

PRIMENA MAŠINSKOG UČENJA ZA PROCENU CENA I KOLIČINA RADOVA PRI IZGRADNJI STAMBENIH I STAMBENO-POSLOVNIH OBJEKATA

AN APPLICATION OF MACHINE LEARNING FOR THE ESTIMATION OF COST AND QUANTITIES OF WORKS IN THE CONSTRUCTION OF RESIDENTIAL AND RESIDENTIAL-COMMERCIAL BUILDINGS

UDK: 338.512:69]:004.4

624:519.876.3

Pregledni rad

Nevena SIMIĆ, MSc.građ.inž.¹⁾

Aleksandar DEVEDŽIĆ, MSc.građ.inž.²⁾

Marija IVANOVIĆ, MSc.građ.inž.³⁾

V. prof. dr Predrag PETRONIJEVIĆ, dip. građ. inž.⁴⁾

REZIME

Ovaj rad se bavi problemom procena potrebnih količina radova, kao i koštanja izgradnje stambenih i stambeno-poslovnih objekata korišćenjem algoritama mašinskog učenja. Osnovni cilj je analiza mogućnosti primene mašinskog učenja za razvoj modela koji će za kratko vreme pružiti dovoljno preciznu preliminarnu procenu potrebnih količina i cena glavnih radova na osnovu malog broja poznatih parametara. Istraživanje je sprovedeno na osnovu podataka o realizovanim projektima izgradnje višeporodičnih stambenih i stambeno-poslovnih objekata koji su izgrađeni u periodu od 2012. do 2020. godine na teritoriji Republike Srbije. U radu je predloženo nekoliko modela za procenu količina i cena pojedinih vrsta radova, kao i ukupne cene građevinskih radova. Rezultati analize su pokazali da se veća tačnost može postići predikcijom količina nego cena pojedinih radova. Razvijeni modeli mogu biti korisni u procesu planiranja troškova i količina potrebnog materijala u ranim fazama razvoja projekta.

Ključne reči: procena troškova, procena količina, mašinsko učenje, veštačke neuronske mreže

ABSTRACT

This paper deals with the problem of estimating the required quantities of works, as well as the cost of construction of residential and residential-commercial buildings using machine learning algorithms. The main goal is to analyze the possibility of applying machine learning for the development of a model that will in a short time provide a sufficiently precise preliminary estimate of the required quantities and cost of major works based on a small number of known parameters. The research was conducted on the basis of data on realized projects of the construction of multi-family residential and residential-commercial buildings that were constructed in the period from 2012 to 2020 on the territory of the Republic of Serbia. The paper proposes several models for estimating the quantities and cost of individual types of works, as well as the total price of construction works. The results of the analysis showed that greater accuracy can be achieved by predicting quantities than the cost of individual works. The developed models can be useful in the process of planning the cost and quantities of required material in the early stages of project development.

Key words: cost estimation, quantities estimation, machine learning, artificial neural networks

Adresa autora: ¹⁾ Asistent – student doktorskih studija, Građevinski fakultet Univerziteta u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd

E-mail: nsimic@grf.bg.ac.rs

²⁾ TX Services, Milutina Milankovića 1i, 11000 Beograd

E-mail: aleksandar.devedzic93@gmail.com

³⁾ Asistent – student doktorskih studija, Građevinski fakultet Univerziteta u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, God. rođenja: 1990.

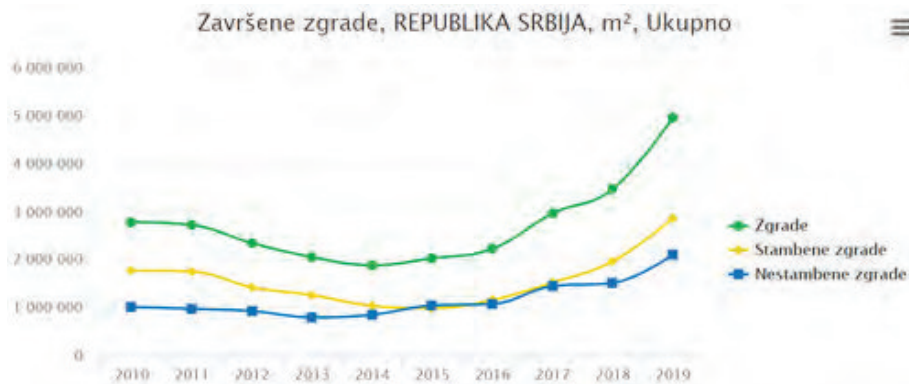
E-mail: mapetrovic@grf.bg.ac.rs

⁴⁾ Vanredni profesor, Građevinski fakultet Univerziteta u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd

E-mail: pecap@grf.bg.ac.rs

UVOD

Stambeni i stambeno-poslovni objekti čine najveći deo građevinske industrije u oblasti visokogradnje, kako u svetu, tako i kod nas. U poslednjih pet godina, u Republici Srbiji se beleži konstantan rast u oblasti građevinske industrije u pogledu broja stambenih i stambeno-poslovnih projekata. Podatak koji govori u prilog tome je da je u 2015. godini izgrađeno nešto manje od milion metara kvadratnih stambenih površina, dok je u 2019. godini izgrađeno oko tri miliona, odnosno tri puta više (slika



Slika 1. Porast broja završenih zgrada u Srbiji (preuzeto sa <https://data.stat.gov.rs/Home/Result/05030201?languageCode=sr-Latn>, pristupljeno 12.07.2021.)

1). Prema podacima Republičkog zavoda za statistiku, vrednost izvedenih radova u građevinarstvu na teritoriji Republike Srbije povećana je za 39,9 % u trećem kvartalu 2019. godine u odnosu na isti period 2018. godine (Republički zavod za statistiku, 2019). Trendu porasta građevinske aktivnosti u većoj meri su doprineli radovi na izgradnji stambenih i nestambenih zgrada (rast od 18,6%) (Republički zavod za statistiku, 2019).

Kako u poslednjih nekoliko godina postoji konstantan rast u ovoj grani građevinarstva, postavlja se pitanje da li se proces planiranja troškova i količina potrebnog materijala u ranim fazama razvoja projekta može unaprediti i ubrzati. Takođe, mnogi investitori koji ulažu svoje resurse u izgradnju stambenih objekata, nemaju tehničko znanje iz oblasti građevinarstva kako bi mogli da planiraju svoje troškove u ranim fazama investicije. Stoga se nameće pitanje kako se ovaj problem može prevazići.

Pored građevinske industrije, oblast koja poslednjih godina beleži drastičan rast u svetu jeste razvoj sistema veštačke inteligencije. Osnovni cilj veštačke inteligencije je razvoj sistema koji će rešavati problem na inteligentan način. Ovakvi sistemi su pogodni za primenu u oblastima u kojima se raspoloživo velikom količinom podataka, a kako je građevinarstvo jedna od oblasti u kojoj se svaki projekat sastoji od mnoštva podataka, spoj veštačke inteligencije i građevinske industrije je potpuno opravdan. Veštačka inteligencija nalazi veliku primenu u prevazilaženju brojnih problema u upravljanju projektima u građevinarstvu (Pan i Zhang, 2021)(Eber, 2020). U literaturi se često mogu sresti dva različita termina, a to su veštačka inteligencija i mašinsko učenje. Mašinsko učenje predstavlja oblast veštačke inteligencije u kojoj se zaključci donose na osnovu prethodnog iskustava.

Pouzdana procena troškova predstavlja dragoceni alat za efikasno planiranje budućih projekata i podlogu za kontrolu troškova u toku izvođenja radova. S obzirom na sveprisutan problem prekoračenja troškova, pouzdanost procene troškova predstavlja kritičan faktor za uspeh bilo kog projekta (Elmoussalami, 2020). Postoji veliki broj razvijenih modela mašinskog učenja za procenu koštanja različitih objekata u građevinarstvu, od kojih se najveći broj odnosi na zgrade (Tayefeh Hashemi, Eba-

dati and Kaur, 2020. (Chandanshive i Kambekar, 2019) su formirali algoritam veštačkih neuronskih mreža na osnovu baze podataka sa 78 zgrada, za koje su podaci prikupljeni od različitih učesnika na projektima iz Indije. (El-Sawalhi i Shehatto, 2014) su na osnovu 11 značajnih parametara koji su služili kao ulazne promenljive razvili model za ranu procenu troškova sa prosečnom procentualnom greškom manjom od 6%. (Du i Li, 2017) su koristili genetske algoritme kako bi optimizovali arhitekturu modela veštačkih neuronskih mreža.

Višestruka statistička regresija je metoda koja se najduže koristi za rešavanje sličnih problema, pa su (Kim, An i Kang, 2004) iz tog razloga izvršili poređenje performansi takvih modela sa modelima veštačkih neuronskih mreža i zaključivanja na osnovu slučajeva. Iako neuronske mreže funkcionišu kao crna kutija, najveća tačnost je postignuta njihovom primenom. (Jiang, 2020) je uticajne parametre za razvoj modela grupisao u tri kategorije vezane za konstrukciju, arhitekturu i hidrotehničke i električne instalacije. (Roxas i Ongpeng, 2014) su zaključili da se primenom neuronskih mreža može postići dobra mogućnost predviđanja uprkos neujednačenoj raspodeli i nepotpunosti skupa podataka.

Metodologija

Predložena metodologija za razvoj modela sadrži osam koraka (slika 2). U narednim poglavljima su detaljnije prikazani glavni koraci.

Skup ulaznih podataka

Istraživanje je sprovedeno na osnovu podataka o realizovanim projektima izgradnje višeporodičnih stambenih i stambeno-poslovnih objekata koji su izgrađeni u periodu od 2012. do 2020. godine. Ukupan broj prikupljenih projekata iznosio je 71. U mašinskom učenju su kvalitet i veličina baze podataka presudni za uspeh predikcije (Mohri, Rostamizadeh i Talwalker, 2012), pa se prikupljanju podataka pristupilo sa posebnom pažnjom. Podaci su prikupljeni kontaktiranjem izvođačkih i projektantskih firmi koje su učestvovala u realizaciji projekata na teritoriji Republike Srbije. Podaci koji su prikupljeni obuhvatali su tehnički opis objekta, projekat



Slika 2. Metodologija razvoja modela

arhitekture i predmer i predračun radova. Zbog poverljivosti podataka neće biti spominjana imena firmi koje su obezbedile projekte, kao ni tačna lokacija (ulica i broj) projekata. Prikupljanje podataka je trajalo 30 dana.

Na početku istraživanja pristupilo se analizi podataka i ograničenja koje investitor ili izvođačka firma ima na raspolaganju u ranim fazama razvoja projekta. Ti podaci predstavljaju input koji se prosleđuje u model za predikciju. Prvi korak obrade podataka je analiza fizičkih karakteristika objekta kao što su površina parcele, bruto i neto površina objekta i površina vertikalne projekcije objekta.

Bruto površine ovih objekata se kreću od 600 do 25.000 m², broj podzemnih etaža se kreće u rasponu od 0 do 2, dok se broj nadzemnih etaža kreće u rasponu od 2 do 9. Još jedan parametar koji je uzet u obzir je gustina naseljenosti u oblasti u kojoj se nalazi objekat. Parametar gustine naseljenosti je klasifikovan u tri kategorije: srednje gust, gust i veoma gust. Objekti koji su analizirani imaju pet različitih vrsta fasade: malterisana, demit, demit-kamena, demit-ventilisana i aluminijska fasada. Na objektima se razlikuju dva tipa međuspratnih konstrukcija: fert tavanica i armiranobetonska tavanica. Što se tiče načina fundiranja, objekti se razlikuju po tome da li su građeni na šipovima ili bez šipova. Kao parametar je uzeto u obzir i to da li objekat ima podzemnu garažu ili ne.

Vrste radova za koje su prikupljeni podaci o količinama radova su:

- Zemljani radovi,
- Betonski radovi,
- Armirački radovi,
- Keramičarski radovi,
- Parketarski radovi, i
- Izolaterski radovi.

Ovi radovi su izabrani iz razloga što imaju standardizovanu jedinicu mere, pa su se na osnovu toga zemljani i betonski radovi obračunavali u kubnim metrima, količina armiračkih radova u tonama, a keramičarski, parketarski i izolaterski radovi u kvadratnim metrima.

Za iste vrste radova su prikupljeni podaci o troškovima, uz dodatak zidarskih radova. U obzir je uzeta i ukupna cena građevinskih radova. Cene izrade električnih, mašinskih kao i instalacija vodovoda i kanalizacije nisu uzete u obzir, odnosno te vrste radova nisu uračunate u ukupnu cenu građevinskih radova.

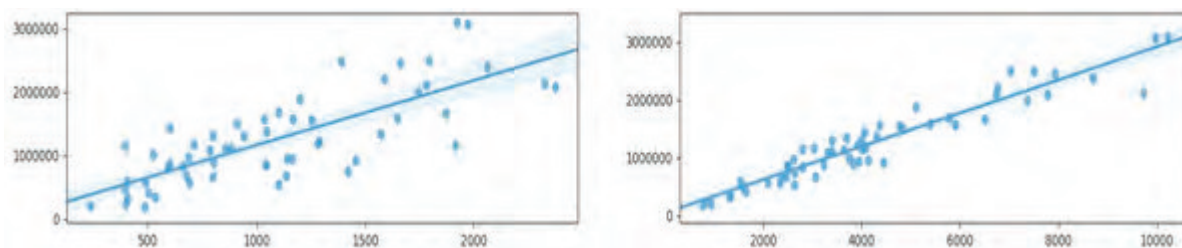
Fizičke karakteristike objekta, kao i količine i cene pojedinih radova su detaljnije analizirane kako bi se sagledala njihova raspodela i uočile eventualne ekstremne vrednosti. Projekti koji imaju bruto površinu veću od 15.000 m², odnosno neto površinu veću od 13.000 m², predstavljaju ekstremne vrednosti u pogledu površina. Takođe, u ekstremne vrednosti spadaju i projekti čija površina parcele prelazi 4.000 m² i čija vertikalna projekcija prelazi 2.000 m². Ovi podaci mogu da utiču na tačnost budućeg modela pa su iz tog razloga eliminisani. Pored toga, projekti čija je cena izgradnje veća od 450 €/m² i manja od 200 €/m² predstavljaju ekstremne vrednosti koji mogu negativno uticati na tačnost modela za procenu, te su isključeni iz dalje analize.

U mašinskom učenju se podaci koji drastično odstupaju od ostalih podataka nazivaju uljezima (engl. *outliers*) i mogu se automatski izbaciti primenom algoritma kao što je algoritam po imenu Izolaciona Šuma (engl. *Isolation Forest*), ali u ovom slučaju korišćenje ovog algoritma nije neophodno zbog malog broja podataka, te su projekti koji predstavljaju uljeze ručno eliminisani. Podaci koji predstavljaju uljeze mogu značajno uticati na ishod treniranja modela mašinskog učenja, pa je ovaj korak od izuzetne važnosti za dalji rad.

PRIPREMA PODATAKA

Međusobna zavisnost podataka

Kako bi se kreirao adekvatan model mašinskog učenja koji primenjuje odgovarajuću matematičku funkciju, potrebno je znati međusobnu zavisnost ulaznih i traženih veličina. Iz tog razloga određena je korelacija između različitih veličina. Na slici 3 je dat primer zavisnosti ukupne cene građevinskih radova u odnosu na površinu parcele i bruto površinu objekta.



Slika 3. Zavisnost ukupne cene u odnosu na: a) površinu parcele, b) bruto površinu objekta

Rad sa kategoričkim podacima

Podaci kao što su površine ili spratnost su numerički podaci, pa se kao takvi direktno mogu proslediti u model. Međutim, podaci kao što su lokacija, vrsta fundiranja ili vrsta fasade su kategorički podaci, te se kao takvi pre prosleđivanja u model moraju enkodirati.

Kategorički podaci koji imaju samo dve vrednosti, kao što je na primer podzemna garaža (da ili ne) ili vrsta fundiranja (sa šipovima ili bez šipova), mogu se jednostavno enkodirati zamenom kategorije za numeričku vrednost (1 ili 0). Međutim, podaci koji sadrže više kategoričkih vrednosti, ne mogu se jednostavno enkodirati brojevima zato što se postavlja pitanje koja vrednost će dobiti veći broj. U ovom slučaju se vrednosti kategoričkih podataka pretvaraju u kolone, a vrednost podatka u koloni koja odgovara određenoj kategoričkoj vrednosti će biti 1, dok će u ostalim kolonama koje predstavljaju kategoričke vrednosti biti 0 (tabela 1).

Skaliranje podataka

Ulazni parametri na osnovu kojih će model vršiti predikciju mogu se dosta razlikovati u pogledu reda veličine. Tako se, na primer, parametar koji predstavlja broj nadzemnih etaža kreće u rasponu od 1 do 9, dok se parametar koji predstavlja bruto površinu objekta kreće i do 14.000. Ovakva razlika u veličinama može doprineti

smanjenju tačnosti modela, pa se iz tog razloga veličine ulaznih parametara moraju skalirati. Skaliraju se samo numeričke veličine, kao što su bruto i neto površina, površina parcele i sl., a ne skaliraju se kategoričke veličine kao što je tip međuspratne konstrukcije ili tip fasade. Najčešći vidovi skaliranja podataka su:

- standardno skaliranje (skalira veličine između -1 i 1 tako da prosečna vrednost bude 0),
- robusno skaliranje (skalira veličine između dve zadate vrednosti eliminišući uljeze odnosno *outlier-e*),
- normalizacija (skalira veličine tako da teže da se prilagode normalnoj distribuciji tj. Gausovoj distribuciji),
- skaliranje između dve veličine (skalira veličine između dve zadate vrednosti, najčešće su to 0 i 1).

U praksi ne postoji jasno definisano uputstvo o tome koji vid skaliranja se kada koristi, već se u najvećem broju slučajeva ispituju sve četiri vrste skaliranja, što će biti slučaj i u ovom radu.

KREIRANJE MODELA

Definisanje ulaznih i rezultujućih promenljivih

Prvi korak u kreiranju odgovarajućeg modela je identifikacija ulaznih i rezultujućih veličina. U tabeli 2

Tabela 1. Način enkodiranja kategoričkih podataka

Vrsta međuspratne konstrukcije	Vrsta međuspratne konstrukcije	Gustina naseljenosti	Srednje gusto	Gusto	Veoma gusto
AB	1	Srednje gusto	1	0	0
Fert	0	Gusto	0	1	0
Fert	0	Veoma gusto	0	0	1
AB	1	Gusto	0	1	0
AB	1	Srednje gusto	1	0	0

Tabela 2. Prikaz ulaznih parametara i parametara koji se predviđaju

Ulazni parametri	Rezultujući parametri
Grad	Količina iskopa (m ³)
Gustina naseljenosti	Količina betonskih radova (m ³)
Tip objekta	Količina armiračkih radova (t)
Površina parcele	Količina keramičarskih radova (m ³)
Bruto površina objekta	Količina parketerskih radova (m ³)
Neto površina objekta	Količina izolaterskih radova (m ³)
Površina osnove	Cena zemljanih radova
Zauzetost parcele	Cena betonskih radova
Vrsta fundiranja	Cena armiračkih radova
Podzemna garaža	Cena zidarskih radova
Broj nadzemnih etaža	Cena keramičarskih radova
Broj podzemnih etaža	Cena parketerskih radova
Vrsta tavanice	Cena izolaterskih radova
Tip fasade	Ukupna cena objekta
Površina vertikalne projekcije objekta	

prikazani su identifikovani ulazni i rezultujući parametri (parametri koji se predviđaju).

Ukupan broj ulaznih parametara nakon enkodiranja kolona koje sadrže kategoričke vrednosti je 15. Pre nego što se vrednosti parametara proslede u model kako bi se počelo sa treniranjem modela, potrebno je odrediti koja kombinacija parametara će dati najveću tačnost. Preveliki broj parametara u modelu može dovesti do problema prevelikog podudaranja, pa je od velike važnosti odrediti koji parametri najviše utiču na tačnost modela. Ovaj problem se rešava iteracijom, koja je automatizovana korišćenjem funkcije rekursivnog uklanjanja parametara (engl. *Recursive Feature Elimination* – RFE), a dostupna je u okviru biblioteke *Scikit-Learn* specijalizovane za kreiranje modela mašinskog učenja.

Cilj rekursivnog uklanjanja parametara (engl. RFE) je odabir parametara koji daju najveću tačnost rekursivnim razmatranjem sve manjih i manjih skupova parametara. Model se prvo trenira sa jednim parametrom iz skupa svih parametara i određuje se tačnost modela preko unakrsne validacije. Potom se broj parametara uvećava i traži se ona kombinacija parametara koja daje najveću tačnost modela. Mana ove metode je veliki utrošak vremena koje se koristi zbog brojnih iteracija. Broj ulaznih parametara koji daju najveću tačnost modelu se razlikuje od modela do modela.

Prilikom razvoja modela na ovom nivou analize kao parametri su uzete u obzir tehničke karakteristike objekta

i karakteristike okruženja objekta, bez razmatranja ekonomskih pokazatelja, kao što je npr. uticaj inflacije.

DEFINISANJE I EVALUACIJA MODELA

Modeli su kreirani i testirani u programskom jeziku *Python* uz pomoć biblioteke *Scikit-Learn* koja je specijalizovana za mašinsko učenje i biblioteke *Keras* koja je specijalizovana za kreiranje neuronskih mreža (Raschka i Mirjalili, 2017). U daljem tekstu, biće predstavljeni modeli koji su imali najbolje performanse u pogledu procene potrebnih količina i cena za pojedine radove.

Evaluacijom modela uz korišćenje različitog tipa skaliranja podataka i različitih parametara koji su određeni primenom rekursivnog uklanjanja parametara, dobijeni su sledeći rezultati za procenu potrebnih količina pojedinih radova gde su zelenom bojom označeni modeli sa dobrim performansama, žutom bojom modeli srednjih performansi, a crvenom bojom modeli loših performansi (tabela 3 i 4).

Iz tabele 3 se može zaključiti da su linearni modeli, konkretno Grebena regresija i Lasso regresija, napravili dobre predikcije za količinu iskopa, količinu betonskih radova i količinu armiračkih radova, tj. za količine glavnih radova koje se odnose na konstrukciju. Za količine koje se odnose na keramičarske, parketerske i izolaterske radove, tj. za završne zanatske radove, modeli nisu napravili dobru predikciju, pa se na osnovu velike razlike između tačnosti dobijene na osnovu treniranih i testira-

Tabela 3. Modeli koji vrše predikciju za određene količine radova

Tražena veličina	Metoda	Tip skaliranja	Broj parametara	Rezultat treniranja	Rezultat unakrsne validacije	Rezultat testiranja
Količina iskopa	Grebena regresija	Robusno skaliranje	Svi parametri	0.942	0.8395	0.832
Količina betonskih radova	Grebena regresija	Robusno skaliranje	14	0.937	0.8833	0.824
Količina armiračkih radova	Lasso regresija	Robusno skaliranje	6	0.954	0.8391	0.855
Količina keramičarskih radova	Grebena regresija	Robusno skaliranje	Svi parametri	0.844	0.7825	0.535
Količina parketerskih radova	Lasso regresija	Robusno skaliranje	Svi parametri	0.876	0.6889	0.297
Količina izolaterskih radova	K – najbližih suseda K=3	Robusno skaliranje	Svi parametri	0.821	0.5751	0.711

Tabela 4. Modeli neuronskih mreža koji vrše predikciju za određene količine radova

Tražena veličina	Metoda	Aktivaciona funkcija	Tip skaliranja	Skriveni slojevi	Broj neurona	Rezultat treniranja	Rezultat testiranja
Količina iskopa	Neuronske mreže	Relu	Standardno skaliranje	5	100	0.960	0.824
Količina betonskih radova	Neuronske mreže	Relu	Standardno skaliranje	5	150	0.927	0.883
Količina armiračkih radova	Neuronske mreže	Relu	Standardno skaliranje	7	150	0.902	0.871
Količina keramičarskih radova	Neuronske mreže	Relu	Standardno skaliranje	7	150	0.862	0.837
Količina parketerskih radova	Neural Network	Relu	Standardno skaliranje	7	150	0.941	0.902
Količina izolaterskih radova	Neuronske mreže	Relu	Standardno skaliranje	5	150	0.789	0.520

Tabela 5. Modeli koji vrše predikciju za određene cene radova

Tražena veličina	Metoda	Tip skaliranja	Broj parametara	Rezultat treniranja	Rezultat unakrsne validacije	Rezultat testiranja
Ukupna cena građevinskih radova	Pojačavanje gradijenta	Robusno skaliranje	Svi parametri	0.989	0.893	0.873
Cena zemljanih radova	K– najbližih suseda K=9	Robusno skaliranje	Svi parametri	0.505	0.3106	0.046
Cena betonskih radova	K– najbližih suseda K=7	Robusno skaliranje	Svi parametri	0.736	0.6971	0.538
Cena armiračkih radova	K– najbližih suseda K=2	Robusno skaliranje	Svi parametri	0.743	0.5369	0.552
Cena zidarskih radova	K– najbližih suseda K=2	Robusno skaliranje	Svi parametri	0.809	0.4346	0.054
Cena keramičarskih radova	K– najbližih suseda K=9	Robusno skaliranje	Svi parametri	0.471	0.3602	0.203
Cena parketerskih radova	K– najbližih suseda K=7	Robusno skaliranje	Svi parametri	0.905	0.6016	0.411
Cena izolaterskih radova	K– najbližih suseda K=2	Robusno skaliranje	Svi parametri	0.535	0.164	0.019

Tabela 6. Modeli neuronskih mreža koji vrše predikciju za određene cene radova

Tražena veličina	Model	Aktivaciona funkcija	Tip skaliranja	Skriveni slojevi	Broj neurona	Rezultat treniranja	Rezultat testiranja
Ukupna cena građevinskih radova	Neuronske mreže	Relu	Standardno skaliranje	6	150	0.937	0.895
Cena zemljanih radova	Neuronske mreže	Relu	Standardno skaliranje	7	100	0.33245	0.445
Cena betonskih radova	Neuronske mreže	Relu	Standardno skaliranje	7	50	0.86475	0.761
Cena armiračkih radova	Neuronske mreže	Relu	Standardno skaliranje	7	75	0.754703	0.698
Cena zidarskih radova	Neuronske mreže	Relu	Standardno skaliranje	8	100	0.610076	0.703
Cena keramičarskih radova	Neuronske mreže	Relu	Standardno skaliranje	7	150	0.431365	0.306
Cena parketerskih radova	Neuronske mreže	Relu	Standardno skaliranje	7	150	0.801246	0.598
Cena izolaterskih radova	Neuronske mreže	Relu	Standardno skaliranje	8	150	0.162353	0.199

nih podataka može zaključiti da se javlja greška velikog podudaranja. Može se primetiti da je u svim slučajevima najveću tačnost dao vid robusnog skaliranja podataka koji je karakterističan po tome što skalira veličine između dve zadate vrednosti eliminišući uljeze odnosno *outlier*-e.

Kada su u pitanju rezultati koji su dobijeni primenom neuronskih mreža, može se primetiti da Relu aktivaciona funkcija daje najbolje rezultate, kao i da je najbolji tip skaliranja standardni tip. Primećuje se da su modeli koji se odnose na procenu količina radova za izradu armiranobetonske konstrukcije dali dobre rezultate, kao i model za procenu količine parketerskih radova, dok su modeli za procenu ostalih količina imali lošije performanse što je bio i slučaj sa modelima iz prethodne tabele.

U tabelama 5 i 6 su prikazani dobijeni rezultati za procenu potrebnih cena pojedinih radova gde su zelenom bojom označeni modeli sa dobrim performansama, žutom bojom modeli srednjih performansi, a crvenom bojom modeli loših performansi.

Iz tabele 5 se može zaključiti da je jedino model po imenu Pojačavanje gradijenta (engl. *Gradient Boosting Regression*) napravio dobru predikciju za ukupnu cenu

građevinskih radova. Za cene koje se odnose na pojedinačne vrste radova najbolje predikcije je dao model K-najbližih suseda, ali i taj algoritam ima jako loše performanse. Ovako loši rezultati se mogu objasniti u razlici u kvalitetu ugrađenih materijala kao i time da li je kompanija sama izvodila radove ili je unajmljivala podizvođače. Ove informacije nisu bile dostupne na početku, pa samim tim nije bilo moguće uvrstiti ih u model.

Primenom neuronskih mreža za procenu cena radova nisu dobijeni bolji rezultati u poređenju sa modelima iz tabele 5. Malo poboljšanje performansi primećeno je kod modela za procenu cena betonskih radova. Primećeno je da je model neuronske mreže za procenu ukupne cene građevinskih radova imao najbolje performanse, što je bio slučaj i sa prethodno predstavljanim modelima.

ANALIZA REZULTATA

Prema vodiču PMBOK (engl. *A Guide to the Project Management Body of Knowledge*) (PMI, 2017) koji je izdao američki Institut za upravljanje projektima (engl. *Project Management Institut*) tačnost procene troškova se povećava za razvojem projekta i u kasnijim fazama ona iznosi -5% do 10%. Ocena tačnosti modela u radu je

Tabela 7. Prikaz veličine srednje apsolutne procentualne greške kod procene količina

Tražena veličina	Model	Aktivaciona funkcija	Tip skaliranja	Skriveni slojevi	Broj neurona	MAPE
Količina iskopa	Neuronske mreže	Relu	Standardno skaliranje	5	100	17.6
Količina betonskih radova	Neuronske mreže	Relu	Standardno skaliranje	5	150	11.7
Količina armiračkih radova	Neuronske mreže	Relu	Standardno skaliranje	7	150	12.9
Količina keramičarskih radova	Neuronske mreže	Relu	Standardno skaliranje	7	150	16.3
Količina parketerskih radova	Neuronske mreže	Relu	Standardno skaliranje	7	150	9.8
Količina izolaterskih radova	K-najbližih suseda K=3	/	Robusno skaliranje	/	/	28.9

Tabela 8. Prikaz veličine srednje apsolutne procentualne greške kod procene cena

Tražena veličina	Model	Aktivaciona funkcija	Tip skaliranja	Skriveni slojevi	Broj neurona	MAPE
Ukupna cena građevinskih radova	Pojačavanje gradijenta	/	Robusno skaliranje	/	/	12.7
Cena zemljanih radova	Neuronske mreže	Relu	Standardno skaliranje	7	100	55.5
Cena betonskih radova	Neuronske mreže	Relu	Standardno skaliranje	7	50	23.9
Cena armiračkih radova	Neuronske mreže	Relu	Standardno skaliranje	7	75	30.2
Cena zidarskih radova	Neuronske mreže	Relu	Standardno skaliranje	8	100	29.7
Cena keramičarskih radova	Neuronske mreže	Relu	Standardno skaliranje	7	150	69.4
Cena parketerskih radova	Neuronske mreže	Relu	Standardno skaliranje	7	150	40.2
Cena izolaterskih radova	Neuronske mreže	Relu	Standardno skaliranje	8	150	80.1

ispitana primenom mere za srednju apsolutnu procentualnu grešku MAPE. MAPE je jedna od najčešće korišćenih mera za ocenu tačnosti prognoze (Kim i Kim, 2016) i upotrebljavana je u velikom broju istraživanja vezanih za procenu troškova i količina u građevinarstvu (npr. (Kovacevic i ostali., 2021)(Li, Baek i Ashuri, 2021)(Karaca i ostali, 2020)). Srednja apsolutna procentualna greška je definisana sledećom formulom:

$$MAPE = \frac{1}{n} * \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| * 100 \quad (1)$$

gde A_t predstavlja stvarnu vrednost, a F_t prognoziranu vrednost.

Analizom rezultata predstavljenih u prethodnom poglavlju, može se zaključiti da se veća tačnost može postići predikcijom količina nego cena pojedinih radova. Količine pojedinih radova ne zavise od kvaliteta radova kao ni od toga da li su angažovane podizvođačke firme, već u najvećoj meri od tipa konstrukcije, površine objekta, spratnosti i sl., pa se samim tim mogu lakše predvideti u ranim fazama planiranja. Slučajevi u kojima količina opada sa porastom gabarita objekta je praktično nemoguća, a veća razlika u količini pojedinih radova za dva objekta koja su slična po površini i spratnosti može se opisati različitim tipom konstrukcije (npr. armiranobetonska tavanica i fert tavanica). Jako loše performanse modela za procenu količina izolaterskih radova mogu se opisati činjenicom da postoje različite vrste izolaterskih proizvoda, pa je zbog kvaliteta materijala u nekim slučajevima zahtevan veći utrošak od prosečnog. Ovo je jedna

od retkih vrsta radova kod kojih kvatitet materijala može drastično da utiče na potrebnu količinu.

Kada je u pitanju predviđanje cena pojedinih radova, zaključak je da klasični modeli mašinskog učenja, kao i neuronske mreže nisu dali zadovoljavajuće rezultate. Samo predviđanje ukupne cene građevinskih radova je dalo dobre rezultate gde je korišćen algoritam po imenu Pojačavanje gradijenta (*Gradient Boosting*). U Tabeli 8 su date veličine srednjih apsolutnih procentualnih grešaka MAPE za modele za procenu ukupne cene građevinskih radova i cena pojedinačnih vrsta radova.

Međutim, ova greška je čak i veća jer se može videti da modeli u većini slučajeva imaju jako niske rezultate treniranja, što znači da modeli nisu uspeali da nađu konkretnu zavisnost između fizičkih karakteristika objekta i cena pojedinih radova.

Za razliku od procene količina, ovakvi rezultati se mogu opisati činjenicom da proračun cene u velikoj meri zavisi od kvaliteta radova, kao i od toga da li su za obavljanje tog posla angažovani podizvođači, a takvi podaci nisu bili na raspolaganju tokom istraživanja.

ZAKLJUČAK I PRAVCI DALJIH ISTRAŽIVANJA

Nakon svih rezultata koji su predstavljeni u ovom radu, postavlja se pitanje na koji način se oni mogu poboljšati. Na osnovu predikcije predloženih modela može se drastično ubrzati i poboljšati donošenje odluke o pristupanju realizaciji projekta u ranim fazama razvoja projekta, a daljim usavršavanjem je moguće poboljšati buduće korake planiranja. Sam princip mašinskog učenja zasniva se na učenju iz velikog broja podataka na osnovu kojih je potrebno doneti zaključke. Za kreiranje modela

o kojima je prethodno bilo reči, korišćeni su podaci sa 52 građevinska projekta što u oblasti mašinskog učenja predstavlja minimalni broj podataka, a naročito u oblasti neuronskih mreža (dubokog učenja). Kod neuronskih mreža, za razliku od klasičnih modela, u većini slučajeva dolazi do povećanja tačnosti sa povećanjem broja podataka, pa se na osnovu toga može zaključiti kako bi povećanje broja projekata doprinelo poboljšanju performansi kreiranih modela.

Izvor podataka za kreiranje modela u ovom istraživanju bili su tehnički opis i predmer i predračun radova koji pružaju limitiran broj podataka. Kako bi se dobio realistični uvid o količinama radova kao i o cenama tih radova, ali i o ukupnoj ceni projekta, umesto podataka iz predmera i predračuna potrebno je koristiti podatke iz projekta izvedenog objekta. Predmer i predračun predstavlja procenu, koja sama po sebi nosi određen nivo greške, dok projekat izvedenog objekta sadrži stvarne podatke o utrošenim količinama i utrošenom novcu, te kao takav predstavlja mnogo relevantniji izvor informacija. Podaci dobijeni iz tehničkog opisa, kao što su podaci o površinama, parceli, spratnosti, vrsti konstrukcije i završne obrade predstavljaju dobru osnovu za pravljenje modela čiji su ulazni parametri fizičke osobine objekta. Međutim u tehničkom opisu se ne mogu naći podaci o zahtevanom kvalitetu radova kao i o tome da li je za izvođenje pojedinih radova angažovana neka podizvođačka firma. Kvalitet materijala za radove kao što su parketerski i izolaterski radovi varira u velikoj meri, a samim tim i cena tih radova, pa se bez podataka o stepenu kvaliteta ne može napraviti adekvatna procena cena, kao i bez podataka o angažovanju podizvođačkih firmi od čega takođe zavisi cena.

Pored 15 korišćenih parametara koji opisuju tehničke karakteristike objekta i karakteristike okruženja objekta, u budućem radu je potrebno uzeti u obzir i ekonomske pokazatelje. Takvi parametri bi opisivali ekonomsko stanje društva u celini, kao i građevinskog sektora kroz vreme. S obzirom na to da su projekti za koje su prikupljeni podaci za izradu ovog istraživanja realizovani u periodu od 2012. do 2020. godine, potrebno je uzeti u obzir inflaciju, kao i promene cena materijala, radne snage i mehanizacije u navedenom periodu. Promene cena navedenih resursa mogu biti značajne i imati veliki uticaj na cene radova.

Na kraju, mora se spomenuti da su podaci za 52 objekta na osnovu kojih se vršila analiza prikupljeni od strane više od 40 firmi što predstavlja veliki problem u vidu kreiranja modela iz razloga velike varijabilnosti. Svaka projektantska firma ima svoj određeni stil projektovanja, a sam projekat diktira način izvođenja radova, što može dovesti do lošijih performansi modela, naročito ukoliko se raspolaže malim brojem podataka. Kada bi se predikcija vršila na osnovu više od 500 projekata koji su izvedeni od strane 20 do 30 kompanija, naziv kompanije bi bitno uticao na rezultat. Međutim, kako ne postoji univerzalni model koji bi imao dobre performanse za sve vrste projekata, najbolje performanse modela bi se postigle kada bi se model kreirao samo na osnovu podataka

koje je jedna kompanija obezbedila, tj. kada bi se model pravio konkretno za određenu kompaniju. Ovo može drastično da limitira primenu ovakvog sistema, pogotovu u Republici Srbiji, iz razloga što je broj firmi koji je izgradio više od 50 stambenih ili stambeno-poslovnih objekata (koji su bili tema istraživanja) veoma mali.

LITERATURA

- [1] Chandanshive, V. and Kambekar, A. R. (2019) 'Estimation of Building Construction Cost Using Artificial Neural Networks', *Journal of Soft Computing in Civil Engineering*, 3(1), pp. 91–107. Available at: http://www.jsoftcivil.com/article_89032.html http://www.jsoftcivil.com/article_89032_9a2e28148ec5b773f259194d5709c658.pdf.
- [2] Du, Z. and Li, B. (2017) 'Construction project cost estimation based on improved BP neural network', *Proceedings – 2017 International Conference on Smart Grid and Electrical Automation, ICSGEA 2017*, 2017-January, pp. 223–226. doi: 10.1109/ICSGEA.2017.162.
- [3] Eber, W. (2020) 'Potentials of artificial intelligence in construction management', *Organization, Technology and Management in Construction*, 12(1), pp. 2053–2063. doi: 10.2478/otmcj-2020-0002.
- [4] El-Sawalhi, N. I. and Shehatto, O. (2014) 'A Neural Network Model for Building Construction Projects Cost Estimating', *Journal of Construction Engineering and Project Management*, 4(4), pp. 9–16. doi: 10.6106/jcepm.2014.4.4.009.
- [5] Elmousalami, H. H. (2020) 'Artificial Intelligence and Parametric Construction Cost Estimate Modeling: State-of-the-Art Review', *Journal of Construction Engineering and Management*, 146(1), p. 03119008. doi: 10.1061/(asce)co.1943-7862.0001678.
- [6] Jiang, Q. (2020) 'Estimation of construction project building cost by back-propagation neural network', *Journal of Engineering, Design and Technology*, 18(3), pp. 601–609. doi: 10.1108/JEDT-08-2019-0195.
- [7] Karaca, I. *et al.* (2020) 'Improving the Accuracy of Early Cost Estimates on Transportation Infrastructure Projects', *Advances in Civil Engineering*, 36(5), pp. 1–11. doi: 10.1061/(ASCE)ME.1943-5479.0000819.
- [8] Kim, G. H., An, S. H. and Kang, K. I. (2004) 'Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning', *Building and Environment*, 39(10), pp. 1235–1242. doi: 10.1016/j.buildenv.2004.02.013.
- [9] Kim, S. and Kim, H. (2016) 'A new metric of absolute percentage error for intermittent demand forecasts', *International Journal of Forecasting*. Elsevier B.V., 32(3), pp. 669–679. doi: 10.1016/j.ijforecast.2015.12.003.

- [10] Kovacevic, M. *et al.* (2021) ‘Construction cost estimation of reinforced and prestressed concrete bridges using machine learning’, *Gradjevinar*, 73(1), pp. 1–13. doi: 10.14256/JCE.2738.2019.
- [11] Li, M., Baek, M. and Ashuri, B. (2021) ‘Forecasting Ratio of Low Bid to Owner’s Estimate for Highway Construction’, *Journal of Construction Engineering and Management*, 147(1), p. 04020157. doi: 10.1061/(asce)co.1943-7862.0001970.
- [12] Mohri, M., Rostamizadeh, A. and Talwalker, A. (2012) *Machine learning, Foundations of Machine Learning*. Cambridge, Massachusetts: The MIT Press.
- [13] Pan, Y. and Zhang, L. (2021) ‘Roles of artificial intelligence in construction engineering and management: A critical review and future trends’, *Automation in Construction*. Elsevier B.V., 122(November 2020), p. 103517. doi: 10.1016/j.autcon.2020.103517.
- [14] PMI (2017) *A guide to the project management body of knowledge (PMBOK guide)*. Project Management Institut.
- [15] Raschka, S. and Mirjalili, V. (2017) *Python Machine Learning – Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow*. 2nd editio, *Taiwan Review*. 2nd editio. Birmingham UK: Packt Publishing Ltd.
- [16] Republički zavod za statistiku (2019) *Trendovi, III kvartal, 2019*. Beograd.
- [17] Roxas, C. L. C. and Ongpeng, J. M. C. (2014) ‘An Artificial Neural Network Approach to Structural Cost Estimation of Building Projects in the Philippines’, *DLSU Research Congress*, pp. 1–8. Available at: <https://www.dlsu.edu.ph/wp-content/uploads/pdf/conferences/research-congress-proceedings/2014/SEE/SEE-I-005-FT.pdf>
http://www.dlsu.edu.ph/conferences/dlsu_research_congress/2014/_pdf/proceedings/SEE-I-005-FT.pdf.
- [18] Tayefeh Hashemi, S., Ebadati, O. M. and Kaur, H. (2020) ‘Cost estimation and prediction in construction projects: a systematic review on machine learning techniques’, *SN Applied Sciences*. Springer International Publishing, 2(10), pp. 1–27. doi: 10.1007/s42452-020-03497-1.