZILIMA | Ostali sudar sa parkiranim vozilom | cvor_304 |
| 305 | 1476092 | BEOGRAD | GROCKA | 19.05.2024. | 15.50 | 20.642465 | 44.613319 | Sa mat.stetom | SN SA NAJMANJE DVA VOZILA - SKRETANJE ILI PRELAZAK | Najmanje dva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje udesno ispred drugog vozila | cvor_305 |
| 306 | 1477150 | BEOGRAD | GROCKA | 19.05.2024. | 15.00 | 20.61529 | 44.66468 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa učešćem jednog vozila i preprekama na ili iznad kolovoza | cvor_306 |
| 307 | 1477242 | BEOGRAD | GROCKA | 23.05.2024. | 09.10 | 20.61311 | 44.664462 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa učešćem jednog vozila i preprekama na ili iznad kolovoza | cvor_307 |
| 308 | 1476616 | BEOGRAD | GROCKA | 24.05.2024. | 11.30 | 20.714937 | 44.670193 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - SKRETANJE ILI PRELAZAK | Najmanje dva vozila koja se kreću različitim putevima uz skretanje ulevo i uz nalazak vozila sleva | cvor_308 |
| 309 | 1477362 | BEOGRAD | GROCKA | 26.05.2024. | 18.30 | 20.5844 | 44.6903993 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - SKRETANJE ILI PRELAZAK | Najmanje dva vozila koja se kreću istim putem u istom smeru uz skretanje, sudar u susretanju | cvor_309 |
| 310 | 1476087 | BEOGRAD | GROCKA | 27.05.2024. | 14.30 | 20.589208 | 44.73954 | Sa povredjenim | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Ostali sudar sa najmanje dva vozila koja se kreću u istom smeru - ostalo | cvor_310 |
| 311 | 1478575 | BEOGRAD | GROCKA | 28.05.2024. | 14.20 | 20.6005 | 44.745012 | Sa mat.stetom | SN SA NAJMANJE DVA VOZILA - BEZ SKRETANJA | Ostali sudar sa najmanje dva vozila koja se kreću u istom smeru - ostalo | cvor_311 |
| 312 | 1477704 | BEOGRAD | GROCKA | 28.05.2024. | 17.00 | 20.580475 | 44.754261 | Sa povredjenim | SN SA JEDINIM VOZILOM | Nezgodu sa jednim vozilom - slazak udesno sa kolovoza na pravcu | cvor_312 |
| 313 | 1478923 | BEOGRAD | GROCKA | 01.06.2024. | 18.30 | 20.586961 | 44.659688 | Sa mat.stetom | SN SA JEDINIM VOZILOM | Nezgodu sa jednim vozilom - slazak ulevo sa kolovoza na pravcu | cvor_313 |

Слика 40. Изглед дела табеле са подацима о саобраћајним незгодима на територији Републике Србије, након трансформације колоне са временским описом

2. Поље у коме се воде подаци о врсти саобраћајне незгоде садржи три јединствене вредности према којима се класификују саобраћајне незгоде. То нам је изузетно битно и неопходна је провера, обзиром да ће се на основу наведених вредности вршити

| | vrsta |
|---|----------------|
| 1 | Sa povredjenim |
| 2 | Sa poginulim |
| 3 | Sa mat.stetom |

прорачун пондерисаних вредности у чворовима и сидрима.

Слика 41. Приказ врста саобраћајних незгода

Из претходно наведеног, можемо закључити да имамо три врсте саобраћајних незгода, и то:

1. Саобраћајна незгода са материјалном штетом;
2. Саобраћајна незгода са повређеним лицима;
3. Саобраћајна незгода са погинулим лицима.

Обзиром да постоји велики број записа којима се евидентирају саобраћајне незгоде, а релативно мали број вредности за опис врсте саобраћајне незгоде, поштујући правила нормализације у складишту података,

то је била једна од димензија на основу које су анализирани подаци и вршен прорачуни.

3. У колони *duzina* воде се подаци о *X* координати *WGS (World Geodetic System, енгл.)* географској дужини, а у колони *sirina* подаци о *Y* координати или географској ширини. Подаци у овим колонама одређени су местом на коме се догодила саобраћајна незгода, и представљаће центре чворова. Приказ података у овим колонама дат је на следећој слици.

| id | identifikator | grad | opstina | datum | vreme | duzina | sirina | vrsta | tip | opis | oznaka | |
|-----|---------------|---------|---------|--------|-------------|--------|-----------|------------|----------------|---|--|----------|
| 285 | 285 | 1475633 | BEOGRAD | GROCKA | 02.05.2024. | 14.30 | 20.627688 | 44.6514846 | Sa mat. stonom | SN SA NAJMANJE DIVA VOZILA – SKRETANJE ILI PRELAZAK | Napmanje diva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje udesno ispred drugog vozila | cvor_285 |
| 286 | 286 | 1474243 | BEOGRAD | GROCKA | 02.05.2024. | 21.00 | 20.581122 | 44.728864 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa jednim vozilom na kolovozu | cvor_286 |
| 287 | 287 | 1474951 | BEOGRAD | GROCKA | 05.05.2024. | 15.50 | 20.583319 | 44.758465 | Sa mat. stonom | SN SA PARKIRANIM VOZILIMA | Sudar sa parkiranim vozilom na deone stare kolovoza | cvor_287 |
| 288 | 288 | 1475627 | BEOGRAD | GROCKA | 06.05.2024. | 15.50 | 20.6405. | 44.6168098 | Sa mat. stonom | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Napmanje diva vozila koja se kreću u istom smeru – sustizanje | cvor_288 |
| 289 | 289 | 1476100 | BEOGRAD | GROCKA | 06.05.2024. | 20.00 | 20.587008 | 44.6740172 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Napmanje diva vozila koja se kreću u istom smeru – sustizanje | cvor_289 |
| 290 | 290 | 1474746 | BEOGRAD | GROCKA | 06.05.2024. | 18.45 | 20.613274 | 44.663376 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa učešćem jednog vozila i preprekama na ili iznad kolovoza | cvor_290 |
| 291 | 291 | 1476120 | BEOGRAD | GROCKA | 07.05.2024. | 05.30 | 20.640493 | 44.606377 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa jednim vozilom – slazak udesno sa kolovoza na pravcu | cvor_291 |
| 292 | 292 | 1478254 | BEOGRAD | GROCKA | 09.05.2024. | 10.50 | 20.618121 | 44.731169 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Ostale nezgode sa najmanje diva vozila – suprotni smerovi bez skretanja | cvor_292 |
| 293 | 293 | 1478540 | BEOGRAD | GROCKA | 10.05.2024. | 15.00 | 20.640165 | 44.737742 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Napmanje diva vozila koja se kreću u istom smeru – ukljuivanje u saobraćaj | cvor_293 |
| 294 | 294 | 1478548 | BEOGRAD | GROCKA | 14.05.2024. | 13.30 | 20.583231 | 44.693993 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa jednim vozilom na kolovozu | cvor_294 |
| 295 | 295 | 1475461 | BEOGRAD | GROCKA | 14.05.2024. | 18.30 | 20.580914 | 44.762572 | Sa mat. stonom | SN SA NAJMANJE DIVA VOZILA – SKRETANJE ILI PRELAZAK | Napmanje diva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje udesno ispred drugog vozila | cvor_295 |
| 296 | 296 | 1476086 | BEOGRAD | GROCKA | 15.05.2024. | 16.20 | 20.589488 | 44.743831 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Napmanje diva vozila koja se kreću u istom smeru – sustizanje | cvor_296 |
| 297 | 297 | 1478549 | BEOGRAD | GROCKA | 15.05.2024. | 05.15 | 20.637673 | 44.628393 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa učešćem jednog vozila – slazak udesno sa kolovoza na pravcu | cvor_297 |
| 298 | 298 | 1478701 | BEOGRAD | GROCKA | 16.05.2024. | 20.00 | 20.637847 | 44.630241 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa jednim vozilom – slazak udesno sa kolovoza na pravcu | cvor_298 |
| 299 | 299 | 1475692 | BEOGRAD | GROCKA | 16.05.2024. | 13.30 | 20.608135 | 44.738986 | Sa mat. stonom | SN SA NAJMANJE DIVA VOZILA – SKRETANJE ILI PRELAZAK | Napmanje diva vozila koja se kreću različitim putevima uz skretanje udesno i uz nalazak vozila sleva | cvor_299 |
| 300 | 300 | 1478425 | BEOGRAD | GROCKA | 16.05.2024. | 19.30 | 20.642403 | 44.611379 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Napmanje diva vozila koja se kreću u istom smeru – sustizanje | cvor_300 |
| 301 | 301 | 1475721 | BEOGRAD | GROCKA | 16.05.2024. | 19.00 | 20.642021 | 44.610018 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa učešćem jednog vozila i preprekama na ili iznad kolovoza | cvor_301 |
| 302 | 302 | 1475882 | BEOGRAD | GROCKA | 17.05.2024. | 16.30 | 20.689967 | 44.721418 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Ostale nezgode sa najmanje diva vozila – suprotni smerovi bez skretanja | cvor_302 |
| 303 | 303 | 1478795 | BEOGRAD | GROCKA | 17.05.2024. | 09.45 | 20.581987 | 44.763627 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Ostali sudar sa najmanje diva vozila koja se kreću u istom smeru – ostalo | cvor_303 |
| 304 | 304 | 1476386 | BEOGRAD | GROCKA | 18.05.2024. | 20.20 | 20.581169 | 44.68808 | Sa mat. stonom | SN SA PARKIRANIM VOZILIMA | Ostali sudar sa parkiranim vozilom | cvor_304 |
| 305 | 305 | 1476992 | BEOGRAD | GROCKA | 19.05.2024. | 15.50 | 20.642465 | 44.613319 | Sa mat. stonom | SN SA NAJMANJE DIVA VOZILA – SKRETANJE ILI PRELAZAK | Napmanje diva vozila koja se kreću istim putem u istom smeru uz skretanje, skretanje udesno ispred drugog vozila | cvor_305 |
| 306 | 306 | 1477150 | BEOGRAD | GROCKA | 19.05.2024. | 15.00 | 20.61129 | 44.66468 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa učešćem jednog vozila i preprekama na ili iznad kolovoza | cvor_306 |
| 307 | 307 | 1477242 | BEOGRAD | GROCKA | 23.05.2024. | 09.10 | 20.6131 | 44.664462 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa učešćem jednog vozila i preprekama na ili iznad kolovoza | cvor_307 |
| 308 | 308 | 1478616 | BEOGRAD | GROCKA | 24.05.2024. | 11.30 | 20.714937 | 44.670193 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – SKRETANJE ILI PRELAZAK | Napmanje diva vozila koja se kreću različitim putevima uz skretanje udesno i uz nalazak vozila sleva | cvor_308 |
| 309 | 309 | 1477362 | BEOGRAD | GROCKA | 26.05.2024. | 18.30 | 20.5844. | 44.6903993 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – SKRETANJE ILI PRELAZAK | Napmanje diva vozila koja se kreću istim putem u istom smeru uz skretanje, sudar u sustizanju | cvor_309 |
| 310 | 310 | 1476087 | BEOGRAD | GROCKA | 27.05.2024. | 14.30 | 20.583208 | 44.76594 | Sa povredjenim | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Ostali sudar sa najmanje diva vozila koja se kreću u istom smeru – ostalo | cvor_310 |
| 311 | 311 | 1478575 | BEOGRAD | GROCKA | 28.05.2024. | 14.20 | 20.6005 | 44.745012 | Sa mat. stonom | SN SA NAJMANJE DIVA VOZILA – BEZ SKRETANJA | Ostali sudar sa najmanje diva vozila koja se kreću u istom smeru – ostalo | cvor_311 |
| 312 | 312 | 1477704 | BEOGRAD | GROCKA | 28.05.2024. | 17.00 | 20.580475 | 44.754261 | Sa povredjenim | SN SA JEDNIM VOZILOM | Nezгода sa jednim vozilom – slazak udesno sa kolovoza na pravcu | cvor_312 |
| 313 | 313 | 1478933 | BEOGRAD | GROCKA | 01.06.2024. | 18.30 | 20.583631 | 44.656388 | Sa mat. stonom | SN SA JEDNIM VOZILOM | Nezгода sa jednim vozilom – slazak udesno sa kolovoza na pravcu | cvor_313 |

Слика 42. Изглед дела табеле са подацима о локацији саобраћаје незгоде на територији Републике Србије

7.3. Екстракција, трансформација и уписивање података у складиште података

7.3.1. Имплементација димензије времена

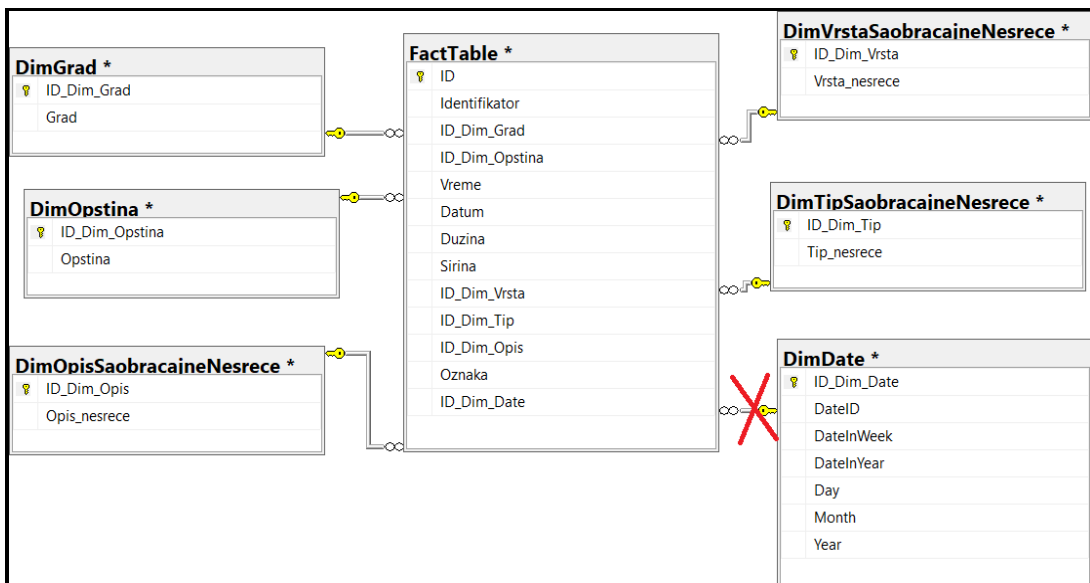
У тексту који следи приказан је начин имплементације димензија у складишту података, на основу доступних и учитаних података. Једна од најважнијих димензија складишта података је димензија времена.

Први корак у имплементацији димензије времена је раскидање везе између *fact* табеле *CentralnaTabela* и табеле *DimDate* која представља димензију времена. Ово се изводи због правила референцијалног интегритета, обзиром да у супротном систем за управљање базом података не би дозволио брисање података из табеле *DimDate*. Након раскидања везе између табела, табела *DimDate* се брише и поново креира, са истим колонама, атрибутима,

типовима података и ограничењима. На тај начин се бришу сви предходни подаци и табела је спремна за уписивање најажурнијег стања из изворишне базе података.

У даљем тексту приказан је програмски код, који прати предходно описани поступак.

```
IF EXISTS (  
SELECT * FROM INFORMATION_SCHEMA.REFERENTIAL_CONSTRAINTS  
WHERE CONSTRAINT_NAME = 'FK_FactTable_DimDate')  
BEGIN  
ALTER TABLE [dbo].[FactTable] DROP CONSTRAINT [FK_FactTable_DimDate] END  
IF EXISTS  
(SELECT * FROM sysobjects  
WHERE id = OBJECT_ID(N'DimDate'))  
BEGIN  
DROP TABLE [dbo].[DimDate] CREATE TABLE [dbo].[DimDate](  
[ID_Dim_Date] [int] IDENTITY(1,1) NOT NULL, [DateID] [int] NOT NULL,  
[DateInWeek] [int] NOT NULL, [DateInYear] [int] NOT NULL, [Day] [int] NOT NULL,  
[Month] [int] NOT NULL, [Year] [int] NOT NULL,  
CONSTRAINT [PK_DimDate] PRIMARY KEY CLUSTERED (  
[ID_Dim_Date] ASC  
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY  
= OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON) ON [PRIMARY]  
) ON [PRIMARY] END
```



Слика 43. Брисање везе између fact табеле и табеле у којој се налазе подаци за димензију времена

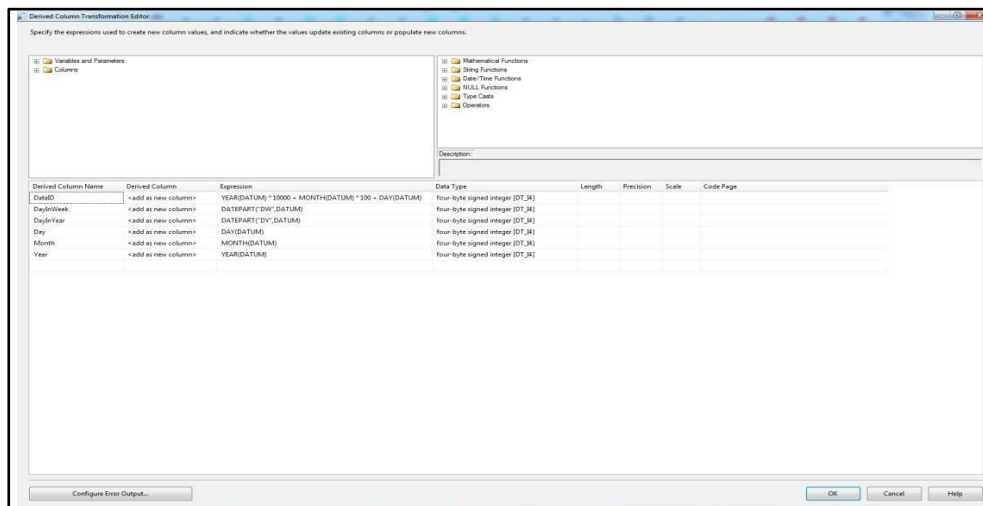
Подаци су конвертовани из оригиналног типа податка у коме су преузети са платформе *nvarchar(255)* у тип податка *date*.

Након тога, на основу добијене вредности, подаци су конвертовани по различитим критеријумима у циљу добијања нових вредности за дефинисане колоне у складишту података, у табели *DimDate*.

Колоне од интереса су *DayID*, *DayInWeek*, *DayInYear*, *Day*, *Month*, *Year*. Колона *DayID* користи се за спајање *fact* табеле и димензије *DimDate*, а креирана је на следећи начин:

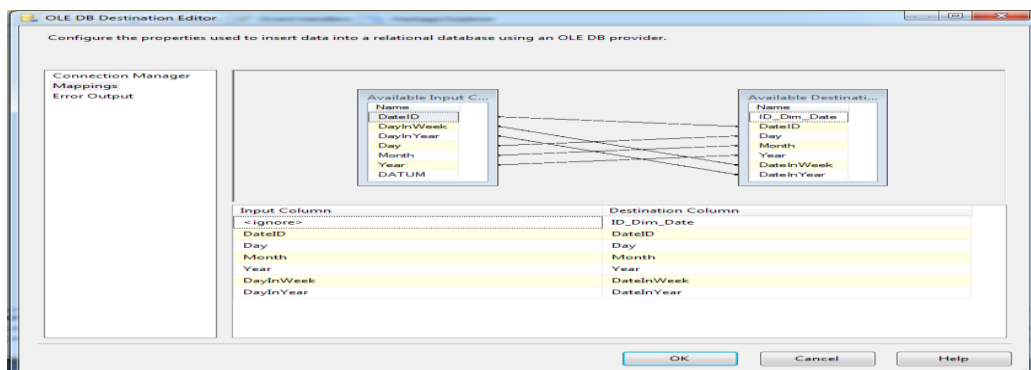
$$YEAR([Datum]) * 10000 + MONTH([Datum]) * 100 + DAY([Datum])$$

DayID од 25.11.2023. је 20231125. На овај начин омогућене су нове карактеристике за претрагу по критеријуму времена, а додатно су побољшане перформансе при сортирању.



Слика 44. Конверовање података и генерисање нових вредности у табели *DimDate*

Након претходне активности, следи сортирање по колони *DayID* и уписивање података у жељено одредиште, а то је табела *DimDay* у складишту података.

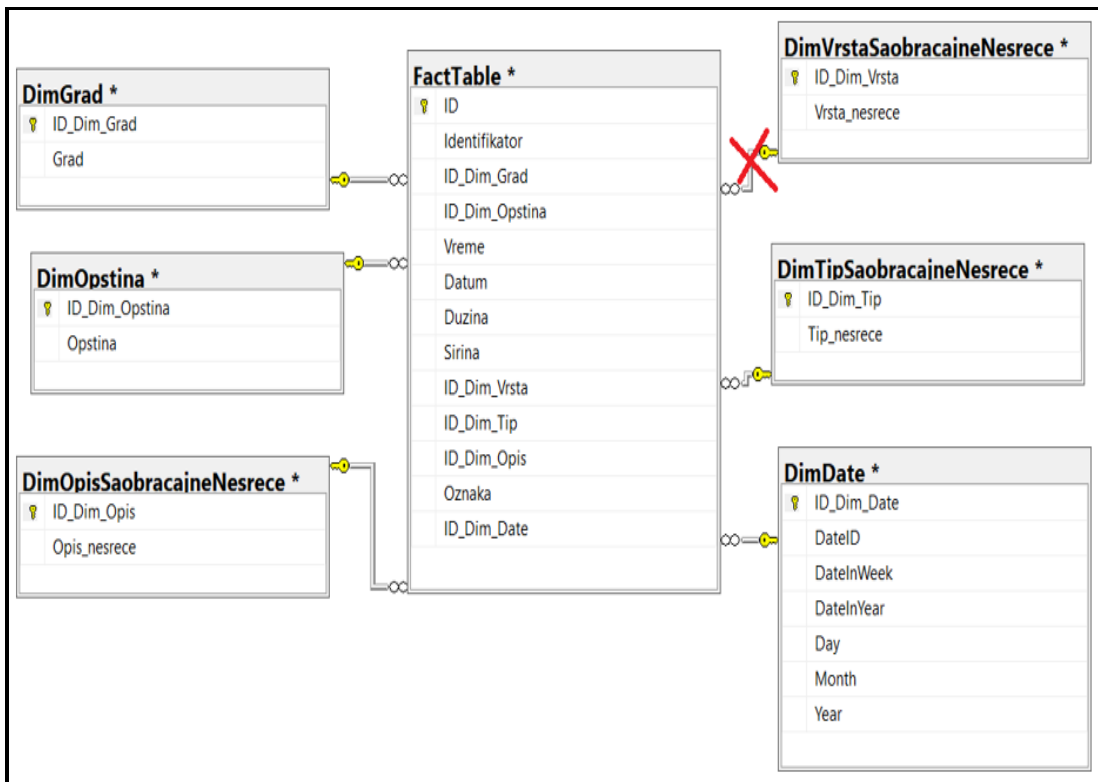


Слика 45. Уписивање података у жељено одредиште

Наведеним поступком имплементирана је димензија којом су описани подаци о времену. Одређено је извориште, развијена је процедура за трансформацију података и одређено је одредиште за уписивање података, чиме су подаци припремљени за даљу употребу.

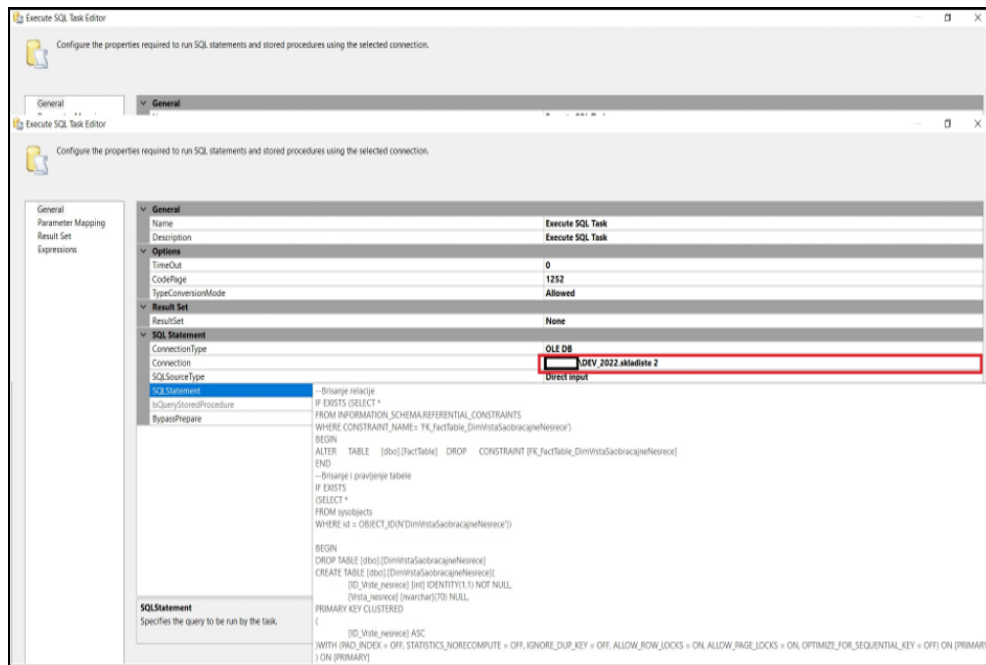
7.3.2. Имплементација димензије Врста саобраћајне незгоде

Једна од димензија која се користи у складишту података је димензија којом су детаљно описане врсте саобраћајних незгода. Као што је претходно наведено, у скупу преузетих података постоје укупно три врсте саобраћајних незгода. У тексту који следи приказан је поступак имплементације наведене димензије.



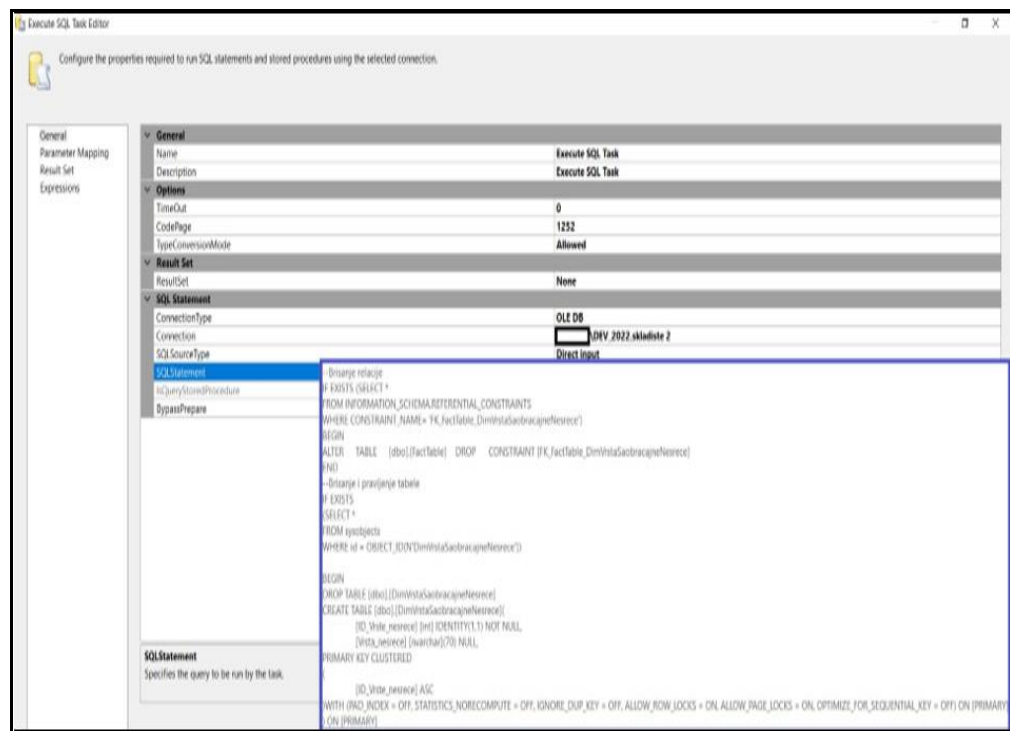
Слика 46. Брисање релације између fact табеле и димензије при уписивању података у димензију

Да би релација била обрисана, неопходно је прво дефинисати назив сервера над којим се ради и одредити дестинацију на којој се налази складиште података.



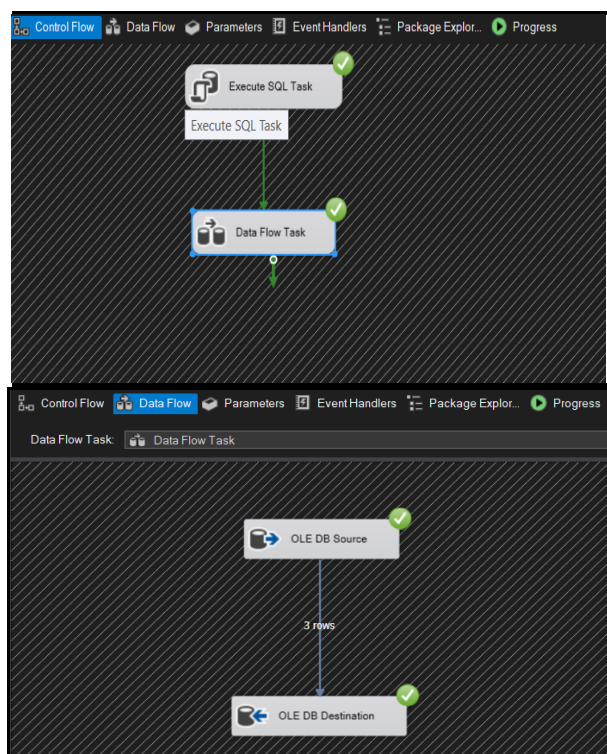
Слика 47. Дефинисање одредишта и успостављање конекције са складиштем података

Након дефинисања параметара везаних за конекцију са складиштем података, дефинише се брисање релације између димензије и *fact* табеле, брисање комплетне табеле и њено поновно креирање, као што је приказано на Слици 48.



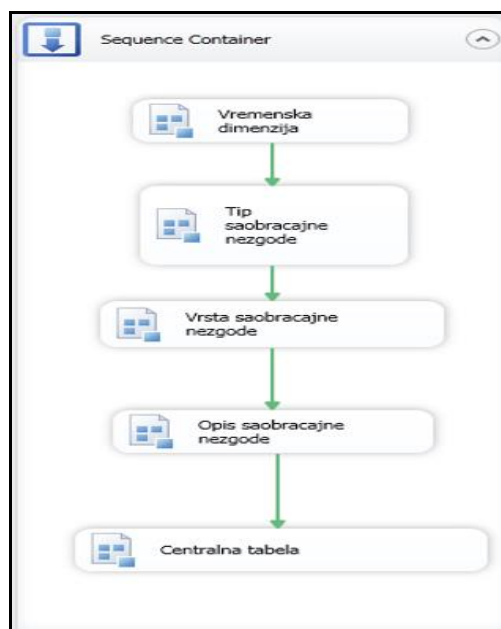
Слика 48. Приказ упита за брисање релације, брисање и креирање димензије

Следећи корак је дефинисање извора података и одредишта на које ће подаци бити уписани. Поступак успешног учитавања и уписивања података приказан је на Слици 49. Тиме је процес креирања наведене димензије завршен.



Слика 49. Приказ процеса учитавања, трансформације и уписивања података у димензију Врста саобраћајне незгоде у складишту податаку

Након што су дефинисане и имплементиране димензије, приступа се имплементацији централне *fact* табеле. Као и у случајевима имплементације димензија, врши се брисање података из централне табеле складишта података. Нема брисања релација ка димензијама, обзиром да су већ обрисане при поновном уносу података у димензије, и при креирању централне табеле, оне се заједно са њом поново креирају. Врши се повезивање кључева помоћу релација, учитавање, трансформација и уписивање података у централну табелу. Изглед корака који се извршавају при наведеном процесу приказан је на Слици 50.



Слика 50. Приказ процеса учитавања, трансформације и уписивања података у димензије и централну табелу при генерисању складишта података

Након наведених поступака, завршен је поступак учитавања, трансформације и уписа података у складиште података. Подаци су обрађени и могу се користити за анализе по различитим критеријумима.

Када су подаци припремљени, корисник може да добије извештај о “црним тачкама”, за локацију и временски интервал који му је од интереса. Изглед форме за унос параметара приказан је на Слици 51.

Идентификација црних тачака

Учитавање података

Изаберите град

Изаберите општину

Унесите почетну вредност полупречника кластера метара

Унесите доњу границу вредности кластера

Изаберите почетни датум

Изаберите крајњи датум

Идентификуј "црне тачке"

Слика 51. Изглед форме на којој корисник уноси податке за идентификацију “црних тачака”

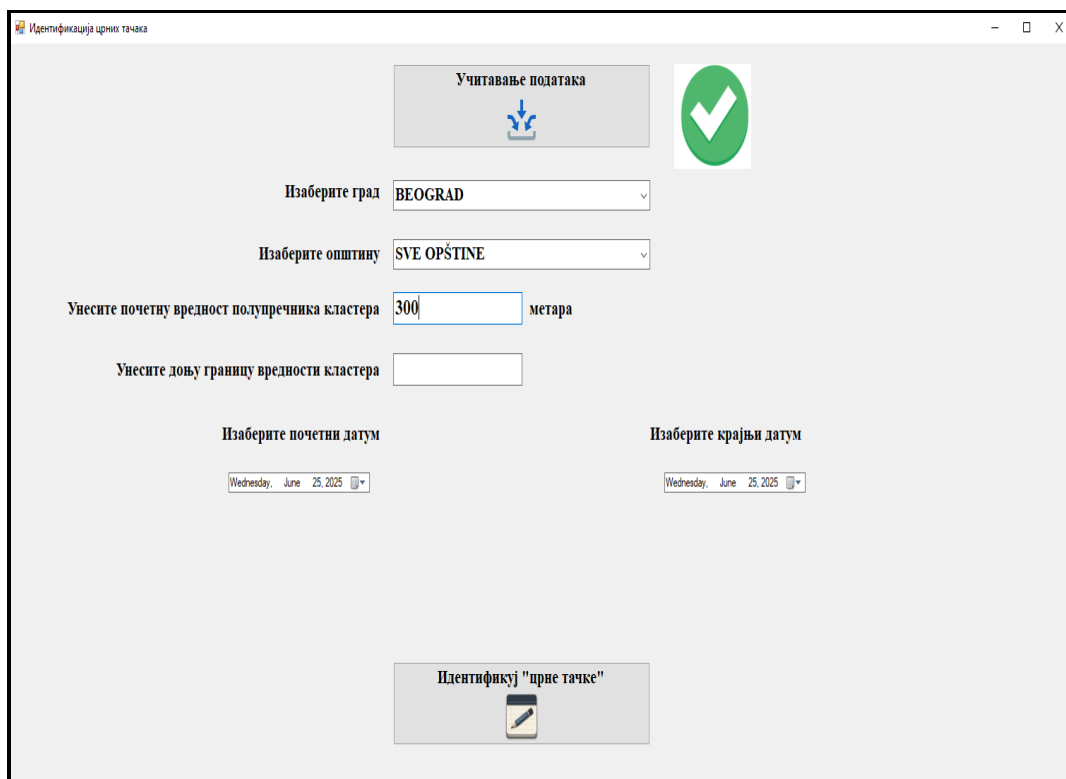
Корисник бира град за који жели да идентификује “црне тачке” или опцију “SVI GRADOVI” уколико жели да идентификује црне тачке на територији целе државе. Изглед форме када корисник бира град приказан је на слици 52.

Слика 52. Изглед форме на којој корисник бира град за који жели идентификацију “црних тачака”

За изабрани град корисник бира општину или све општине у том граду за које жели да идентификује “црне тачке”, што је приказано на Слици 53.

Слика 53. Изглед форме на којој корисник бира општину за коју жели идентификацију “црних тачака”

У следећем кораку корисник уноси почетну вредност величине кластера, као што је описано у делу начина функционисања предложеног модификованог алгоритма. Предлог је да почетна вредност буде 300 метара, а та вредност ће се постепено повећавати у сваком од корака рада алгоритма, док се не дође до најоптималнијих кластера.



Слика 54. Изглед форме на којој корисник уноси почетну вредност полупречника кластера

Након тога, корисник уноси минималну вредност величине суме пондерисаних вредности, као што је описано у делу о начину функционисања предложеног модификованог алгоритма. Предлог је да то буде 85, обзиром да је то гаранција да се на том месту догодила најмање једна саобраћајна незгода са страдалим лицима или девет саобраћајних незгода са повређеним лицима.

Идентификација црних тачака

Учитавање података

Изаберите град BEOGRAD

Изаберите општину SVE OPŠTINE

Унесите почетну вредност полупречника кластера 300 метара

Унесите доњу границу вредности кластера 85

Изаберите почетни датум Wednesday, June 25, 2025

Изаберите крајњи датум Wednesday, June 25, 2025

Идентификуј "црне тачке"

Слика 55. Изглед форме на којој корисник уноси доњу границу вредности кластера

У следећем кораку корисник бира трогодишњи период за који жели идентификацију “црних тачака” на изабраном подручју.

Идентификација црних тачака

Учитавање података

Изаберите град BEOGRAD

Изаберите општину SVE OPŠTINE

Унесите почетну вредност полупречника кластера 300 метара

Унесите доњу границу вредности кластера 85

Изаберите почетни датум

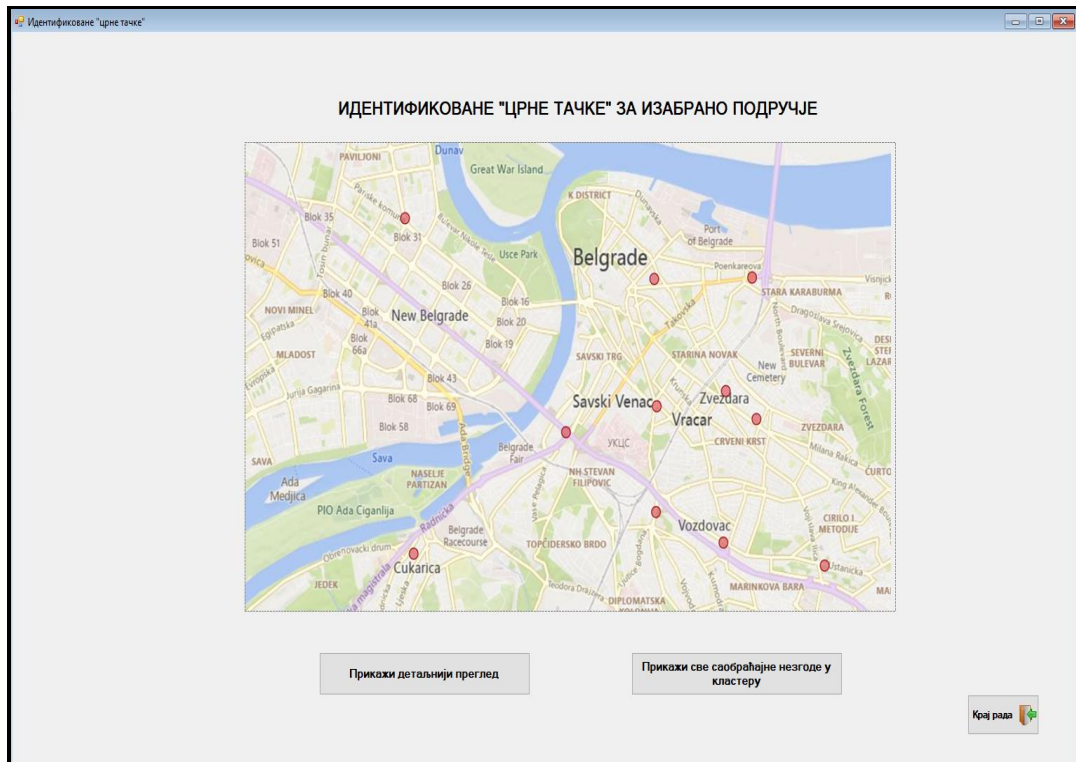
Изаберите крајњи датум

Идентификуј "црне тачке"

Слика 56. Изглед форме на којој корисник бира временски оквир за идентификацију “црних тачака” на изабраном подручју

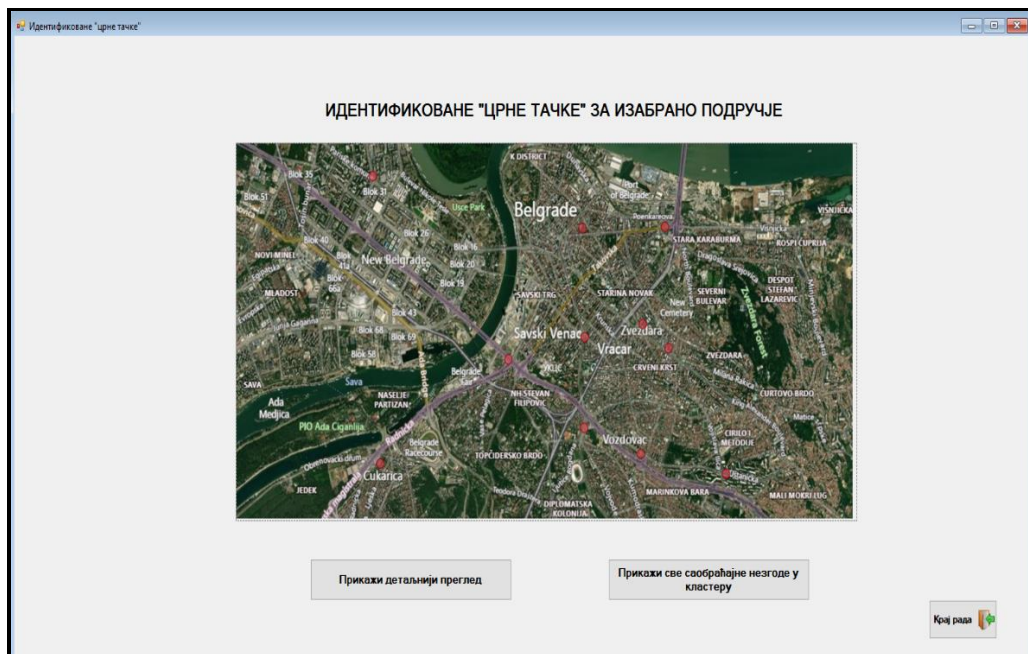
Кликом на дугме “Идентификуј “црне тачке”” корисник добија форму

са прегледом предложених и извојених опасних места, што је приказано на Слици 57.



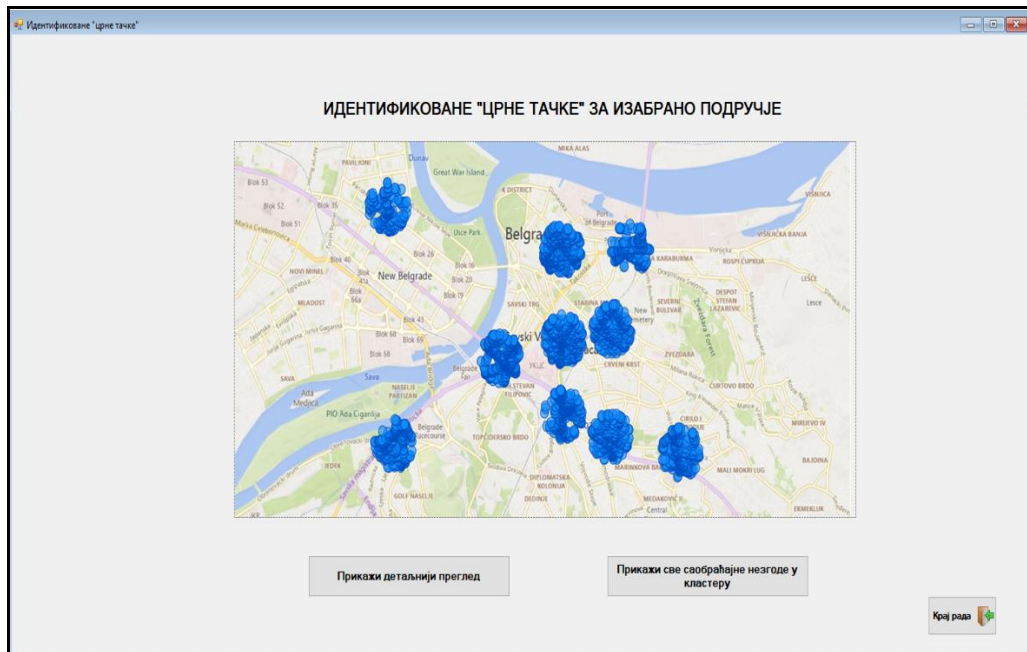
Слика 57. Изглед форме са идентификованим “црним тачкама” на изабраном подручју

Кликом на дугме “Прикажи детаљнији преглед” добија се мапа као на слици број 58.



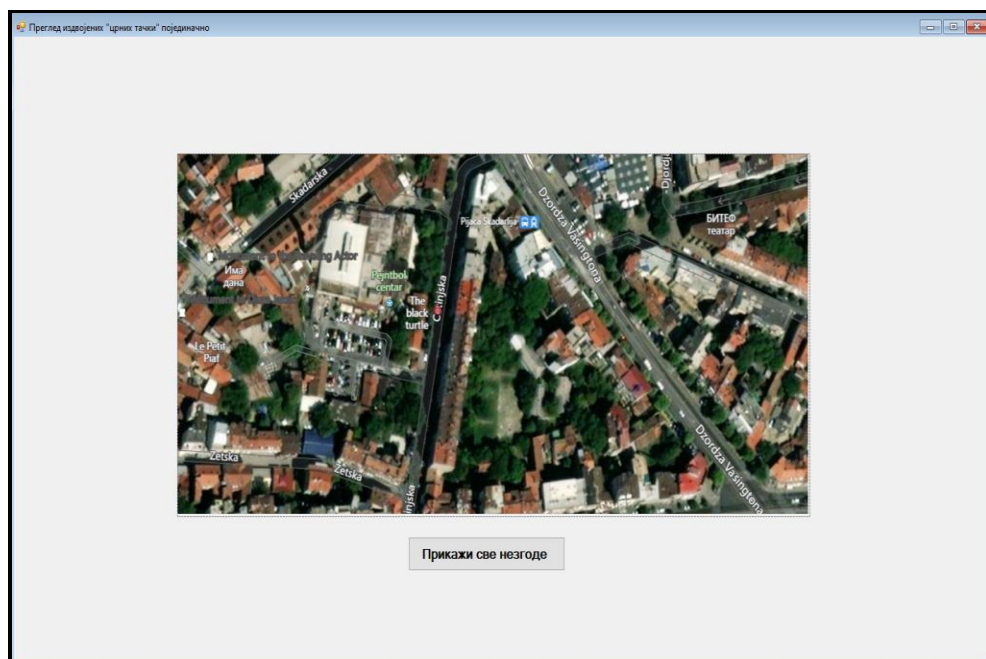
Слика 58. Изглед форме са идентификованим “црним тачкама” на изабраном подручју са опцијом детаљнијег прегледа

Кликом на дугме “Прикажи све саобраћајне незгоде у кластери” добија се мапа као на слици број 59.

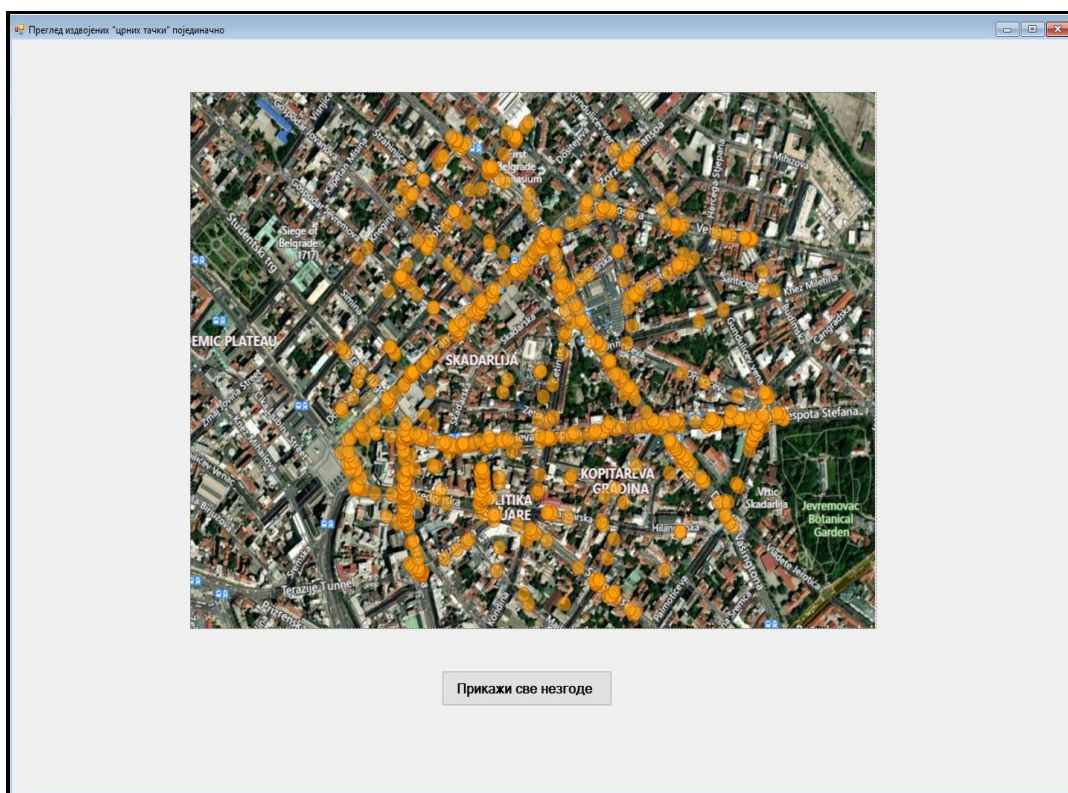


Слика 59. Изглед форме са идентификованим “црним тачкама” на изабраном подручју са прегледом свих саобраћајних незгода у кластери

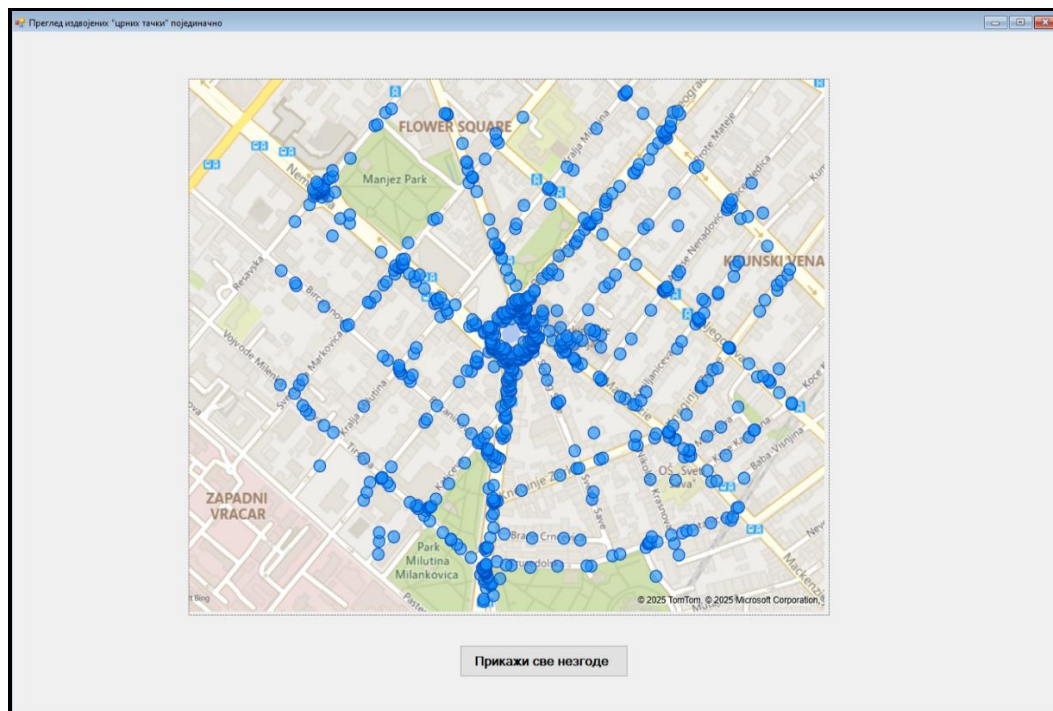
Апликација нуди могућност приказа издвојених локација на којима се догодио највећи број саобраћајних незгода за одређени временски период. Преглед је приказан на сликама од 60 до 63.



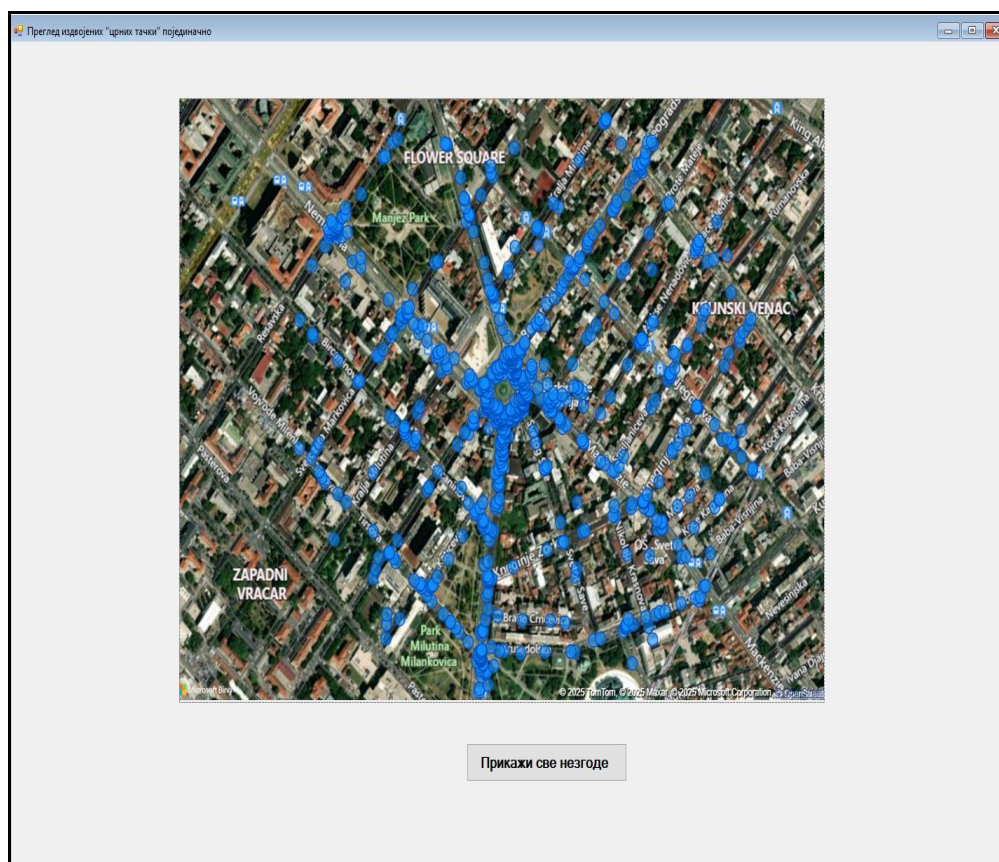
Слика 60. Изглед форме са идентификованом “црном тачком” са највећим бројем саобраћајних незгода у кластери полупречника 479 метара, улица Цетињска



Слика 61. Изглед форме са идентификованом “црном тачком” са највећим бројем саобраћајних незгода у кластеру полупречника 479 метара, са приказом локација свих саобраћајних незгода



Слика 62. Изглед форме са идентификованом “црном тачком” на кружном току Славија са приказом свих саобраћајних незгода у кластеру полупречника 479 метара

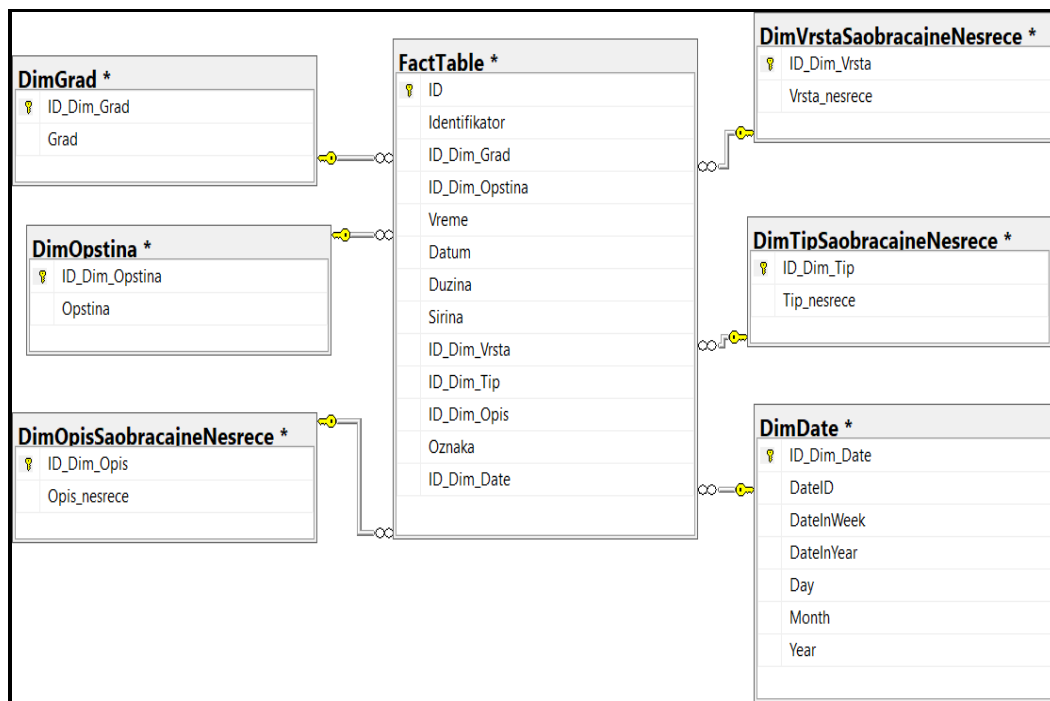


Слика 63. Изглед форме са идентификованом “црном тачком” на кружном току Славија са детаљним приказом свих саобраћајних незгода у кластеру полупречника 479 метара

8. Доказ хипотеза

Сагледавањем рада алгоритма машинског учења за одређивање кластера на основу просторних података, проучавањем доступне литературе и применом на конкретном примеру, уочено је да се модификацијама може проширити могућност његове примене. Унапређење се огледа у томе да се у скупу параметара пронађу такве варијабле, како нумеричке, тако и текстуалне, које се могу конвертовати у нумеричке вредности, и да се пронађу такви кластери код којих ће се, поред критеријума нагомилавања, као параметар за њихову идентификацију и издвајање користити и збирне вредности наведених пондерисаних вредности. Пожељно је да параметри са којима алгоритам ради буду што прецизнији, како би били добијени веродостојнији резултати. На тај начин могу се добити праведније вредности и прецизније идентификовати кластери на основу којих се касније могу доносити стратешке одлуке. Компаративна анализа рада алгоритма и резултати добијени над истим скупом података приказани су у поглављу 6.3. Ове чињенице су уједно и ***доказ посебне хипотезе да предложени и модификовани DBSCAN алгоритам користи параметре математичког модела и не захтева унапред познат број кластера, што доводи до прецизнијих резултата.***

У оквиру реализације система пословне интелигенције извршена је израда складишта података које је савремено конципирано и засновано на аналитичком моделу података, а у коме су обједињени, трансформисани и пречишћени подаци из више различитих извора. Модел података складишта података дат је на Слици 64. Оваква структура нам омогућава идентификацију кластера на основу ускладиштених података. Ове чињенице су уједно и ***доказ посебне хипотезе да се предложени и модификовани DBSCAN алгоритам може применити у системима пословне интелигенције заснованим на концепту складишта података.***



Слика 64. Модел података складишта података

У система пословне интелигенције укључени су сви параметри из изворног система, који су били доступни у оквиру скупа отворених података.

Обзиром да је локација саобраћајне незгоде приказана у виду координата у посебним колонама, додатно би могли бити приказани и подаци о путу (нпр. врста и стање коловоза, подаци о опреми пута, запрљаност коловоза). Такође, могли би бити интегрисани и подаци о општем стању возила, уколико се возило упућује на технички преглед (нпр. стање уређаја за кочење) и сл. Посебна категорија би могла да буде везана за возаче и опште податке о њима (нпр. године старости, занимање, дужина возачког стажа, да ли је викенд возач, да ли је био под утицајем алкохола или опијата, колико дуго је користио возило са којим је учествовао у незгоди). Такође, могу бити додате напомене, које би садржале податке за које органи који врше увиђај сматрају да су од значаја, а да нису обухваћени у некој од претходних колона.

На основу наведених чињеница, очигледно је да реализовани систем пружа могућност свеобухватне анализе и доношење одлука засновано на доступним улазним параметрима. Такође, отворене су могућности и за даља проширења, у складу са доступним подацима. Оваквом поставком података је **доказана и посебна хипотеза да се применом модификованог DBSCAN алгоритма побољшава идентификација опасних тачака на путевима и омогућава боља анализа података за доношење одлука.** Додатак овој тврдњи представља и генерисана табела која представља списак идентификованих опасних места на путевима у Републици Србији за изабрани трогодишњи временски период.

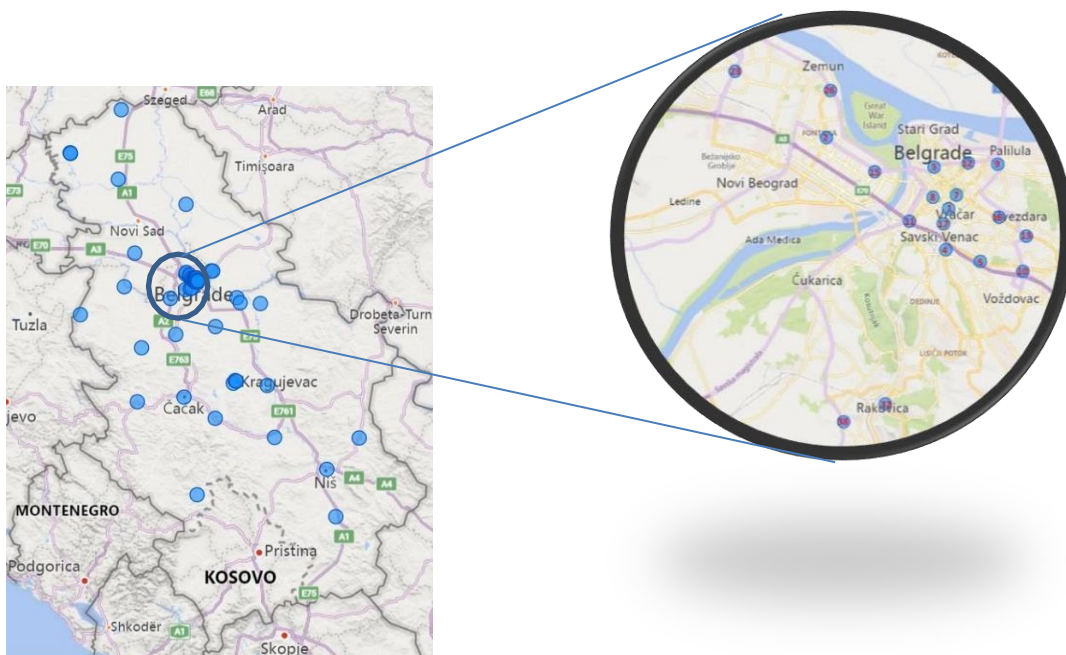
Табела 10. Списак идентификованих опасних места на путевима у Републици Србији за изабрани трогодишњи временски интервал

| Ранг ОМ | Општина ОМ | ПУ ОМ | Х коор. | У коор. | ПБСН | ПОГИНУЛИ | ПОВРЕЂЕНИ | МАТЕРИЈАЛНА ШТЕТА | ОПИС |
|------------|--------------|-----------|-------------|-------------|------|----------|-----------|----------------------|---|
| 1 | Палилула | Београд | 20.47682200 | 44.8171750 | 1290 | 1 | 90 | 305 | Улица Цвијићева |
| 2 | Врачар | Београд | 20.46786000 | 44.8024620 | 1277 | 2 | 83 | 277 | Улица Макензијева, код Трга Славија |
| 3 | Нови Београд | Београд | 20.41269300 | 44.8252820 | 1034 | 0 | 90 | 134 | Кружни ток који спаја улице Париске комуне, Улицу Омладинских бригада и Булевар Михаила Пупина |
| 4 | Стари Град | Београд | 20.46127000 | 44.81580000 | 848 | 0 | 67 | 178 | Раскрсница Улице Коларчеве и Трга Републике |
| 5 | Вождовац | Београд | 20.46660900 | 44.7893210 | 775 | 0 | 52 | 255 | Булевар ослобођења, деоница од Аутокоманде до укључења на аутопут |
| 6 | Вождовац | Београд | 20.48226800 | 44.7855610 | 707 | 1 | 51 | 112 | Раскрсница Улица Устаничке и Крушевачке |
| 7 | Чукарица | Београд | 20.41702500 | 44.7868650 | 686 | 1 | 45 | 151 | Деоница од раскрснице улица Петровићева и Радничка, до раскрснице улица Светолика Лазаревића Лазе и Савске магистрале |
| 8 | Врачар | Београд | 20.47140000 | 44.8068800 | 669 | 2 | 39 | 109 | Раскрсница Београдске и Булевара Краља Александра (Вуков споменик) |
| 9 | Савски Венац | Београд | 20.46078600 | 44.8061480 | 665 | 0 | 48 | 185 | Раскрсница Улице Кнеза Милоша и Улица Немањина (Лондон) |
| 10 | Палилула | Београд | 20.49021200 | 44.8167340 | 658 | 1 | 44 | 133 | Панчевачки мост |
| 11 | Вождовац | Београд | 20.50162500 | 44.7824000 | 642 | 0 | 53 | 112 | Раскрсница Војислава Илића и Устаничке улице |
| 12 | Савски Венац | Београд | 20.45017800 | 44.7984720 | 631 | 0 | 40 | 231 | Мостарска петља |
| 13 | Звездара | Београд | 20.50319000 | 44.7936650 | 580 | 1 | 41 | 85 | Раскрсница Булевара Краља Александра и Улице Господара Вучића |
| 14 | Ниш | Ниш | 21.90772200 | 43.3212730 | 579 | 3 | 29 | 34 | Деоница од раскрснице улица Војда Карађорђа и улице Војводе Мишића и раскрснице Улице Јована Скерлића |
| 15 | Нови Београд | Београд | 20.43445000 | 44.8143700 | 559 | 1 | 36 | 114 | Раскрсница улица Ушће, Милентија Поповића и Булевара Михаила Пупина |
| 16 | Звездара | Београд | 20.49078000 | 44.7997500 | 533 | 0 | 44 | 93 | Булевар Револуције |
| 17 | Врачар | Београд | 20.46551000 | 44.7976600 | 531 | 1 | 37 | 76 | Раскрсница улице Пастерове и Булевара ослобођења |
| 18 | Шабац | Шабац | 19.70242700 | 44.7489170 | 500 | 2 | 30 | 30 | Раскрсница Улица Милоша Поцерца и Улице Краља Александра |
| 19 | Пожаревац | Пожаревац | 21.18412300 | 44.61910000 | 448 | 0 | 38 | 68 | Раскрсница Улице Косовске и улице Моше Пијаде |

| | | | | | | | | | |
|----|------------|------------|-------------|-------------|-----|---|----|----|--|
| 20 | Обреновац | Обреновац | 20.20323200 | 44.6585330 | 438 | 1 | 32 | 33 | Раскрсница улица Београдски батаљони и |
| 21 | Ваљево | Ваљево | 19.89018000 | 44.27545000 | 438 | 1 | 33 | 23 | Раскрсница улица Пастерове и Синђелићеве |
| 22 | Смедерево | Смедерево | 20.92807600 | 44.66422400 | 438 | 0 | 39 | 48 | Раскрсница улаца Слободе и Кнеза Михаила |
| 23 | Земун | Београд | 20.37138000 | 44.84652000 | 434 | 0 | 34 | 94 | Искључење за Добановце, са аутопута за Нови Сад |
| 24 | Лесковац | Лесковац | 21.94405400 | 42.99631000 | 430 | 1 | 33 | 15 | Раскрсница Булевара ослобођења и улице Ивана Милутиновића |
| 25 | Краљево | Краљево | 20.69183300 | 43.72359800 | 428 | 1 | 31 | 33 | Раскрсница Улица Цара Душана и Војводе Путника |
| 26 | Земун | Београд | 20.41446300 | 44.84036400 | 411 | 0 | 35 | 61 | Раскрсница улица Карађорђева и Стевана Марковића |
| 27 | Лозница | Лозница | 19.22408300 | 44.53260900 | 397 | 0 | 38 | 17 | Деоница између раскрснице улица Пашићеве и Вере Благојевић и Кружног тока Вук Караџић |
| 28 | Јагодина | Јагодина | 21.25834800 | 43.98167100 | 391 | 2 | 17 | 51 | Раскрсница Улице Југ Бодганове и Кнеза Лазара |
| 29 | Земун | Београд | 20.73067600 | 44.84036400 | 370 | 1 | 22 | 65 | Раскрсница Улице Цара Душана и искључења за Пупинов мост |
| 30 | Крушевац | Крушевац | 21.33827900 | 43.57305700 | 362 | 1 | 27 | 7 | Кружни ток који спаја улице Благоја Паровића, Паје Јовановића, Војводе Степе, Цара Душана и Хајдук Вељка |
| 31 | Крагујевац | Крагујевац | 20.91033200 | 44.01996400 | 356 | 0 | 33 | 26 | Раскрсница улица потпоручника Говедарице и Војводе Путника |
| 32 | Раковица | Београд | 20.43925000 | 44.73988000 | 356 | 2 | 15 | 36 | Раскрсница Улица Патријарха Димитрија и Улице Ослободилаца Раковице |
| 33 | Сомбор | Сомбор | 19.11298700 | 45.77455500 | 354 | 1 | 26 | 9 | Раскрсница Венца Војводе Радомира Путника и Улице Војвођанске |
| 34 | Чачак | Чачак | 20.35245100 | 43.89012500 | 347 | 1 | 25 | 12 | Кружни ток који спаја Улицу Девет Југовића, улицу Железничку и Улицу Кужељеву. |
| 35 | Панчево | Панчево | 20.66487100 | 44.87040100 | 332 | 2 | 15 | 12 | Раскрсница Првوماјске и Милоша Обреновића |
| 36 | Крагујевац | Крагујевац | 20.92222000 | 44.01451000 | 328 | 1 | 23 | 13 | Раскрсница улице Града Сирена и улице Саве Ковачевића |
| 37 | Панчево | Панчево | 20.64746300 | 44.87048100 | 324 | 1 | 22 | 19 | Раскрсница улица Ђуре Ђаковића и Немањине |
| 38 | Раковица | Београд | 20.42042500 | 44.73413300 | 310 | 1 | 21 | 15 | Укључење на Ибарску Магистралу |
| 39 | Књажевац | Зајечар | 22.25942748 | 43.56849574 | 309 | 2 | 13 | 9 | Раскрсница Улица Кнеза Милоша и Капларова |
| 40 | Врбас | Нови Сад | 19.63672671 | 45.57318329 | 300 | 1 | 21 | 5 | Раскрсница улица Маршала Тита и Ивана Лоле Рибара |
| 41 | Сомбор | Сомбор | 19.11649600 | 45.76894100 | 297 | 0 | 28 | 17 | Улица Мите Поповића |
| 42 | Крагујевац | Крагујевац | 20.88888300 | 43.99947700 | 296 | 2 | 12 | 6 | Улица Краљевачког батаљона |

| | | | | | | | | | |
|----|------------|------------|-------------|-------------|-----|---|----|----|---|
| 43 | Рума | Рума | 19.81600100 | 45.00807700 | 296 | 0 | 29 | 6 | Раскрсница Улица Главна и Павловачка |
| 44 | Нови Пазар | Нови Пазар | 20.49532900 | 43.12006700 | 281 | 2 | 11 | 1 | Раскрсница улица Мур и Златиборска |
| 45 | Раља | Смедерево | 20.96416000 | 44.62529400 | 276 | 2 | 10 | 6 | Раскрсница Цариградске и Смедеревске |
| 46 | Младеновац | Београд | 20.69495147 | 44.44038588 | 268 | 0 | 23 | 38 | Раскрсница улица Краља Александра и Јанка Катића |
| 47 | Лазаревац | Београд | 20.26219000 | 44.37886000 | 259 | 1 | 16 | 14 | Раскрсница Улица Дула Караклајића и Бранка Радичевића |
| 48 | Суботица | Суботица | 19.66942000 | 46.10006400 | 251 | 0 | 23 | 21 | Раскрсница улица Максима Горког и Буре Баковића |
| 49 | Зрењанин | Зрењанин | 20.37401300 | 45.38184100 | 251 | 1 | 16 | 6 | Раскрсница улица Новосадска и Милана Станивуковића |
| 50 | Ужице | Ужице | 19.84456700 | 43.85363600 | 250 | 1 | 13 | 35 | Раскрсница Улице Липа и Улице Омладинске |

Предлог решења који је дат у раду реализован је системом пословне интелигенције и аутоматизованим процесом који учитава, трансформише, уписује, прорачунава, анализира, упоређује и исписује податке. Обзиром да је омогућено табеларно и графичко приказивање добијених резултата коришћењем креиране апликације, **доказана је хипотеза да се применом модификованог DBSCAN алгоритма омогућава лакше креирање и дељење знања о опасним тачкама међу доносиоцима одлука.**



Слика 65. Преглед идентификованих опасних места на путевима у Републици Србији за изабрани трогодишњи период

Код идентификације „црних тачака“ у постојећој методологији примењује се исти принцип за целу територију Републике Србије. Предложеним модификованим алгоритмом и креираном апликацијом могуће је идентификовати опасна места на путевима за одређену микролокацију, као што су општина, град или регион. У зависности од величине територије, може бити дефинисана одговарајућа гранична вредност пондерисаног броја саобраћајних незгода – k , при чему промена граничне вредности не би утицала на сам концепт примене модификованог алгоритма. У зависности од тога да ли се идентификација опасних места врши на националном нивоу или за одређене делове територије, одговарајући параметри у дефиницији опасног места – „црне тачке“ могу, али не морају нужно, бити промењени. Све наведено указује да је могућа релативно једноставна надоградња већ постојећих система пословне интелигенције, чиме је **доказана посебна хипотеза да**

модификовани DBSCAN алгоритам побољшава идентификацију опасних тачака и омогућава скалабилну примену на већим сетовима података.

Реализацијом истраживања о начину рада и примене алгоритама машинског учења за кластеровање над просторним подацима, који се интензивно развијају и све више примењују у области анализе података, истраживањем постојећих методологија за идентификацију опасних места на путевима и предлогом њихове симбиозе у циљу визуелизације и олакшаног коришћења при доношењу правовремених и праведнијих одлука у реалном времену, испуњен је циљ истраживања и дат је конкретан научно-истраживачки допринос у области информационих система, односно система за подршку у одлучивању.

На основу потврђених посебних хипотеза везаних за предлог примене модификованог алгоритма машинског учења за идентификацију опасних места на путевима коришћењем система пословне интелигенције, може се закључити да се његовом применом омогућава дефинисање и анализа података, дизајн и имплементација решења за анализу података, а све са циљем доношења одлука потребних управљању, ***што доказује основну хипотезу да се применом модификованог алгоритма машинског учења за идентификацију кластера над просторним подацима проширује могућност примене стандардног DBSCAN алгоритма и успешно су идентификована опасна места, “црне тачке” на путевима на територији Републике Србије, коришћењем јавно доступних података, што се успешно може користити за доношење одлука значајних за управљање.***

9. Научни допринос и дискусија резултата

У овом раду, полазиште је био *DBSCAN* алгоритам, један од најзначајнијих густинских метода кластеровања, али је кроз систематичне измене и увођење нових елемената развијена верзија која омогућава обраду комплексних, хетерогених и динамичких података.

У раду је развијен модел заснован на модификованом *DBSCAN* алгоритму, чија је основна новина увођење пондерисаних вредности у процес кластеризације. Док класични *DBSCAN* алгоритам идентификује кластере искључиво на основу густине тачака у простору, у модификованом моделу кластери се одређују и према “тежини” тачака, односно значају у контексту. На овај начин обезбеђена је прецизнија и поузданија идентификација кластера, јер резултати не одражавају само количину података, већ и њихов квалитет и релевантност.

Пондерисане вредности су уведене са циљем да се добије кластеризација која боље одражава реалне услове и омогућава да се кластери интерпретирају као структуре које имају и просторну густину и тежину. На тај начин постигнута је оптималнија идентификација и анализа, што је у раду и експериментално потврђено.

Модел је уједно и параметризован, што значи да се може прилагодити различитим типовима података и апликационим сценаријима. Ова флексибилност омогућава широку примену, али истовремено захтева пажљиво дефинисање параметара ради постизања веродостојних резултата.

Ограничење са којим сам се сусрео односи се на потребу за стандардизованим пондерима. Они не могу бити произвољно дефинисани, већ морају бити засновани на релевантним критеријумима и проверљивим изворима. То обезбеђује објективност модела, али истовремено захтева додатни напор у прикупљању и верификацији података.

Сумирано, у раду је:

- развијен модификовани *DBSCAN* алгоритам са пондерисаним вредностима,
- формулисани параметризовани модел који комбинује густину и тежину тачака,
- експериментално потврђено да модификовани алгоритам даје прецизније и поузданије резултате од класичног,
- и указано на ограничења која се односе на потребу за стандардизацијом пондера и пажљивим дефинисањем параметара.

9.1. Модификације у односу на стандардни DBSCAN

Класични *DBSCAN* алгоритам, иако снажан у откривању кластера произвољног облика и отпоран на шум, показује ограничења у условима хетерогених података и различитих густина. У модификованој верзији алгоритма уведено је динамичко подешавање ϵ параметра, које омогућава да се радијус околине прилагођава локалној густини података. Овај приступ омогућава да алгоритам истовремено препозна кластере у зонама високе и ниске густине, чиме се превазилази проблем фиксног ϵ . Поред тога, уведена је тежински приступ који свакој тачки додељује релевантност на основу њених атрибута, чиме се обезбеђује да кластери са значајнијим тачкама буду препознати као приоритетни. Алгоритам је проширен и временском димензијом, тако да кластероване више није ограничено на просторне координате, већ укључује и динамику кроз време. Коначно, развијена је мулти-тежинска верзија која комбинује више критеријума кроз систем тежина, чиме се добија свеобухватнији модел.

9.2. Нови елементи и параметри модела

Да би се превазишла ограничења класичног *DBSCAN-a*, у модификовани модел уведени су нови елементи који га чине прилагодљивијим и прецизнијим. Најзначајнији је динамички ϵ параметар, који се израчунава у односу на локалну густину и омогућава кластероване у зонама различите густине. Поред тога, уведени су тежински коефицијенти који свакој тачки додељују релевантност засновану на њеним атрибутима, чиме се обезбеђује да

анализа не фаворизује само бројност, већ и значајност података. Новина је и временска компонента, која омогућава кластеровање у временској димензији и идентификацију динамичких кластера. Коначно, мулти-тежински систем комбинује више фактора кроз систем тежина, што омогућава свеобухватнију анализу и добијање кластера који истовремено узимају у обзир густину, релевантност и динамику података.

9.3. Разлике у односу на постојећа решења

У литератури постоје бројне варијанте *DBSCAN-a*, као што су *Weighted DBSCAN*, *DBSCAN++* за стриминг податке, *Adaptive DBSCAN* за спатио-темпоралне скупове и хибридни модели који комбинују *DBSCAN* са другим техникама. Међутим, предложени модел се разликује у томе што интегрише више модификација у јединственом оквиру. Док постојеће варијанте решавају појединачне проблеме, предложени модел комбинује адаптивни ϵ , тежински приступ и временску компоненту у једном алгоритму. За разлику од класичних *Weighted DBSCAN* варијанти, овде се тежине не додељују само на основу једног критеријума, већ кроз комбинацију више фактора. Модел је дизајниран тако да може да се примени на различите типове података, што га чини универзалнијим у односу на специјализоване варијанте. На тај начин, предложени приступ представља интеграцију и проширење постојећих решења, а не само њихову модификацију.

9.4. Конкретна побољшања

Предложени модел доноси конкретна побољшања у односу на постојеће методе кластеровања. Адаптивни ϵ омогућава прецизније кластеровање у хетерогеним скуповима података, где класични *DBSCAN* није могао да постигне задовољавајуће резултате. Тежински приступ обезбеђује да кластери са релевантнијим тачкама буду препознати као приоритетни, а не само они са највећим бројем тачака. Интеграција временске димензије омогућава динамичку анализу кроз различите интервале, чиме се добија богатија и реалнија слика. Мулти-тежински систем омогућава свеобухватност кроз комбинацију више фактора, чиме се добија модел који истовремено узима у обзир густину, релевантност и динамику података. Ова побољшања чине модел практично применљивијим и научно релевантнијим, јер превазилази

ограничења класичног *DBSCAN*-а и постојећих варијанти, нудећи нови методолошки оквир за кластеровање комплексних података.

9.5. Допринос рада

Допринос рада може се јасно формулисати кроз више међусобно повезаних елемената који заједно чине новину истраживања. Први и основни допринос огледа се у развоју модификованог *DBSCAN* алгоритма. Уведен је нови модел кластеризације који укључује пондерисане параметре објеката, односно предмета кластеризације. Тежина сваке тачке дефинисана је као комбинација пондерисаних вредности, чиме кластеризација више не зависи искључиво од броја тачака, већ од њиховог реалног значаја у контексту. На овај начин обезбеђена је прецизнија идентификација кластера и боља репрезентација података.

Други допринос представља увођење математичког модела за процену значаја кластера. Формулисан је модел који комбинује пондерисану вредност кластера, број унутрашњих и спољашњих тачака, као и просечну вредност растојања између преклапајућих кластера. Овај модел омогућава да се кластери рангирају по тежини и сложености, чиме се добија прецизнија процена њиховог значаја. На тај начин кластери се не процењују само по количини података, већ и по квалитету информација које садрже, што представља значајан корак напред у односу на постојеће приступе.

Трећи допринос огледа се у интеграцији алгоритма са системом пословне интелигенције и складиштем података. Алгоритам је повезан са *Data Warehouse* системом кроз *ETL* процес, где се текстуалне категорије конвертују у нумеричке пондере. Резултати кластеризације смештени су у *OLAP* структуре, што омогућава интерактивну анализу и визуелизацију кроз *BI* алате. Ова интеграција обезбеђује да алгоритам није само теоријски модел, већ део практичног система за подршку одлучивању, чиме се његова примена проширује на реалне пословне сценарије.

Четврти допринос представља развијена методологија за идентификацију и визуелизацију „црних тачака“. Користећи јавно доступне податке, кроз примену *MW-DBSCAN* алгоритма добијају се кластери који се рангирају по значају и приказују на географским мапама. Овај приступ омогућава да се резултати кластеризације представе у визуелно разумљивом

формату, што значајно доприноси процесу доношења одлука и повећава практичну вредност истраживања.

Сумирано, допринос рада огледа се у развоју модификованог *DBSCAN* алгоритма са пондерисаним параметрима, увођењу математичког модела за процену значаја кластера, интеграцији са системом пословне интелигенције и складиштем података, као и у развоју методологије за идентификацију и визуелизацију критичних тачака. Ови елементи заједно чине јединствену целину која обезбеђује нову вредност у области кластеризације и пословне аналитике.

9.6. Потврда научног доприноса кроз практичну примену

Сагледавањем домаће и стране литературе, која се бави наведеном проблематиком, дат је предлог једаног новог приступа за идентификацију “црних тачака”. Предлог је да се коришћењем моћних система за обраду података, јавно доступним подацима да нова улога у идентификацији опасних места, првенствено коришћењем система пословне интелигенције, где се корак по корак пролазити кроз алгоритам, идентификују и по потреби елиминишу одређене локације, да би се на крају, изолованим поступцима дошло до најучесталијих места где се догађају саобраћајне незгоде. Међусобна компарација се врши према последицама које су изазване саобраћајном незгодом. Једна од битних карактеристика је та што се за одређивање опасних места не узимају само локације на којима су се догодиле саобраћајне незгоде са погинулим лицима, што је до сада рађено, већ се опсег посматрања проширује и на места на којима се учестало догађају и саобраћајне незгоде са повређенима, и по последицама су рангиране након саобраћајних незгода са страдалим лицима. Ово представља знатно унапређење, али са друге стране знатно усложњава поступак идентификације опасних места, јер је потребно извршити анализу огромне количине података. У Табели 11 приказан је предлог идентификованих опасних места, а црвеном бојом су означена опасна места, добијена предложеном методологијом, која нису могла бити одређена применом других методологија.

Табела 11. Списак идентификованих опасних места на путевима у Републици Србији за изабрани трогодишњи временски интервал са обележеним локацијама које не би биле идентификоване са тренутно примењеном методологијом

| Ранг ОМ | Општина ОМ | ПУ ОМ | Х коор. | У коор. | ПБСН | ПОГИНУЛИ | ПОВРЕЂЕНИ | МАТЕРИЈАЛНА ШТЕТА | ОПИС |
|------------|--------------|-----------|-------------|-------------|------|----------|-----------|----------------------|---|
| 1 | Палилула | Београд | 20.47682200 | 44.8171750 | 1290 | 1 | 90 | 305 | Улица Цвијићева |
| 2 | Врачар | Београд | 20.46786000 | 44.8024620 | 1277 | 2 | 83 | 277 | Улица Макензијева, код Трга Славија |
| 3 | Нови Београд | Београд | 20.41269300 | 44.8252820 | 1034 | 0 | 90 | 134 | Кружни ток који спаја улице Париске комуне, Улицу Омладинских бригада и Булевар Михаила Пупина |
| 4 | Стари Град | Београд | 20.46127000 | 44.81580000 | 848 | 0 | 67 | 178 | Раскрсница Улице Коларчеве и Трга Републике |
| 5 | Вождовац | Београд | 20.46660900 | 44.7893210 | 775 | 0 | 52 | 255 | Булевар ослобођења, деоница од Аутокоманде до укључења на аутопут |
| 6 | Вождовац | Београд | 20.48226800 | 44.7855610 | 707 | 1 | 51 | 112 | Раскрсница Улица Устаничке и Крушевачке |
| 7 | Чукарица | Београд | 20.41702500 | 44.7868650 | 686 | 1 | 45 | 151 | Деоница од раскрснице улица Петровићева и Радничка, до раскрснице улица Светолика Лазаревића Лазе и Савске магистрале |
| 8 | Врачар | Београд | 20.47140000 | 44.8068800 | 669 | 2 | 39 | 109 | Раскрсница Београдске и Булевара Краља Александра (Вуков споменик) |
| 9 | Савски Венац | Београд | 20.46078600 | 44.8061480 | 665 | 0 | 48 | 185 | Раскрсница Улице Кнеза Милоша и Улица Немањина (Лондон) |
| 10 | Палилула | Београд | 20.49021200 | 44.8167340 | 658 | 1 | 44 | 133 | Панчевачки мост |
| 11 | Вождовац | Београд | 20.50162500 | 44.7824000 | 642 | 0 | 53 | 112 | Раскрсница Војислава Илића и Устаничке улице |
| 12 | Савски Венац | Београд | 20.45017800 | 44.7984720 | 631 | 0 | 40 | 231 | Мостарска петља |
| 13 | Звездара | Београд | 20.50319000 | 44.7936650 | 580 | 1 | 41 | 85 | Раскрсница Булевара Краља Александра и Улице Господара Вучића |
| 14 | Ниш | Ниш | 21.90772200 | 43.3212730 | 579 | 3 | 29 | 34 | Деоница од раскрснице улица Војда Карађорђа и улице Војводе Мишића и раскрснице Улице Јована Скерлића |
| 15 | Нови Београд | Београд | 20.43445000 | 44.8143700 | 559 | 1 | 36 | 114 | Раскрсница улица Ушће, Милентија Поповића и Булевара Михаила Пупина |
| 16 | Звездара | Београд | 20.49078000 | 44.7997500 | 533 | 0 | 44 | 93 | Булевар Револуције |
| 17 | Врачар | Београд | 20.46551000 | 44.7976600 | 531 | 1 | 37 | 76 | Раскрсница улице Пастерове и Булевара ослобођења |
| 18 | Шабац | Шабац | 19.70242700 | 44.7489170 | 500 | 2 | 30 | 30 | Раскрсница Улица Милоша Поцерца и Улице Краља Александра |
| 19 | Пожаревац | Пожаревац | 21.18412300 | 44.61910000 | 448 | 0 | 38 | 68 | Раскрсница Улице Косовске и улице Моше Пијаде |

| | | | | | | | | | |
|----|------------|------------|-------------|-------------|-----|---|----|----|--|
| 20 | Обреновац | Обреновац | 20.20323200 | 44.6585330 | 438 | 1 | 32 | 33 | Раскрсница улица Београдски батаљони и |
| 21 | Ваљево | Ваљево | 19.89018000 | 44.27545000 | 438 | 1 | 33 | 23 | Раскрсница улица Пастерове и Синђелићеве |
| 22 | Смедерево | Смедерево | 20.92807600 | 44.66422400 | 438 | 0 | 39 | 48 | Раскрсница улаца Слободе и Кнеза Михаила |
| 23 | Земун | Београд | 20.37138000 | 44.84652000 | 434 | 0 | 34 | 94 | Искључење за Добановце, са аутопута за Нови Сад |
| 24 | Лесковац | Лесковац | 21.94405400 | 42.99631000 | 430 | 1 | 33 | 15 | Раскрсница Булевара ослобођења и улице Ивана Милутиновића |
| 25 | Краљево | Краљево | 20.69183300 | 43.72359800 | 428 | 1 | 31 | 33 | Раскрсница Улица Цара Душана и Војводе Путника |
| 26 | Земун | Београд | 20.41446300 | 44.84036400 | 411 | 0 | 35 | 61 | Раскрсница улица Карађорђева и Стевана Марковића |
| 27 | Лозница | Лозница | 19.22408300 | 44.53260900 | 397 | 0 | 38 | 17 | Деоница између раскрснице улица Пашићеве и Вере Благојевић и Кругног тока Вук Караџић |
| 28 | Јагодина | Јагодина | 21.25834800 | 43.98167100 | 391 | 2 | 17 | 51 | Раскрсница Улице Југ Бодганове и Кнеза Лазара |
| 29 | Земун | Београд | 20.73067600 | 44.84036400 | 370 | 1 | 22 | 65 | Раскрсница Улице Цара Душана и искључења за Пупинов мост |
| 30 | Крушевац | Крушевац | 21.33827900 | 43.57305700 | 362 | 1 | 27 | 7 | Кружни ток који спаја улице Благоја Паровића, Паје Јовановића, Војводе Степе, Цара Душана и Хајдук Велка |
| 31 | Крагујевац | Крагујевац | 20.91033200 | 44.01996400 | 356 | 0 | 33 | 26 | Раскрсница улица потпоручника Говедарице и Војводе Путника |
| 32 | Раковица | Београд | 20.43925000 | 44.73988000 | 356 | 2 | 15 | 36 | Раскрсница Улица Патријарха Димитрија и Улице Ослободилаца Раковице |
| 33 | Сомбор | Сомбор | 19.11298700 | 45.77455500 | 354 | 1 | 26 | 9 | Раскрсница Венца Војводе Радомира Путника и Улице Војвођанске |
| 34 | Чачак | Чачак | 20.35245100 | 43.89012500 | 347 | 1 | 25 | 12 | Кружни ток који спаја Улицу Девет Југовића, улицу Железничку и Улицу Кужељеву. |
| 35 | Панчево | Панчево | 20.66487100 | 44.87040100 | 332 | 2 | 15 | 12 | Раскрсница Првوماјске и Милоша Обреновића |
| 36 | Крагујевац | Крагујевац | 20.92222000 | 44.01451000 | 328 | 1 | 23 | 13 | Раскрсница улице Града Сирена и улице Саве Ковачевића |
| 37 | Панчево | Панчево | 20.64746300 | 44.87048100 | 324 | 1 | 22 | 19 | Раскрсница улица Ђуре Ђаковића и Немањине |
| 38 | Раковица | Београд | 20.42042500 | 44.73413300 | 310 | 1 | 21 | 15 | Укључење на Ибарску Магистралу |
| 39 | Књажевац | Зајечар | 22.25942748 | 43.56849574 | 309 | 2 | 13 | 9 | Раскрсница Улица Кнеза Милоша и Капларова |
| 40 | Врбас | Нови Сад | 19.63672671 | 45.57318329 | 300 | 1 | 21 | 5 | Раскрсница улица Маршала Тита и Ивана Лоле Рибара |
| 41 | Сомбор | Сомбор | 19.11649600 | 45.76894100 | 297 | 0 | 28 | 17 | Улица Мите Поповића |
| 42 | Крагујевац | Крагујевац | 20.88888300 | 43.99947700 | 296 | 2 | 12 | 6 | Улица Краљевачког батаљона |

| | | | | | | | | | |
|----|------------|------------|-------------|-------------|-----|---|----|----|---|
| 43 | Рума | Рума | 19.81600100 | 45.00807700 | 296 | 0 | 29 | 6 | Раскрсница Улица Главна и Павловачка |
| 44 | Нови Пазар | Нови Пазар | 20.49532900 | 43.12006700 | 281 | 2 | 11 | 1 | Раскрсница улица Мур и Златиборска |
| 45 | Раља | Смедерево | 20.96416000 | 44.62529400 | 276 | 2 | 10 | 6 | Раскрсница Цариградске и Смедеревске |
| 46 | Младеновац | Београд | 20.69495147 | 44.44038588 | 268 | 0 | 23 | 38 | Раскрсница улица Краља Александра и Јанка Катића |
| 47 | Лазаревац | Београд | 20.26219000 | 44.37886000 | 259 | 1 | 16 | 14 | Раскрсница Улица Дула Караклајића и Бранка Радичевића |
| 48 | Суботица | Суботица | 19.66942000 | 46.10006400 | 251 | 0 | 23 | 21 | Раскрсница улица Максима Горког и Ђуре Ђаковића |
| 49 | Зрењанин | Зрењанин | 20.37401300 | 45.38184100 | 251 | 1 | 16 | 6 | Раскрсница улица Новосадска и Милана Станивуковића |
| 50 | Ужице | Ужице | 19.84456700 | 43.85363600 | 250 | 1 | 13 | 35 | Раскрсница Улице Липа и Улице Омладинске |

Приступ проблематици модификације алгоритма машинског учења и проширење могућности његове примене је мултидисциплинаран и повезује савремене ИКТ и науку о подацима. Наука о подацима, као и алати који су коришћени како за сакупљање, тако и за обраду података у самом систему, могу дати боље увиде у места која су опасна и предњаче по броју и последицама саобраћајних незгода. Самим тим, омогућено је ефикасније доношење одлука у систему, а по питању правовременог и одговарајућег одговора самог система.

Као продукт предложеног научноистраживачког рада издваја се:

- Дефинисање и модификација алгоритма машинског учења за кластеровање над просторним подацима, чиме је проширена могућност његове примене и створени нови услови за његово коришћење у оквиру система пословне интелигенције и може се применити за идентификацију опасних места на путевима у Републици Србији;
- Прикупљање, систематизација и критичка анализа постојећих решења, знања, искустава и научних резултата који се односе на идентификацију опасних места на путевима у Републици Србији;
- Развој и примена статистичких модела за анализу података.

Захваљујући адекватно планираном приступу, добијени су поуздани резултати. Својом флексибилношћу и све већом популарношћу алгоритми машинског учења омогућавају ефикасан развој и примену, за добијање информација у најкраћем времену.

Практичан допринос лежи у чињеници да је резултат дисертације модификовани алгоритам машинског учења за кластеровање над просторним подацима и омогућава анализу и откривање знања у оквиру доступних података. То подразумева стављање информационих ресурса на располагање доносиоцима одлука који могу у сваком тренутку, интерактивно да добијају жељене извештаје уз минималну подршку информатичке службе. Тиме се значајно смањује временско кашњење између идентификације евентуалних проблема и предузимања неопходних акција како би се ти проблеми решили.

Основна поставка дисертације је везана за проширење могућности

примене алгоритма машинског учења за идентификацију кластера над просторним подацима, који се може користити у оквиру система пословне интелигенције и на тај начин је могуће остваривање бенефита при доношењу одлука на највишем нивоу. Примена је прилагођена развоју система пословне интелигенције у условима када већ постоји претходно акумулирано знање о проблему.

9.7. Ограничења актуелног истраживања и могући правци за будуће истраживање

Иако је дисертација дала значајан допринос кроз развој модификованог *DBSCAN* алгоритма, формулисање математичког модела за процену значаја кластера и интеграцију са системом пословне интелигенције, постоје одређена ограничења која је неопходно нагласити ради објективне процене резултата и дефинисања будућих истраживачких праваца.

Прво ограничење огледа се у избору параметара и пондера. Иако предложени математички модел омогућава процену значаја кластера на основу више критеријума, вредности пондера су у великој мери условљене контекстом у коме се примењују. То значи да резултати могу бити осетљиви на промену тежина и да је потребно додатно истражити методе за динамичко или аутоматско одређивање пондера. У будућим истраживањима посебан акценат треба ставити на развој адаптивних механизма који би омогућили да се тежине параметара прилагођавају специфичним условима и захтевима апликационог окружења.

Друго ограничење односи се на интеграцију са системом пословне интелигенције. Иако је алгоритам успешно повезан са *Data Warehouse* и *OLAP* структурама, степен аутоматизације *ETL* процеса и скалабилност решења нису у потпуности истражени. У пракси то значи да се при већим количинама података или у условима реалног времена могу јавити проблеми у перформансама и одрживости система. Будућа истраживања треба да се усмере на оптимизацију *ETL* процеса, повећање степена аутоматизације и испитивање могућности примене модела у условима обраде података у реалном времену.

Могући правци будућег истраживања укључују проширење модела на друге типове података и апликационе домене (нпр. здравство, финансије, енергетика), развој адаптивних механизма за динамичко одређивање пондера,

унапређење *ETL* процеса и интеграције са *VI* системима ради постизања веће скалабилности и аутоматизације, као и примену модела у реалном времену. Поред тога, потребно је наставити истраживање визуелизације резултата кроз напредне интерактивне интерфејсе, који би омогућили корисницима да брже и лакше интерпретирају добијене кластере.

Сумирано, ограничења актуелног истраживања не умањују његов допринос, већ представљају полазну основу за будућа истраживања која могу додатно унапредити алгоритам, проширити његову примену и обезбедити још већи степен поузданости и корисности у различитим областима.

10. Закључак

Резултати истраживања показали су да је могуће унапредити постојећи *DBSCAN* алгоритам кроз систематичне модификације које га чине применљивијим у условима хетерогених и динамичких података. Адаптивно подешавање ε параметра омогућило је да се кластери формирају у зонама различите густине без губитка прецизности, тежински приступ је обезбедио да кластери са релевантнијим тачкама буду препознати као приоритетни, док је интеграција временске димензије омогућила идентификацију динамичких образаца који се мењају током различитих интервала. На тај начин потврђене су постављене хипотезе да је могуће развити алгоритам који истовремено узима у обзир густину, релевантност и динамику података.

Анализа постојећих решења у литератури показала је да предложени модел представља интеграцију више модификација у јединственом оквиру, чиме се разликује од варијанти које решавају само појединачне проблеме. Практична примена потврдила је да нови модел доноси конкретна побољшања у односу на класични *DBSCAN* и сродне методе, јер омогућава свеобухватнију анализу, повећава тачност идентификације кластера и пружа нови методолошки оквир за кластероване комплексних података.

Посебан значај истраживања огледа се у чињеници да је предложени приступ у складу са савременим трендовима у области пословне интелигенције. Истраживања показују да се у предузећима без система пословне интелигенције анализира тек око 10% расположивих података, док се увођењем напредних аналитичких метода тај проценат повећава и до 90% [135]. Овај налаз потврђује да развој нових алгоритамских решења није само теоријски допринос, већ има и практичну вредност у повећању искоришћености података и подршци процесу доношења одлука.

Сумирано, дисертација је дала јасне доказе о предностима модификованог *DBSCAN* модела, потврдила постављене хипотезе и показала да је могуће развити универзални и флексибилан алгоритам кластероване који интегрише више критеријума. Научни допринос огледа се у томе што предложени модел омогућава добијање поузданијих резултата, пружа нову вредност за даља истраживања и има практичну примену у различитим областима где је потребна анализа великих и сложених скупова података.

Литература

- [1] Вулић И. : Примена ГИС у лоцирању извора радио таласа, магистарски рад, Електронски факултет, Ниш, 1998.
- [2] Babić I, Miljković A, Čabarkapa M, Nikolić V, Đorđević A, Ranđelović M, Ranđelović D. Triple Modular Redundancy Optimization for Threshold Determination in Intrusion Detection Systems. *Symmetry*. 2021; 13(4):557. <https://doi.org/10.3390/sym13040557>
- [3] Pandey, T. N., Jagadev, A. K., Mohapatra, S. K., & Dehuri, S. (2017, August). Credit risk analysis using machine learning classifiers. In 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS) (pp. 1850-1854). IEEE
- [4] Data Mining From A to Z: How to Discover Insights and Drive Better Opportunities, SAS Institute Inc., Cary, North Carolina, USA, 2016.
- [5] Veljović, A., Radojičić, M., & Vesić, J. (2006). Menadžment informacioni sistemi. Viša tehnička škola za industrijski menadžment, ICIM-Izdavački centar za industrijski menadžment.
- [6] Bojčić, I. M. (2017). Modelom vođen razvoj skladišta podataka zasnovanog na Data Vault pristupu (Doctoral dissertation, Univerzitet u Beogradu, Fakultet organizacionih nauka).
- [7] Wickham, H. (2019). Data science: how is it different to statistics?. *IMS Bulletin*, 48.
- [8] Battula, B. P., Rama Krishna, K. V. S. S., & Kim, T. H. (2015). An efficient approach for knowledge discovery in decision trees using inter quartile range transform. *International Journal of Control and Automation*, 8(7), 325-334.
- [9] Развој софтверске апликације и методологије за одређивање опасних места – „црних тачака“ на путевима Републике Србије, Агенција за безбедност саобраћаја, Београд 2016.
- [10] Computing Curricula 2005, IEEE Computer Society, Los Alamitos, California, USA, 2005.
- [11] Computer Science Curricula 2013, IEEE Computer Society, Los Alamitos, California, USA, 2013.
- [12] Stolic, P.; Stevic, Z.; Petronic, S.; Nikolic, V.; Stevic, M.; Kreculj, D.; Milosevic, D. Modeling, Simulation, and Computer Control of a High-Frequency Wood Drying System. *Electronics* 2023, 12, 226. <https://doi.org/10.3390/electronics12010226>
- [13] Elvik, R. (2008). A survey of operational definitions of hazardous road locations in some European countries. *Accident Analysis & Prevention*, 40(6), 1830-1835.
- [14] P. Subramanian, S. Palaniappan, M.Vedamanikam, N. Ujwal Bharat, and S.Sinnasamy, “EXPLORING THE EFFECTIVENESS OF DATA MINING TECHNIQUES FOR LUNG CANCER PREDICTION,” 2016.
- [15] Bhargava, N., Sharma, G., Bhargava, R., & Mathuria, M. (2013). Decision tree analysis on j48 algorithm for data mining. *Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering*, 3(6).
- [16] Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, 1165-1188
- [17] Wang, J., & Gu, L. (2016). Challenges of teaching data science in a business

- school. *Issues in Information Systems*, 17(3).
- [18] Mikulik, J., & Hollo, P. (2007). Road accident investigation guidelines for road engineers. World Road Association PIRAC Technical Committee.
- [19] Lloyd, D., Wilson, D., Mais, D., Deda, W., & Bhagat, A. (2015). Reported road casualties Great Britain: 2014 annual report.
- [20] Кукић, Д., Милошевић, Ј., Милинковић, Б., Цвијовић, Р. (2016). Унапређење базе података о саобраћајним незгодама у складу са CADaS препорукама Европске комисије – значај за локалну заједницу. XI међународна конференција „Безбедност саобраћаја у локалној заједници“. Врњачка Бања. Србија.
- [21] Elvik, R. (2008). Comparative analysis of techniques for identifying locations of hazardous roads. *Transportation Research Record*, 2083(1), 72-75.
- [22] Нешић, М. (2015). Методи идентификације опасних места-црних тачака на ванградским путевима – докторска дисертација, Саобраћајни факултет у Београду.
- [23] Анђелковић, Д. (2019). Идентификација опасних места на путевима применом континуалне дисперзионе анализе (Докторска дисертација, Универзитет у Новом Саду, Факултет техничких наука у Новом Саду). <https://nardus.mpn.gov.rs/handle/123456789/11405>
- [24] Bíl, M., Andrášik, R., & Sedoník, J. (2019). A detailed spatiotemporal analysis of traffic crash hotspots. *Applied geography*, 107, 82-90. DOI: <https://doi.org/10.1016/j.apgeog.2019.04.008>
- [25] Sørensen, M. (2007). Best practice guidelines on black spot management and safety analysis of road networks. Oslo: Transportøkonomisk institutt.
- [26] Karamanlis, I., Kokkalis, A., Profillidis, V., Botzoris, G., & Galanis, A. (2023). Identifying Road Accident Black Spots using Classical and Modern Approaches. *WSEAS Transactions on Systems*, 22, 556-565. DOI: 10.37394/23202.2023.22.56
- [27] Debrabant, B., Halekoh, U., Bonat, W. H., Hansen, D. L., Hjelmberg, J., & Lauritsen, J. (2018). Identifying traffic accident black spots with Poisson-Tweedie models. *Accident Analysis & Prevention*, 111, 147-154. DOI: <https://doi.org/10.1016/j.aap.2017.11.021>
- [28] Zhang, C., Shu, Y., & Yan, L. (2019). A novel identification model for road traffic accident black spots: a case study in ningbo, China. *IEEE Access*, 7, 140197-140205. DOI: <https://doi.org/10.1109/ACCESS.2019.2942647>
- [29] Setiawan, A., Ramlan, R., Labaso, E. R., & Nur, A. A. (2022, November). Identification and treatment of black spots in the road network around Universitas Tadulako. In *IOP Conference Series: Earth and Environmental Science* (Vol. 1075, No. 1, p. 012042). IOP Publishing. DOI: 10.1088/1755-1315/1075/1/012042
- [30] Treeranurat, W., & Suanmali, S. (2021). Determination of black spots by using accident equivalent number and upper control limit on rural roads of Thailand. *Engineering Management in Production and Services*, 13(4), 57-71. DOI: [10.2478/emj-2021-0031](https://doi.org/10.2478/emj-2021-0031)
- [31] Yadav, D. K., Ghodmare, S. D., & Kumar, N. N. (2022). Mitigation of Blackspots on Highways by the Application of Safe System Approach. *Materials Today: Proceedings*, 52, 1228-1235. DOI: <https://doi.org/10.1016/j.matpr.2021.11.042>
- [32] Elvik, R. (2009). Comparative analysis of techniques for identifying locations of hazardous roads. *Transportation Research Record*, 2084(1), 82-85.

- [33] Cheng, W., & Washington, S. P. (2005). Experimental evaluation of hotspot identification methods. *Accident Analysis & Prevention*, 37(5), 870-881.
- [34] Miranda-Moreno, L. F., Labbe, A., & Fu, L. (2007). Bayesian multiple testing procedures for hotspot identification. *Accident Analysis & Prevention*, 39(6), 1192-1201.
- [35] Montella, A. (2010). A comparative analysis of hotspot identification methods. *Accident Analysis & Prevention*, 42(2), 571-581.
- [36] Ambros, J., Havránek, P., Valentová, V., Křivánková, Z., & Striegler, R. (2016). Identification of hazardous locations in regional road network—Comparison of reactive and proactive approaches. *Transportation Research Procedia*, 14, 4209-4217.
- [37] Zhang, C., He, J., Wang, H., Ye, Y., Yan, X., Wang, C., & Zhang, X. (2025). A systematic review of the application and prospect of road accident blackspots identification approaches. *Transportation Letters*, 17(6), 1114-1137.
- [38] Wan, Y., He, W., & Zhou, J. (2021). Urban road accident black spot identification and classification approach: A novel grey verhuls–Empirical bayesian combination method. *Sustainability*, 13(20), 11198.
- [39] Dhawale, N. (2022). Modifying an Old Rugged Jeep to Develop a Platform for Agriculture Purpose and for Proximal and Remote Soil Sensing Usage. *International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)*.
- [40] Dong, C., & Chang, N. (2023). Overview of the identification of traffic accident-prone locations driven by big data. *Digital Transportation and Safety*, 2(1), 67-76. DOI: <https://doi.org/10.48130/DTS-2023-0006>
- [41] Laanaoui, M. D., Lachgar, M., Mohamed, H., Hamid, H., Villar, S. G., & Ashraf, I. (2024). Enhancing urban traffic management through real-time anomaly detection and load balancing. *Ieee Access*, 12, 63683-63700.
- [42] Liang, G., Sun, X., Zhang, Y., Chen, M., & Zhang, W. (2021). Identifying expressway accident black spots based on the secondary division of road units. *Promet-Traffic&Transportation*, 33(5), 731-743. <https://doi.org/10.7307/ptt.v33i5.3680>
- [43] Elvik, R. (2007). State-of-the-art approaches to road accident black spot management and safety analysis of road networks. Oslo, Norway: Transportøkonomisk institutt.
- [44] Haque, M. I., Khan, A. A., Singh, G., Dass, S., & Jaglan, S. (2022, September). Accident analysis and road safety audit: a case study on NH-76. In *IOP Conference Series: Earth and Environmental Science* (Vol. 1086, No. 1, p. 012025). IOP Publishing. DOI 10.1088/1755-1315/1086/1/012025 https://statdataanalytics.files.wordpress.com/2015/02/dwbi_post1.pdf, pristupljeno 18.07.2025.
- [45] Haghghi, F., & Karimi, E. (2018). Evaluation and statistical validation of black-spots identification methods. *International Journal of Transportation Engineering*, 6(1), 1-15. DOI: <https://doi.org/10.22119/ijte.2017.52977>
- [46] Jibril, M. S., Aule, D. S., & Garba, B. D. (2023). The Spatial Patterns of Road Traffic Crash Black Spots and Emergency Facilities in Federal Capital City. *Journal of Geoscience and Environment Protection*, 11, 121-134. DOI: <https://doi.org/10.4236/gep.2023.112008>
- [47] PIARC. (2007). Road Accident Investigation Guidelines for Road Engineers.
- [48] Yuan, T., Zeng, X., & Shi, T. (2020). Identifying urban road black spots with a novel method based on the firefly clustering algorithm and a geographic

- information system. Sustainability, 12(5), 2091. DOI: <https://doi.org/10.3390/su12052091>
- [49] PIARC. (2008). Road Safety Manual. Recommendations from the World Road Association. Chapter 7 (Priority ranking).
- [50] Пројекат: „Идентификација опасних места на државним путевима Републике Србије са предлогом мера санације“, ЈП „Путеви Србије“, МУП, Министарство за инфраструктуру, службе надзора и предузећа за путеве, 2007.
- [51] Пројекат: „Упоредна анализа опасних места и ризичних деоница на државним путевима“, Криминалистичко-полицијска академија, 2010.
- [52] Пројекат: „Идентификација и класификација опасних места „црних тачака“ у саобраћају на подручју града Београда, са предлогом санационих мера“, Криминалистичко-полицијска академија, 2010.
- [53] Пројекат: „Методологија идентификације и управљања опасним местима (црним тачкама) на државним путевима Републике Србије“, Саобраћајни факултет, 2010.
- [54] Cui, H., Dong, J., Zhu, M., Li, X., & Wang, Q. (2022). Identifying accident black spots based on the accident spacing distribution. Journal of traffic and transportation engineering (English edition), 9(6), 1017-1026. DOI: <https://doi.org/10.1016/j.jtte.2021.02.006>
- [55] Кукић, Д., Петровић, Д., Тешић, М., Милинковић, Б. (2014). Поређење алата за оцену безбедности пута IA-2 „Ибарска магистрала“ од Београда до Чачка, IX међународна конференција „Безбедност саобраћаја у локалној заједници“, Зборник радова стр. 331-336, Зајечар.
- [56] Leskovec, J., Rajaraman, A., & Ullman, J. D. (2020). Mining of massive data sets. Cambridge university press.
- [57] Отворени подаци, <https://data.gov.rs/sr/discover/>, приступљено 30.06.2025. године.
- [58] Mundy, J., & Thornthwaite, W. (2011). The Microsoft data warehouse toolkit: with SQL Server 2008 R2 and the Microsoft Business Intelligence toolset. John Wiley & Sons.
- [59] Isazad Mashinchi, M., Ojo, A., & Sullivan, F. J. (2019, January). Analysis of Business Intelligence Applications in Healthcare Organizations. In Proceedings of the 52nd Hawaii International Conference on System Sciences.
- [60] <https://samoobrazovanje.rs/sta-je-inteligencija/>, pristupljeno 21.07.2025.
- [61] Suknović, M., & Delibašić, B. (2010). Poslovna inteligencija i sistemi za podršku odlučivanju. Beograd: FON.
- [62] Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., & Becker, B. (2008). The data warehouse lifecycle toolkit. John Wiley & Sons.
- [63] Gao, J., Lin, S., & Koronios, A. (2006). Data Quality in Engineering Asset Management Organizations-Current Picture in Australia. In ICIQ (pp. 467-484).
- [64] Panian, Ž. (2003). Klepac G. Poslovna inteligencija, Masmedia.
- [65] Atzeni, P., Jensen, C. S., Orsi, G., Ram, S., Tanca, L., & Torlone, R. (2013). The relational model is dead, SQL is dead, and I don't feel so good myself. ACM SIGMOD Record, 42(2), 64-68. doi: 10.1145/2503792.2503808
- [66] Milicev, D. (2009). Model-driven development with executable UML. John Wiley & Sons.
- [67] Lazarević, B., Marjanović, Z., Aničić, N., & Babarogić, S. (2006). Baze podataka (treće izdanje).

- [68] Bečejski-Vujaklija, D. (2009). Uvod u informacione sisteme. Fakultet organizacionih nauka.
- [69] Moss, L. T., & Atre, S. (2003). Business intelligence roadmap: the complete project lifecycle for decision-support applications. Addison-Wesley Professional.
- [70] Kalakota, R., Robinson, M., & Tapscott, D. (2001). *E-business 2.0: Roadmap for Success* (Vol. 261). Boston: Addison-Wesley.
- [71] Wyatt, I. (2009). *The small-cap investor: secrets to winning big with small-cap stocks*. John Wiley & Sons.
- [72] Badami, V. (2003). Payback on Business Intelligence. *DM Review*.
- [73] Luhn, H. P. (1958). A business intelligence system. *IBM Journal of research and development*, 2(4), 314-319.
- [74] Building Data-Driven Decision Support Systems, <http://www.dssresources.com>, pristupljeno 18.07.2025. godine
- [75] Rud, O. P. (2009). Business intelligence success factors: tools for aligning your business in the global economy (Vol. 18). John Wiley & Sons.
- [76] Dresner, H. (2009). *Profiles in performance: Business intelligence journeys and the roadmap for change*. John Wiley & Sons.
- [77] Vidojković, D., Savić, J., & Bečejski-Vujaklija, D. (2002). Business intelligence portal: Police college decision support system. *Info M*, 1(3-4), 17-19.
- [78] Димитријевић, Н. (2014). Прилог развоју Business Intelligence система на примеру управљања успехом студената академије (Doctoral dissertation, Универзитет одбране, Војна академија).
- [79] Arias, R., Jave, Y., Saldaña, R., & Garces, E. (2024, April). Methodology for a Business Intelligence Platform by Using Oracle 19C Database Engine and its Limitations. In *Computer Science On-line Conference* (pp. 67-87). Cham: Springer Nature Switzerland.
- [80] Nichols, D. M., Chan, C. H., Bainbridge, D., McKay, D., & Twidale, M. B. (2008, June). A lightweight metadata quality tool. In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries* (pp. 385-388).
- [81] Fotache, M., & Strimbei, C. (2015). SQL and data analysis. Some implications for data analysis and higher education. *Procedia Economics and Finance*, 20, 243-251.
- [82] Implementation of data warehouse in analysis of QMS inconsistencies. *Metal. Int*, 17(3), 5-10.
- [83] Aravind Kumar Balasubramaniam, „Methodological Standpoint: Inmon Vs Kimball“, 2015. Dostupno na: <https://www.astera.com/type/blog/data-warehouse-concepts>, pristupljeno 18.07.2025. godine.
- [84] Kimball R., *The Data Warehouse Toolkit – Practical Techniques for Building Dimensional Data Warehouses*, John Wiley & Sons, 1996.
- [85] Kimball, R., & Ross, M. (2016). *Dimension Table Core Concepts*. The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence, 327-384.
- [86] Steven Finlay, *Predictive Analytics, Data Mining and Big Data-Myths, Misconceptions and Methods*, Palgrave Macmillan, London, England, UK, 2014.
- [87] Gojgic, N., Veljovic, A., Zahorjanski, M., Stanojevic, L., & Cvijovic, N. (2012). Implementation of data warehouse in analysis of QMS inconsistencies. *Metal. Int*, 17(3), 5-10.

- [88] Golfarelli, M., & Rizzi, S. (2009). *Data warehouse design: Modern principles and methodologies*. McGraw-Hill, Inc.
- [89] The future of data warehousing: 7 industry experts share their predictions“, <https://www.betterbuys.com/bi/future-of-data-warehousing/>, pristupljeno 18.07.2025. godine
- [90] Golfarelli, M., Maio, D., & Rizzi, S. (1998, January). Conceptual design of data warehouses from E/R schemes. In *Proceedings of the thirty-first Hawaii international conference on system sciences* (Vol. 7, pp. 334-343). IEEE.
- [91] The future of data warehousing: 8 industry experts share their predictions“, <https://www.betterbuys.com/bi/future-of-data-warehousing/>, pristupljeno 18.07.2025. godine
- [92] Data warehousing articles, <http://www.dbta.com/Categories/Data-Warehousing-333.aspx>, pristupljeno 18.07.2025. godine
- [93] Tarko, A. P., & Kanodia, M. (2004). *Hazard Elimination Program-Manual on Improving Safety of Indiana Road Intersections and Sections; Volume 1: Research Report and Volume 2: Guidelines for Highway Safety Improvements in Indiana*.
- [94] Sendín-Raña, P., González-Castaño, F. J., Pérez-Barros, E., Rodríguez-Hernández, P. S., Gil- Castiñeira, F., & Pousada-Carballo, J. M. (2009). Improving the performance and functionality of Mondrian open-source OLAP systems. *Software: Practice and Experience*, 39(3), 279-298.
- [95] Skaug, L., Nojournian, M., Dang, N., & Yap, A. (2025). Road crash analysis and modeling: A systematic review of methods, data, and emerging technologies. *Applied Sciences*, 15(13), 7115.
- [96] Butt, M. S., & Shafique, M. A. (2025). A literature review: AI models for road safety for prediction of crash frequency and severity. *Discover Civil Engineering*, 2(1), 99.
- [97] Lacherre, J., Castillo-Sequera, J. L., & Mauricio, D. (2024). Factors, prediction, and explainability of vehicle accident risk due to driving behavior through machine learning: A systematic literature review, 2013–2023. *Computation*, 12(7), 131.
- [98] Ma, Z., Mei, G., & Cuomo, S. (2021). An analytic framework using deep learning for prediction of traffic accident injury severity based on contributing factors. *Accident Analysis & Prevention*, 160, 106322.
- [99] Behboudi, N., Moosavi, S., & Ramnath, R. (2024). Recent advances in traffic accident analysis and prediction: A comprehensive review of machine learning techniques. *arXiv preprint arXiv:2406.13968*.
- [100] Hoang, D. T., Achache, M., & Jain, V. K. (2024). Deep spectral network for time series clustering. *EAI Endorsed Transactions on Scalable Information Systems*, 11(3).
- [101] Hu, H., & Lee, M. J. (2022, February). Graph neural network-based clustering enhancement in VANET for cooperative driving. In *2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)* (pp. 162-167). IEEE.
- [102] Wani, A. A. (2024). Comprehensive analysis of clustering algorithms: exploring limitations and innovative solutions. *PeerJ Computer Science*, 10, e2286.
- [103] Yin, H., Aryani, A., Petrie, S., Nambissan, A., Astudillo, A., & Cao, S. (2024). A rapid review of clustering algorithms. *arXiv preprint arXiv:2401.07389*.
- [104] Zakur, Y., & Flaih, L. (2023). Apriori algorithm and hybrid Apriori algorithm

- in the data mining: a comprehensive review. In *E3S Web of Conferences* (Vol. 448, p. 02021). EDP Sciences.
- [105] Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996, August). A density-based algorithm for discovering clusters in large spatial databases with noise. *Inkdd* (Vol. 96, No. 34, pp. 226-231).
- [106] Jiang, H., Jang, J., & Lacki, J. (2020). Faster DBSCAN via subsampled similarity queries. *Advances in Neural Information Processing Systems*, 33, 22407-22419.
- [107] Deng, D. (2020, September). DBSCAN clustering algorithm based on density. In *2020 7th international forum on electrical engineering and automation (IFEAA)* (pp. 949-953). IEEE.
- [108] Gholizadeh, N., Saadatfar, H., & Hanafi, N. (2021). K-DBSCAN: An improved DBSCAN algorithm for big data. *The Journal of supercomputing*, 77(6), 6214-6235.
- [109] Zhang, Y., Wang, X., & Li, J. (2022). *DBSCAN++: Efficient density-based clustering for streaming data*. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 1234–1243). ACM. <https://doi.org/10.1145/3534678.3539123>
- [110] Tareq, M., Sundararajan, E. A., Harwood, A., & Bakar, A. A. (2021). A systematic review of density grid-based clustering for data streams. *Ieee Access*, 10, 579-596.
- [111] Çeven, S., & Albayrak, A. (2024). Traffic accident severity prediction with ensemble learning methods. *Computers and Electrical Engineering*, 114, 109101.
- [112] Chen, J., & Zhang, X. (2024). Accident Black Spots Identification of Electric Scooter Based on Self-Adaptive DBSCAN Algorithm. *Academic Journal of Computing & Information Science*, 7(10), 119-124.
- [113] Elyassami, S., Hamid, Y., & Habuza, T. (2021, June). Road crashes analysis and prediction using gradient boosted and random forest trees. In *2020 6th IEEE Congress on Information Science and Technology (CiSt)* (pp. 520-525). IEEE.
- [114] Gu, Y., Liu, Y., Liu, D., Han, L. D., & Jia, X. (2024). Spatiotemporal kernel density clustering for wide area near Real-Time pothole detection. *Advanced Engineering Informatics*, 60, 102351.
- [115] Khaled, A., Elsir, A. M. T., Wang, P., Shen, Y., & Zhang, Q. (2024). A graph-based approach for traffic prediction using similarity and causal relations between nodes. *Knowledge-Based Systems*, 296, 111913.
- [116] Kushwaha, A. K., Kumar, H., Arora, H., & Kumar, B. A. (2025). A deep embedded clustering method for location-specific driving safety profiling using trajectory data. *Scientific Reports*, 15(1), 41635.
- [117] Rajabpour, A., & Yarahmadi, F. (2024). Remote sensing, geographic information system (GIS), and machine learning in the pest status monitoring. In *Decision system in agricultural pest management* (pp. 247-353). Singapore: Springer Nature Singapore.
- [118] Thean, L. H., & Chor, W. T. (2025, December). Advancing Cyclist Safety: A Robust Collision Avoidance System with DBSCAN-Enhanced Clustering and Bayesian-Optimized Adaptive Extended Kalman Filter. In *2025 IEEE 13th Conference on Systems, Process & Control (ICSPC)* (pp. 13-18). IEEE
- [119] Zhang, X., Chen, H., Chen, J., Feng, H., Liu, M., Zhang, X., ... & Chen, F. (2025). A hybrid machine learning-enhanced MCDM model for transport

- safety engineering. *Scientific Reports*, 15(1), 36467.
- [120] Wang, Y., Zhao, Q., & Xiu, W. (2024, July). Blackspots identification of abnormal driving behavior based on improved DBSCAN algorithm. In Third International Conference on Electronic Information Engineering, Big Data, and Computer Technology (EIBDCT 2024) (Vol. 13181, pp. 292-299). SPIE.
- [121] Delibašić, B., Suknović, M., & Jovanović, M. (2009). Algoritmi mašinskog učenja za otkrivanje zakonitosti u podacima. Fakultet organizacionih nauka.
- [122] Szénási, S., & Csiba, P. (2014). Clustering algorithm in order to find accident black spots identified by GPS coordinates. Proceedings of the 14th GeoConference on Informatics, Geoinformatics, and Remote Sensing, Ilza, Poland, 19-25.
- [123] Dewson, R. (2008). Beginning SQL Server 2008 Express for Developers: From Novice to Professional. Apress.
- [124] Zhang, J., Wan, C., He, A., Zhang, D., & Soares, C. G. (2021). A two-stage black-spot identification model for inland waterway transportation. *Reliability Engineering & System Safety*, 213, 107677. DOI: <https://doi.org/10.1016/j.res.2021.107677>
- [125] Zhang, M., Ma, Y., Li, J., & Zhang, J. (2023). A density connection weight-based clustering approach for dataset with density-sparse region. *Expert Systems with Applications*, 230, 120633.
- [126] Susilo, B. H. (2016). Guideline for survey, investigation, and design of black spot location (SID-BSL) and its application in Lampung Province, Indonesia. *Civil Engineering Dimension*, 18(1), 49-56. DOI: <https://doi.org/10.9744/ced.18.1.49-56>
- [127] Jordan ATANASIJEVIC, Danijela MILOSEVIC, Upgrading the Business Intelligence System by Implementing the Decision Tree Model in the R Software Package, *Studies in Informatics and Control*, ISSN 1220-1766, vol. 29(2), pp. 243-254, 2020. DOI: <https://doi.org/10.24846/v29i2y202009>
- [128] J. Atanasijević, Analiza podataka o saobraćajnim nezgodama korišćenjem sistema poslovne inteligencije, Zbornik sa XXIX naučne i biznis konferencije Yu Info 2023, 12-15. mart 2023, Kopaonik, ISBN 978-86-85525-29-2
- [129] Портал отворених података, <https://data.gov.rs/sr/datasets/podatsi-o-saobratshajnim-nezgodama-po-politsijskim-upravama-i-opshtinama/>, приступљено 30.06.2025. године
- [130] SolidIT (2018). DB-Engines Ranking. Доступно на: <https://db-engines.com/en/ranking>, приступљено 21.07.2025.
- [131] Ballard, C., Farrell, D. M., Gupta, A., Mazuela, C., & Vohnik, S. (2012). Dimensional Modeling: In a Business Intelligence Environment. IBM Redbooks.
- [132] Lazarević, B., Marjanović, Z., Aničić, N., & Babarogić, S. (2006). Baze podataka (šesto izdanje).
- [133] Bojčić, I. M. (2017). Modelom vođen razvoj skladišta podataka zasnovanog na Data Vault pristupu (Doctoral dissertation, Univerzitet u Beogradu, Fakultet organizacionih nauka).
- [134] Fotache, M. (2016). Data processing languages for business intelligence. SQL vs. R. *Informatica Economica*, 20(1), 48.
- [135] Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996, August). A density-based algorithm for discovering clusters in large spatial databases with noise. *Inkdd* (Vol. 96, No. 34, pp. 226-231).

Прилози

Списак слика

| | |
|---|----|
| Слика 1. Идентификација „црних тачака“ коришћењем приступа „клизећег прозора“ [26] | 29 |
| Слика 2. Приказ екрана програма Microsoft Excel са детаљима опасног места [101] .. | 33 |
| Слика 3. Алгоритам процеса управљања опасним местима у Београду | 36 |
| Слика 4. Веза између контекста и разумевања знања | 46 |
| Слика 5. Трослојна архитектура складишта података [35] | 49 |
| Слика 6. Пример просторних скупова података..... | 59 |
| Слика 7. Пример просторних скупова података..... | 59 |
| Слика 8. Тачка p је директно доступна кроз густину од q , кружница представља Eps околину од q и $MinPts$ је 4, што показује да је p централна тачка у кластеру | 60 |
| Слика 9. Тачка p је повезана кроз густину од тачке q | 61 |
| Слика 10. Граничне тачке p и q су повезане кроз густину од централне тачке o | 61 |
| Слика 11. Сортиран 4-dist граф и преломна тачка | 64 |
| Слика 12. Сортирани 4-dist график и преломна тачка..... | 68 |
| Слика 13. Део података саобраћајних незгода у општини Ариље за период од 2021. до 2024. године | 71 |
| Слика 14. Преглед израчунатих података..... | 73 |
| Слика 15. Сортирани 4-dist график и преломна тачка..... | 77 |
| Слика 16. Облик кластера који користи модификовани алгоритам | 78 |
| Слика 17. Преглед примера скупа тачака на које се примењује алгоритам | 79 |
| Слика 18. Преглед растојања за једну тачку из примера скупа тачака на које се примењује алгоритам | 80 |
| Слика 19. Преглед кластера за сваку од тачки из примера скупа тачака на које се примењује алгоритам | 81 |
| Слика 20. Преглед тачки обухваћених једним од кластера из примера скупа тачака на које се примењује алгоритам..... | 81 |
| Слика 21. Приказ сценарија када се два кластера преклапају, а растојање њихових центроида је мање од полупречника кластера | 82 |
| Слика 22. Приказ сценарија када се два кластера преклапају, а растојање њихових центроида је веће од полупречника кластера..... | 83 |
| Слика 23. Приказ сценарија када се два кластера не преклапају | 84 |
| Слика 24. Модификовани алгоритам..... | 86 |
| Слика 25. Визуелна презентација изгледа једног чвора | 87 |
| Слика 26. Скуп података са којим алгоритам започиње рад | 88 |
| Слика 27. Центроиди издвојених кластера који задовољавају иницијалне критеријуме | 88 |
| Слика 28. Резултати добијени модификованим алгоритмом након сажимања издвојених кластера..... | 89 |
| Слика 29. Први издвојени независни кластери | 90 |
| Слика 30. Кластери добијени у следећој итерацији са проширеним полупречником | 90 |
| Слика 31. Кластери добијени у следећој итерацији са проширеним полупречником | 91 |
| Слика 32. Коначни центроиди добијени у последњем кораку рада модификованог алгоритма..... | 91 |
| Слика 33. Преглед добијених кластера радом модификованог алгоритма | 92 |
| Слика 34. Изглед веб стране Портала отворених Података Републике Србије | 95 |
| Слика 35. Изглед табеле преузете са портала отворених података [98] | 96 |
| Слика 36. Изглед форме при покретању апликације..... | 98 |
| Слика 37. Учитавање података из изворних фајлова у складиште података | 99 |

| | |
|---|-----|
| Слика 38. Изглед форме након успешног учитавања података за анализу | 99 |
| Слика 39. Изглед дела табеле са подацима о саобраћајним незгодама на територији Републике Србије, након уписивања преузетих података са Портала отворених података Републике Србије..... | 100 |
| Слика 40. Изглед дела табеле са подацима о саобраћајним незгодама на територији Републике Србије, након трансформације колоне са временским описом.. | 101 |
| Слика 41. Приказ врста саобраћајних незгода | 101 |
| Слика 42. Изглед дела табеле са подацима о локацији саобраћаје незгоде на територији Републике Србије..... | 102 |
| Слика 43. Брисање везе између fact табеле и табеле у којој се налазе подаци за димензију времена..... | 103 |
| Слика 44. Конверовање података и генерисање нових вредности у табели DimDate . | 104 |
| Слика 45. Уписивање података у жељено одредиште | 105 |
| Слика 46. Брисање релације између fact табеле и димензије при уписивању података у димензију..... | 105 |
| Слика 47. Дефинисање одредишта и успостављање конекције са складиштем података..... | 106 |
| Слика 48. Приказ упита за брисање релације, брисање и креирање димензије..... | 106 |
| Слика 49. Приказ процеса учитавања, трансформације и уписивања података у димензију Врста саобраћајне незгоде у складишту податаку..... | 107 |
| Слика 50. Приказ процеса учитавања, трансформације и уписивања података у димензије и централну табелу при генерисању складишта података..... | 108 |
| Слика 51. Изглед форме на којој корисник уноси податке за идентификацију “црних тачака”..... | 108 |
| Слика 52. Изглед форме на којој корисник бира град за који жели идентификацију “црних тачака”..... | 109 |
| Слика 53. Изглед форме на којој корисник бира општину за коју жели идентификацију “црних тачака”..... | 109 |
| Слика 54. Изглед форме на којој корисник уноси почетну вредност полупречника кластера | 110 |
| Слика 55. Изглед форме на којој корисник уноси доњу границу вредности кластера . | 111 |
| Слика 56. Изглед форме на којој корисник бира временски оквир за идентификацију “црних тачака” на изабраном подручју | 111 |
| Слика 57. Изглед форме са идентификованим “црним тачкама” на изабраном подручју | 112 |
| Слика 58. Изглед форме са идентификованим “црним тачкама” на изабраном подручју са опцијом детаљнијег прегледа..... | 112 |
| Слика 59. Изглед форме са идентификованим “црним тачкама” на изабраном подручју са прегледом свих саобраћајних незгода у кластеру | 113 |
| Слика 60. Изглед форме са идентификованом “црном тачком” са највећим бројем саобраћајних незгода у кластеру полупречника 479 метара, улица Цетињска | 113 |
| Слика 61. Изглед форме са идентификованом “црном тачком” са највећим бројем саобраћајних незгода у кластеру полупречника 479 метара, са приказом локација свих саобраћајних незгода | 114 |
| Слика 62. Изглед форме са идентификованом “црном тачком” на кружном току Славија са приказом свих саобраћајних незгода у кластеру полупречника 479 метара | 114 |
| Слика 63. Изглед форме са идентификованом “црном тачком” на кружном току Славија са детаљним приказом свих саобраћајних незгода у кластеру полупречника 479 метара | 115 |
| Слика 64. Модел података складишта података | 117 |

Слика 65. Преглед идентификованих опасних места на путевима у Републици Србији
за
изабрани трогодишњи период 122

Списак табела

| | |
|--|-----|
| Табела 1. Преглед дефиниција опасних места на путевима у одабраним европским државама [27] | 25 |
| Табела 2. Критичне вредности за идентификацију „црних тачака“ у Немачкој [26] .. | 27 |
| Табела 3. Граничне вредности за одређивање „црних тачака“ [26]..... | 32 |
| Табела 4. Преглед реализованих студија, пројеката и истраживања у којима је обрађивана тема идентификације опасних места – „црних тачака“ на путевима Републике Србије [104] | 33 |
| Табела 5. Скуп података са државама и очекиваним животним веком [22] | 65 |
| Табела 6. Нормализовани подаци у опсегу вредности од 0 до 1 | 66 |
| Табела 7. Сортиране вредности четвртог најмањег одстојања за сваку државу | 67 |
| Табела 8. Сортиране вредности четвртог најмањег одстојања за сваку тачку | 73 |
| Табела 9. Називи о објашњења променљивих које се користе из Скупа отворених података..... | 96 |
| Табела 10. Списак идентификованих опасних места на путевима у Републици Србији за изабрани трогодишњи временски интервал | 119 |
| Табела 11. Списак идентификованих опасних места на путевима у Републици Србији за изабрани трогодишњи временски интервал са обележеним локацијама које не би биле идентификоване са тренутно примењеном методологијом | 130 |