

ANALIZA FINANSIJSKIH POKAZATELJA I PREDVIĐANJE KREDITNIH USLOVA STAMBENIH KREDITA PRIMENOM METODA I TEHNIKA DATA MINING-A

Nemanja Jovanović¹Nenad Stefanović²

Rezime: U ovom radu prikupljeni su podaci koji se odnose na različite finansijske pokazatelje koji utiču na uslove pod kojima se nude stambeni krediti u evrima. Razvijen je jedan Business Intelligence sistem počev od integracije i transformacije podataka, zatim analize primenom tehnika Data Mining-a pa sve do vizuelizacije, publikovanja i deljenja informacija. Kao rezultat samog istraživanja primenom metoda linearne regresije, ustanovljeno je da se povećanjem finansijskih pokazatelja kao što su stopa Euribora i inflacije zatim smanjenjem realnog indeksa zarada, otvara mogućnost za povećanjem kamatnih stopa banaka, koje imaju nepovoljan uticaj. Krajnji rezultat istraživanja upotpunjen je primenom metode višestruke linearne regresije koja finansijskim institucijama pruža mogućnost dugoročne prognoze visine kamatnih stopa stambenih kredita na tržištu u zavisnosti od stope inflacije i šestomesečnog Euribora. Na osnovu razvijenog Business Intelligence sistema banke i druge finansijske institucije mogu se bolje prilagođavati tržištu, održavati konkurentnost i stabilnost.

Ključne reči: Business Intelligence, podaci, kamate, finansije, banke, regresija, krediti

ANALYSIS OF FINANCIAL INDICATORS AND PREDICTION OF CREDIT TERMS OF HOUSING LOANS USING DATA MINING METHODS AND TECHNIQUES

Abstract: In this paper, data related to various financial indicators that influence the conditions under which housing loans in euros are offered have been collected. A Business Intelligence system was developed, starting with the integration and transformation of data, then analysis using Data Mining techniques, all the way to visualization, publication and sharing of new information. As a result of the research itself, using the linear regression method, it was established that by increasing financial indicators such as the Euribor rate and inflation, then by decreasing the real wage index, the possibility of increasing bank interest rates, which have an unfavorable impact, opens up. The final result of the research was completed by the application of the multiple linear regression method, which provides financial institutions with the possibility of long-term forecasting of interest rates for housing loans on the market, depending on the inflation rate and the six-month Euribor. Based on the developed Business Intelligence system, banks and other financial institutions can better adapt to the market, maintain competitiveness and stability.

Key words: Business Intelligence, data, interest, finance, banks, regression, loans

1. UVODNA RAZMATRANJA

Razvojem računarske tehnologije u savremenom poslovnom svetu, stvorila se potreba za raznovrsnim načinima istraživanja podataka na internetu. Jedan od načina istraživanja podataka jeste primena Data Mining tehnika putem kojih se vrši izdvajanje korisnih informacija, što predstavlja jedan vid takozvanog rudarenja, odnosno istraživanja podataka. Postoji više vrsta Data Mining tehnika, a jedna koja je odabrana u svrhu istraživanja jeste tehnika regresije, koja u sebi sadrži više različitih metoda za predikciju pojava i događaja od izuzetne važnosti poslovnom svetu, ali i široj javnosti. Uloga Data Mining-a predstavlja sastavni deo poslovanja svake banke i druge finansijske institucije u analizi i predviđanju različitih finansijskih pokazatelja. Za banku, kao finansijsku instituciju, najbitnije je da svojim uslugama privuče veći broj klijenata sa kojima će ostvariti uspešnu saradnju. Jedna od različitih usluga u setu bankarskih ponuda jeste i stambeni kredit. Analizom finansijskih pokazatelja u domenu stambenih kredita stiže se sveobuhvatna slika trenutnog stanja na tržištu koja banci ali i drugim finansijskim institucijama omogućava održavanje konkurentnosti kao i prilagođavanje

¹Student doktorskih studija, Fakultet tehničkih nauka, Čačak, Svetog Save 65, e-mail: 515-2019@ftn.edu.rs

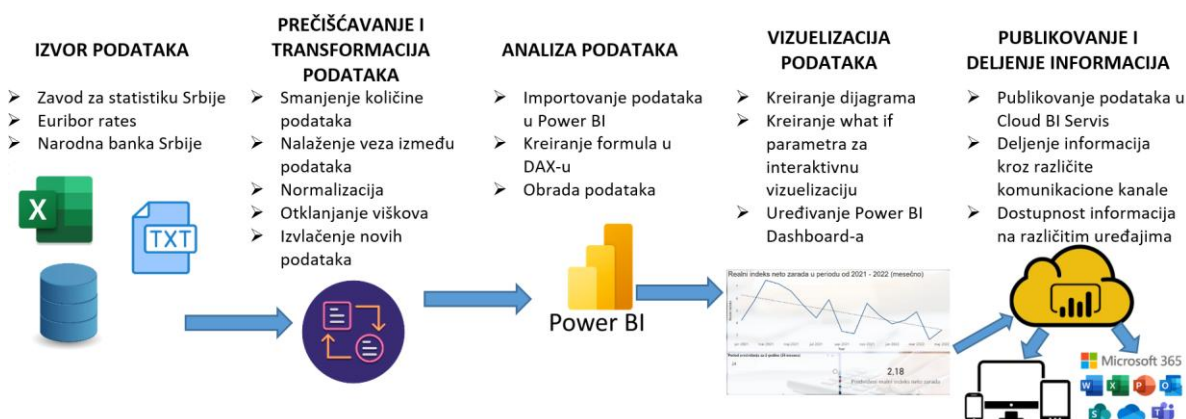
²Redovni profesor, Fakultet tehničkih nauka, Čačak, Svetog Save 65, e-mail: nenad@ftn.edu.rs

novonastalim uslovima poslovanja. Poslednjih godina, počev od kraja 2019. godine, svesni smo različitih događanja počev od pandemije korona virusa, pa do različitih političkih i ekonomskih sankcija koja su uslovia promenu tržišnih uslova na međunarodnom nivou. Sve ove promene itekako imaju uticaja na finansijski sektor u kome je od ključnog značaja da se na svaku navedenu promenu pravovremeno deluje. Sproveden je veliki broj istraživanja na datu temu, dok su za potrebe rada izdvojena dva koja su najpribližnija. Autori prvog istraživanja su došli do zaključka da je indeks potrošačkih cena (indeks inflacije) varijabla sa najvećim značajem koji dalje objašnjava povećanje cena stanova [1]. Drugo istraživanje navodi autore na rezultat jake veze između negativnih kamatnih stopa i indeksa cena nekretnina [2]. Upravo prethodno izlaganje predstavlja cilj istraživanja, i prikupljeni podaci biće analizirani primenom metoda regresije, kao jedne od Data Mining tehnika, uz pomoć softverskog alata Power BI, kako bi se pružile najkvalitetnije informacije, odnosno prognoze za dalji razvoj finansijskih događaja na tržištu po pitanju stambenih kredita. Drugim rečima, naglasiće se značaj svake regresione metode u analizi podataka i uporediće se njihov pojedinačni uticaj na dobijene rezultate samog istraživanja sa drugim srodnim istraživanjima.

U prethodnom tekstu predstavljena je tematika samog istraživanja dok će se u narednom delu rada prikazati sama metodologija istraživanja kao i ostvareni rezultati i diskusija istih, što je upravo i cilj samog istraživačkog rada.

2. METODOLOGIJA ISTRAŽIVANJA

Početak metodologije istraživanja zasniva se na izgradnji Business Intelligence sistema sa definisanim aktivnostima, koji je prikazan na slici 1.



Slika 1 – Arhitektura Business Intelligence sistema

Nakon prikazane arhitekture Business Intelligence sistema, u narednim potpoglavljima sledi objašnjenje svake pojedinačne aktivnosti.

2.1. Izvor podataka

Podaci prikupljeni za analizu preuzeti su .xlsx i tekstualnom formatu sa različitih izvora, u zavisnosti od vrste podataka korišćeni su sledeći izvori:

1. Republički zavod za statistiku Srbije; Podaci prikupljeni sa sajta Republičkog zavoda za statistiku Srbije u .xlsx formatu odnose se na: prosečne cene kvadrata novogradnje, indeksa neto zarada i inflacije u Republici Srbiji [3].
2. Euribor rates (engl. Euro Interbank Offered Rate); Podaci prikupljeni sa sajta Euribor rates odnose se na prikaz kamatnih stopa šestomesečnog Euribora na mesečnom nivou u periodu od januara 2021. do Juna 2022. godine. Podaci su prikupljeni u tekstualnom formatu i isti predstavljani u .xlsx formatu zbog bolje obrade alatu Power BI za analizu podataka [4].

3. Narodna banka Srbije; Podaci prikupljeni sa sajta Narodne banke Srbije odnose se na prikaz prosečnih kamatnih stopa izdatih na stambene kredite u tekstualnom formatu i isti su formatirani u .xlsx formatu [5].

2.2. Prečišćavanje i transformacija podataka

Nakon uspešno prikupljenih podataka pristupa se sledećem koraku koji se odnosi na prečišćavanje i transformaciju podataka. Prečišćavanje podataka predstavlja jedan od najvažnijih zadataka Data Mining-a i uključuje pripremu i transformaciju podataka u odgovarajući šablon, koji je pogodan istraživačkim metodama [6]. Prvobitni skup podataka je u sebi sadržao veliki broj nepotrebnih kolona, zatim kolone nisu bile u jedinstvenom vremenskom redosledu što je dalje dovodilo do gubljenja hronologije samog analitičkog šablona. Takođe svaka ćelija u .xlsx formatu morala je da se pravilno formatira na način kojim bi se obezbedila tačnost pri analizi podataka. Nakon uspešnog prečišćavanja i transformacije podataka sledi proces pripreme za analizu istih.

2.3. Analiza podataka

Prečišćeni i transformisani skup podataka uvozi se u softver za analizu. Power BI (Desktop) je interaktivni softverski proizvod za vizuelizaciju podataka koji je razvio Microsoft sa primarnim fokusom na poslovnu inteligenciju [7]. Sama aplikacija je deo Microsoft Power Platforme i predstavlja kolekciju softverskih usluga, aplikacija i konektora koji zajedno rade na pretvaranju nepovezanih izvora podataka u koherentne, vizuelno impresivne i interaktivne izveštaje.

Kreiranjem funkcija u DAX-u (engl. Data Analysis Expressions) koji je podržan od strane alata Power BI, formiraju se modeli koji su prilagođeni analizi podataka putem Data Mining tehnike. Data Mining predstavlja proces izdvajanja validnih i nepoznatih informacija iz velikih baza podataka, koje bi se mogle koristiti za donošenje strateških odluka u poslovanju [8]. Tehnike Data Mining-a mogu se podeliti na nadgledani (engl. Supervised) i nenadgledani (engl. Unsupervised) tip učenja. Kao što je već pomenuto, u ovom istraživanju koristiće se regresiona tehnika koja spada u nadgledan tip učenja. Regresiona analiza predstavlja skup statističkih tehnika koje služe kao osnova za donošenje zaključaka o odnosima između međusobno povezanih varijabli. Pošto su ove tehnike primenljive u skoro svim oblastima studija regresiona analiza je za sada najkorišćenija od svih metoda analize podataka [9]. Tehnika regresije sadrži više metoda, dok je za potrebe istraživanja odabrana jednostavna i višestruka linearna regresija. Najjednostavniji oblik regresije jeste jednostavna linearna regresija koja sadrži jednu zavisnu i jednu nezavisnu promenljivu, na osnovu kojih se može kreirati model linearne regresije. Dati model se može posmatrati kao linija koja umanjuje stopu greške između stvarne predviđene vrednosti i tačaka na liniji [10]. Slučaj sa jednom nezavisnom promenljivom naziva se jednostavna linearna regresija. Kada je uključeno više od jedne nezavisne promenljive, proces se naziva višestruka (kompleksna) linearna regresija [11].

Nakon izgradnje modela za analizu pomoću tehnika Data Mining-a sledi završna aktivnost vizuelizacije i publikovanja podataka.

2.4. Vizuelizacija podataka i publikovanje informacija

Najvažniji finalni deo aktivnosti jeste da se krajnjim korisnicima na najbolji način vizuelno predstave obrađeni podaci. Power BI nudi izuzetna rešenja za vizuelizaciju kao i niz drugih mogućnosti da se podaci predstave na najbolji način što se može videti u narednom poglavlju gde su prikazani rezultati analize. Takođe za samo publikovanje informacija moguće je iste nakon analize postaviti na Cloud i time omogućiti njihovo deljenje na različite uređaje (mobilni telefon, tablet, računar) ili sredstva komunikacije (Microsoft Office 365, Teams i slično).

Detaljan uvid u samu vizuelizaciju podataka sledi u narednom poglavlju koje se odnosi na ostvarene rezultate i diskusiju istih.

3. REZULTATI I DISKUSIJA

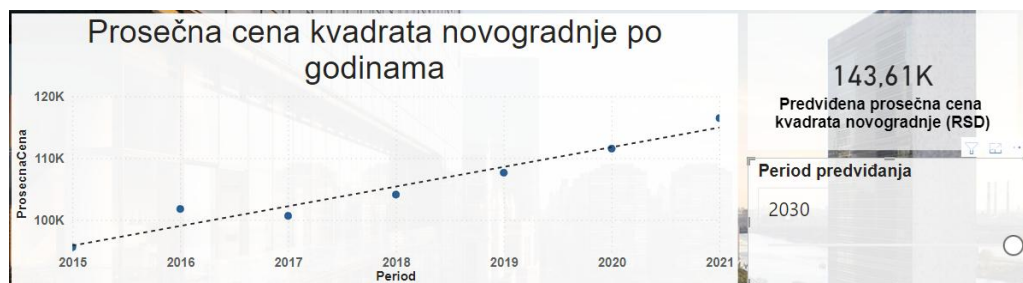
U narednim potpoglavljima biće prikazani ostvareni rezultati kao i diskusija istih primenom navedenih metoda jednostavne i višestruke linearne regresije.

3.1. Rezultati i diskusija primenom metode jednostavne linearne regresije

Za analizu rezultata kretanja **prosečne cene kvadrata novogradnje** u Republici Srbiji odabrani su sledeći parametri:

- Za x (nezavisna promenljiva) – meseci u godini
- Za y (zavisna promenljiva) – prosečna cena kvadrata novogradnje

Na osnovu podataka za period od 2015. do 2021. godine, odnosno definisanih kolona za x i y, primenom linearne regresije, kreiran je grafički prikaz koji prikazuje tok kretanja prosečne cene kvadrata novogradnje na godišnjem nivou, što se može videti na slici 2.



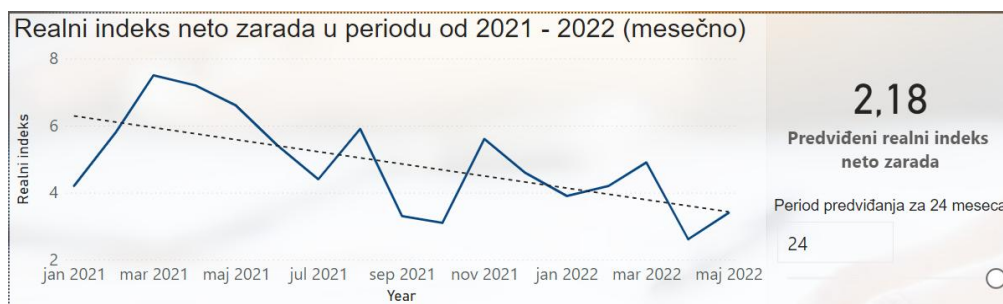
Slika 2 - Podaci o predviđenoj prosečnoj ceni kvadrata novogradnje za 2023. i 2030. godinu

Sa grafika se može zaključiti da će prosečna cena po kvadratu novogradnje rasti iz godine u godinu i da se pad cena nekretnina ne očekuje u skorijem periodu. Ovaj pokazatelj sam po sebi govori da je tržište nekretnina nestabilno i da se pravo na ostvarenje uslova za stambeni kredit dovodi u pitanje jer se povećava verovatnoća da zarade stanovništva ne mogu opravdati povećanja cena.

Za analizu rezultata kretanja **realnog indeksa neto zarada** u Republici Srbiji odabrani su sledeći parametri:

- Za x (nezavisna promenljiva) – meseci u godini
- Za y (zavisna promenljiva) – realni indeks neto zarada

Na osnovu podataka za period od godinu dana (2021 – 2022), odnosno definisanih kolona za x i y, primenom linearne regresije, kreiran je grafički prikaz koji prikazuje tok kretanja realnog indeksa neto zarada na mesečnom nivou, što se može videti na slici broj 3.



Slika 3 - Podaci o predviđenom realnom indeksu neto zarada do kraja 2022. godine

Sa grafika se jasno može videti da linearna regresija prognozira smanjenje realnog indeksa neto zarada i da će do kraja 2022. godine realni indeks neto zarada pasti na 2,18%. Poskupljenjem cena

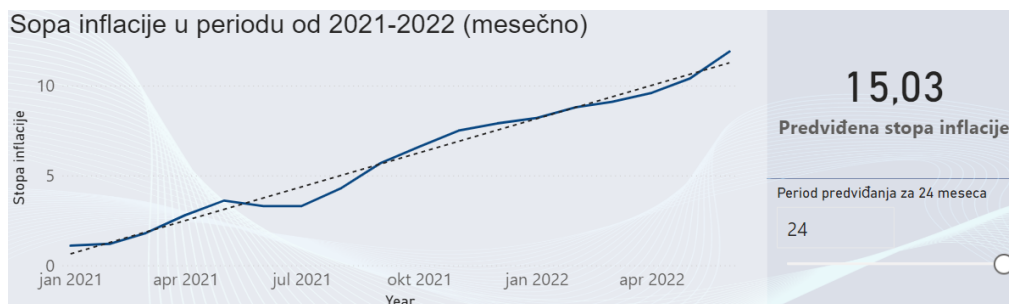
18. KONFERENCIJA SA MEĐUNARODNIM UČEŠĆEM RIZIK I BEZBEDNOSNI INŽENJERING

nekretnina i manjim realnim indeksom zarade, zaposleni (klijenti) imaju manje mogućnosti da ispune sve uslove po pitanju stambenih kredita.

Za analizu rezultata kretanja **stope inflacije** u Republici Srbiji odabrani su sledeći parametri:

- Za x (nezavisna promenljiva) – meseci u godini
- Za y (zavisna promenljiva) – stopa inflacije

Na osnovu podataka za period od godinu dana (2021 – 2022), odnosno definisanih kolona za x i y, primenom linearne regresije, kreiran je grafički prikaz koji prikazuje tok kretanja stope inflacije na mesečnom nivou, što se može videti na slici broj 4.



Slika 4 – Podaci o predviđenoj stopi inflacije do kraja 2022. godine

Sa grafika se jasno može videti da linearna regresija prognozira povećanje stope inflacije, gde će ista do kraja 2022. godine dostići čak dvocifreni nivo, odnosno 15,03%. Smanjenjem realnog indeksa neto zarada i povećanjem stope inflacije kao i prosečne cene kvadrata novogradnje smanjuje se životni standard i kupovna moć, usled čega se i smanjuje mogućnost za plasiranjem i odobravanjem stambenih bankarskih kredita.

Euribor (engl. Euro Interbank Offered Rate) predstavlja različite stope zasnovane na prosečnim kamatnim stopama po kojima veliki broj evropskih banaka pozajmljuje sredstva jedna od druge. Za analizu rezultata uzeta je stopa šestomesečnog Euribora, kao najčešće ponuđena mogućnost prilikom uzimanja stambenih kredita od banaka, gde su odabrani su sledeći parametri:

- Za x (nezavisna promenljiva) – meseci u godini
- Za y (zavisna promenljiva) – stopa šestomesečnog Euribora

Na osnovu podataka za period od godinu dana (2021. – 2022.), odnosno definisanih kolona za x i y, primenom linearne regresije, kreiran je grafički prikaz koji prikazuje tok kretanja stope šestomesečnog Euribora na mesečnom nivou, što se može videti na slici broj 5.



Slika 5 – Podaci o predviđenoj stopi šestomesečnog Euribora do kraja 2022. godine

Sa grafika se jasno može videti da linearna regresija prognozira povećanje stope šestomesečnog Euribora, gde iz negativne stope, Euribor postaje pozitivan kao i da će se do kraja 2022. godine stopa istog održati u iznosu većem od 0,12%. Povećavanje stope šestomesečnog Euribora dovodi takođe i do povećanja kamata za stambene kredite uzete u evrima što smanjuje mogućnost za plasiranjem i naplatom stambenih bankarskih kredita.

3.2. Rezultati i diskusija primenom metode višestruke linearne regresije

Praćenjem pojedinačnih trendova finansijskih pokazatelja ne dozvoljava se stvaranje sveobuhvatne slike istih, te se iz tog razloga mora formirati jedinstveni pokazatelj po kome se banke i druge finansijske institucije mogu upravljati. Formiranjem grupe parametara u kojoj se zavisna promenljiva (kamatna stopa) stavlja u odnos sa dve nezavisne promenljive (stopa inflacije i stopa Euribora) poznata je kao višestruka regresija. Na slici 6. prikazana je prognoza prosečne kamatne stope stambenih kredita primenom višestruke linearne regresije na osnovu prikupljenih mesečnih podataka za 2021. i 2022. godinu.



Slika 6 - Prognoza prosečne kamatne stope stambenih kredita primenom višestruke linearne regresije za kraj 2022. godine

Promenom navedenih parametara dobijaju se različite vrednosti prosečne kamatne stope na stambene kredite. Prema prognozama koje predviđaju rast inflacije i stope šestomesečnog Euribora, predviđa se i rast prosečne kamatne stope na stambene kredite do kraja 2022. godine gde prosečna kamata stopa dostiže vrednost od 3,22%. Drastičan rast ovih kamatnih stopa loše utiče na plasman stambenih kredita od strane banaka i drugih finansijskih institucija. Kako bi banke u napetim tržišnim uslovima obezbedile konkurentnost i stabilnost u poslovanju, u ovom konkretnom slučaju mogu prilagoditi svoju kamatnu stopu kao nižu od prikazane. Manjim i prilagođenim kamatnim stopama banke imaju veću mogućnost plasmana stambenih kredita, dok se gubitak od sniženih kamatnih stopa može nadoknaditi kroz druge bankarske poslove.

4. ZAKLJUČAK

Data Mining tehnika regresije je sa svojim setom metoda omogućila da se predvide trendovi daljih tokova finansijskih pokazatelja u oblasti stambenih kredita. Iz rezultata samog istraživanja ustanovljen je nepovoljan tok finansijskih pokazatelja po pitanju uslova stambenih kredita, odnosno predviđa se otežan plasman istih. U poređenju sa drugim srodnim istraživanjima jedina razlika bi bila u tome što u ovom radu skup podataka obuhvata veći broj krucijalnih parametara. Kada je reč o višestrukoj linearnoj regresiji, stiče se utisak da su parametri za analizu u srodnim istraživanjima izabrani nasumično, dok je u ovom radu uzeto u obzir da su parametri (prediktorske varijable) u strogoj korelaciji kako bi se tačnost višestrukog linearnog modela podigla na viši nivo. Takođe ono što je drugačije u procesu analize podataka i prikaza rezultata u odnosu na srodna istraživanja jeste da u ovom radu postoji mogućnost interaktivnih izveštaja za jednostavnu i višestruku linearnu regresiju čime se povećava mogućnost za novim informacijama, odnosno prognozama različitih finansijskih pokazatelja.

Dalji razvoj istraživanja bazirao bi se u praćenju tokova finansijskih pokazatelja kako bi se za određeni vremenski period izvršila ključna predviđanja. Banke i druge finansijske institucije mogle bi razviti i optimizovati prikazani Business Intelligence model u skladu sa svojom poslovnom politikom kako bi održale stabilnost i konkurentnost na tržištu koja bi doprinela smanjenju rizika od gubitaka nastalim u neizvesnim uslovima prilikom plasmana stambenih kredita.

5. LITERATURA

- [1] Li, J. (2020).: The effects of macroeconomic factors on housing prices in China: Empirical research and linear regression analysis. Proceedings of the 2020 2nd International Conference on Economic Management and Cultural Industry (ICEMCI 2020).
- [2] Hamdar, B.; Skheita, Y.; & Bassam Hamdar, M. (2022).: An econometric approach to assess the impact of negative interest rate policy (NIRP) on real estate price inflation in the eurozone. Economics, 11(1), 49. <https://doi.org/10.11648/j.eco.20221101.17>
- [3] Republički zavod za statistiku Srbije, pristup Jun 20, 2022. dostupno na: <https://www.stat.gov.rs/>
- [4] Euribor-rates.eu, zadnji pristup Jun 22, 2022. dostupno na: <https://www.euribor-rates.eu/en/>
- [5] Narodna banka Srbije, zadnji pristup Jul 1, 2022. dostupno na: https://nbs.rs/sr_RS/indeks/
- [6] Suad, A. and Alasadi, W. S. (2017) "Review of Data Preprocessing Techniques in Data Mining," Journal of Engineering and Applied Sciences, 12, pp. 4102–4107.
- [7] Powerbi.microsoft.com, kratak pregled informacija o samom alatu, zadnji pristup Jun 10, 2022. dostupno na <https://powerbi.microsoft.com/en-us/what-is-power-bi/>
- [8] Agha, S. & Haider, A. (2014). An Introduction to Data Mining Technique. International Journal of Advancement in Engineering Technology, Management & Applied Science, 1(3), 66.
- [9] Golberg, M., Cho, H. (2010). Introduction to Regression Analysis. Department of Mathematical Sciences, University of Nevada, Las Vegas, USA, 1.
- [10] Halili, F., Rustemi, A. (2016). Predictive Modeling: Data Mining Regression Technique Applied in a Prototype, 212-213.
- [11] Freedman, D.A. (2009). Statistical Models: Theory and Practice. Cambridge University Press, 26.