

Razvoj metodologije procjene prosječnih i medijalnih tržišnih cijena stambenih nekretnina i najamnina po jedinicama lokalne samouprave primjenom metoda strojnog učenja

Zagreb, siječanj 2026.

Studije

Ekonomski institut, Zagreb

STUDIJA

Razvoj metodologije procjene prosječnih i medijalnih tržišnih cijena stambenih nekretnina i najamnina po jedinicama lokalne samouprave primjenom metoda strojnog učenja

Urednica i autorica:

Prof. dr. sc. Maruška Vizek

Naručitelj:

Ministarstvo prostornoga uređenja, graditeljstva
i državne imovine Republike Hrvatske

Zagreb, siječanj 2026.

Sadržaj

1. Uvod	4
2. Ograničenja dostupnih pokazatelja o cijenama stanova i najmaninama	7
3. Metode strojnog učenja korištene u procjeni medijalnih i prosječnih cijena stanova te medijalnih najamnina	11
3.1. Podjela uzorka i treniranje modela	11
3.2. Modeli korišteni u strojnom učenju	12
3.3. Evaluacija modela	15
3.4. Izbor optimalnih objašnjavajućih varijabli za konačni prediktivni model	15
3.5. Izbor zaključnog modela i metode za procjenu cijene stanova	16
4. Podaci korišteni u procjeni medijalnih i prosječnih cijena stanova te medijalnih najamnina primjenom metoda strojnog učenja	18
Demografska obilježja	19
Stambene karakteristike	20
Prostorne i geografske značajke	21
Infrastrukturna dostupnost	22
Ekonomski i tržišni pokazatelji	22
Prostorni pomak ciljane varijable	23
5. Procjena medijalnih cijena stanova za jedinice lokalne samouprave primjenom metoda strojnog učenja	24
5.1. Konstrukcija prostornog pomaka ciljane varijable	24
5.2. Formiranje analitičkog uzorka i podjela na trening i test skup	26
5.3. Usporedba performansi modela strojnog učenja sa proširenim skupom objašnjavajućih varijabli radi identifikacije ključnih varijabli za predikciju medijalnih cijena stanova	27
5.4. Identifikacija ključnih varijabli za predikciju medijalnih cijena stanova	33
5.5. Evaluacija modela medijalnih cijena stanova s izabranim objašnjavajućim varijablama	36
5.6. Zaključni model za procjenu medijalnih cijena stanova u jedinicama lokalne samouprave i rezultati procjene modela	37

6. Procjena medijalnih najamnina za jedinice lokalne samouprave	44
6.1. Opis podataka i metodološki pristup procjeni medijalnih najamnina	44
6.2. Rezultati procjene medijalnih najamnina	50
7. Procjena prosječnih cijena stanova za jedinice lokalne samouprave primjenom metoda strojnog učenja	56
7.1. Konstrukcija prostornog pomaka ciljane varijable	57
7.2. Formiranje analitičkog uzorka i podjela na trening i test skup	58
7.3. Usporedba performansi modela strojnog učenja sa proširenim skupom objašnjavajućih varijabli radi identifikacije ključnih varijabli za predikciju medijalnih cijena stanova	59
7.4. Identifikacija ključnih varijabli za predikciju medijalnih cijena stanova	64
7.5. Evaluacija modela medijalnih cijena stanova s izabranim objašnjavajućim varijablama	65
7.6. Zaključni model za procjenu medijalnih cijena stanova u jedinicama lokalne samouprave i rezultati procjene modela	67
8. Određivanje minimalnog praga broja transakcija za robustnu procjenu lokalnih tržišnih vrijednosti stambenih nekretnina	74
9. Zaključak	82
Literatura	86
Popis kratica	88
Popis slika	89
Popis tablica	90
Prilozi Studiji	91

1. Uvod

Pouzdani i prostorno cjeloviti podaci o kretanjima cijena stambenih nekretnina i najamnina na lokalnim tržištima predstavljaju ključnu infrastrukturu za oblikovanje učinkovitih, pravednih i fiskalno održivih stambenih politika. Tržišne cijene kupoprodaje i najma izravno utječu na dostupnost stanovanja, regionalni razvoj, socijalnu uključenost i mobilnost stanovništva, ali i na dizajn i provedbu brojnih instrumenata javnih politika, uključujući sustave potpora, subvencija i poreznih olakšica. U uvjetima izraženih prostornih razlika među jedinicama lokalne samouprave te heterogenosti lokalnih tržišta nekretnina, postojanje konzistentnih i metodološki utemeljenih procjena tržišnih cijena na lokalnoj razini preduvjet je za donošenje informiranih i ujednačenih odluka u području stanovanja.

Nepostojanje pouzdanih podataka o cijenama stambenih nekretnina i najamnina na razini pojedinih jedinica lokalne samouprave ima višestruke negativne posljedice na društveni i gospodarski život te na provedbu javnih politika. U nedostatku relevantnih tržišnih informacija, javne institucije nisu u mogućnosti objektivno procijeniti stvarnu razinu priuštivosti stanovanja u pojedinim sredinama, što može dovesti do nejednake raspodjele potpora i situacija u kojima se javna sredstva dodjeljuju na temelju statistički nestabilnih ili nereprezentativnih pokazatelja. Istodobno, izostanak lokalno specifičnih podataka otežava planiranje stambene izgradnje, razvoj komunalne i društvene infrastrukture te procjenu potreba stanovništva, osobito u demografski osjetljivim i slabije razvijenim područjima. Na razini kućanstava, manjak transparentnih informacija o lokalnim tržišnim uvjetima povećava neizvjesnost pri donošenju odluka o kupnji ili najmu stana, dok na nacionalnoj razini produbljuje regionalne nejednakosti i ograničava mogućnost praćenja učinaka javnih politika u prostoru.

Analiza dostupnih administrativnih i tržišnih izvora podataka u Republici Hrvatskoj pokazuje da informacijski sustav eNekretnine, unatoč svojoj podatkovnoj iscrpnosti, ne omogućuje izravan izračun prosječnih i medijalnih tržišnih cijena stanova i najamnina za sve jedinice lokalne samouprave. Štoviše, samo manji broj jedinica bilježi dovoljan broj kupoprodajnih transakcija za statistički pouzdan izračun prosječnih i medijalnih cijena, dok su podaci o najamnima izrazito rijetki i u većini jedinica u potpunosti izostaju. Posljedično, bez dodatnog analitičkog okvira nije moguće osigurati konzistentnu i objektivnu primjenu propisa i programa koji se oslanjaju na tržišne cijene nekretnina u svim dijelovima zemlje.

Na potrebu uspostave pouzdanog i metodološki ujednačenog sustava procjene tržišnih cijena dodatno upućuju i aktualni instrumenti stambene politike Republike Hrvatske proizašli iz **Nacionalnog plana stambene politike Republike Hrvatske do 2030. godine**,

strateškog dokumenta usvojenog radi unapređenja priuštivosti, održivosti i dostupnosti stanovanja. **Program priuštivog najma** usmjeren je na aktivaciju postojećeg, često neiskorištenog stambenog fonda kroz davanje stanova u dugoročni najam uz potporu države, uključujući i pozivanje privatnih vlasnika da svoje prazne nekretnine stave u funkciju najma uz financijsku potporu i upravljanje od strane Agencije za pravni promet i posredovanje nekretninama (APN). Program obuhvaća postupke javnih poziva, kriterije uključivanja, bodovanja i način utvrđivanja lokalno priuštive najamnine, pri čemu je razlika između priuštive i medijalne najamnine u određenoj jedinici lokalne samouprave financirana iz državnog proračuna, čime se izravno pretpostavlja postojanje pouzdanih medijalnih najamnina kao referentnih vrijednosti.

S druge strane, **Pravilnik o potporama za stjecanje prve stambene nekretnine** (Narodne novine 86/2025) pruža poticaje poput povrata plaćenog poreza i dijela PDV-a pri kupnji prve nekretnine, osobito za mlade obitelji do 45 godina, uz definiciju prihvatljivih tržišnih cijena nekretnina koje se temelje na prosječnim cijenama stanova po jedinicama lokalne samouprave. Obje potpore pretpostavljaju postojanje pouzdanih, lokalno specifičnih pokazatelja tržišnih cijena kupoprodaje i najma – pokazatelja koji trenutačno nisu dostupni za velik dio jedinica lokalne samouprave. Ova strukturalna odsutnost podataka dodatno naglašava potrebu za razvojem metodologije koja omogućuje objektivne i replikabilne procjene tržišnih cijena i najamnina, što je ključno za dosljednu provedbu stambene politike prema ciljevima Nacionalnog plana do 2030. godine.

U tom kontekstu, cilj ove studije je razvoj integrirane metodologije procjene tržišnih cijena stanova i najamnina na razini svih jedinica lokalne samouprave primjenom suvremenih statističkih metoda i metoda strojnog učenja. Metodološki pristup za procjenu medijalnih najamnina temelji se na procjeni medijalnih tržišnih cijena stanova za referentnu godinu za sve jedinice lokalne samouprave korištenjem metoda strojnog učenja. Proces procjene započinje procjenom modela medijalnih cijena nekretnina za dio jedinica lokalne samouprave za koje postoje dostupni podaci. Modeli koriste dostupna demografska, ekonomska, prostorna, geografska, infrastrukturna i stambena obilježja jedinica lokalne samouprave u svrhu predikcije medijalnih cijena nekretnina u tim jedinicama.

U tu svrhu testiraju se i uspoređuju prediktivne sposobnosti različitih regresijskih algoritama strojnog učenja, uključujući klasične linearne modele i njihove penalizirane varijante (LASSO, Ridge i Elastic Net), regresijska stabla, metode ansambala poput slučajnih šuma (Random Forest) te gradijentno pojačane modele (XGBoost). U modele je dodatno uključen i prostorni pomak, odnosno vrijednost medijalne cijene stana u prvoj najbližoj jedinici lokalne samouprave koja je izračunata primjenom metode K najbližih susjeda (KNN). Model koji je u tom procesu pokazao najbolje prediktivne mogućnosti, korišten je zatim za predikciju medijalnih cijena stanova u svim jedinicama lokalne

samouprave u kojima taj podatak nije dostupan. Nakon toga se na odnosu medijalnih cijena stanova i medijalnih iznosa najamnina u jedinicama u kojima su dostupna oba podatka, računa udio medijalne najamnine u cijeni stana. Taj udio je zatim primijenjen kako bi se na osnovi što procijenjenih, a što stvarno ostvarenih medijalnih cijena stanova izračunala medijalna najamnina za sve jedinice lokalne samouprave koji ne bilježe taj podatak. Ovako dobivena replikabilna kombinacija procijenjenih i stvarnih podataka o medijalnim najamninama u jedinicama lokalne samouprave može se koristiti za potrebe primjene Programa priuštivog najma.

Opisani analitički okvir i logika modeliranja primjenjuju se i na procjenu prosječnih tržišnih cijena stanova, pri čemu se razlikuje isključivo definicija ciljne varijable. Drugim riječima, isti set eksplanatornih varijabli i iste metode strojnog učenja su korištene (i testirane) kako bi se dobila procjena prosječnih cijena stanova za one jedinice lokalne samouprave koje ne bilježe taj podatak. Takav pristup osigurava konzistentnost i usporedivost procjena prosječnih i medijalnih cijena stanova među jedinicama lokalne samouprave te omogućuje fleksibilnu primjenu metodologije u ovisnosti o zahtjevima pojedinih propisa ili programa javnih politika. Ovako procijenjene prosječne cijene stanova za sve jedinice lokalne samouprave se zatim mogu koristiti za potrebe primjene Programa potpora za stjecanje prve stambene nekretnine.

Dodatno, studija se bavi pitanjem statističke pouzdanosti izravnih izračuna tržišnih pokazatelja (medijalne i prosječne cijene stanova) u jedinicama s manjim brojem transakcija. Iako važeći zakonski okvir dopušta objavu izračuna u jedinicama s najmanje deset kupoprodajnih transakcija s ciljem zaštite prava na privatnost podataka individualnih sudionika na tržištu nekretnina (odnosno da bi se spriječila rekonstrukcija podataka o individualnim transakcijama), empirijska analiza pokazuje da takav prag često nije dovoljan za dobivanje stabilnih, reprezentativnih i vremenski konzistentnih procjena. Stoga se razvija analitički okvir za utvrđivanje minimalnog broja transakcija potrebnih za pouzdan izračun prosječnih i medijalnih cijena, uzimajući u obzir varijabilnost cijena i osjetljivost statističkih pokazatelja na male uzorke. Rezultat je argumentirana preporuka praga ispod kojeg se izravni izračuni smatraju nepouzdanima te je opravdano koristiti modelirane procjene u službene svrhe.

Razvijena metodologija pruža transparentan, replikabilan i operativno primjenjiv okvir koji Ministarstvu prostornoga uređenja, graditeljstva i državne imovine Republike Hrvatske omogućuje sustavno praćenje tržišnih cijena stanova i najamnina u svim jedinicama lokalne samouprave, uključujući one u kojima izravni tržišni podaci ne postoje ili nisu statistički pouzdani, čime se stvara čvrsta analitička osnova za pravednu i učinkovitu provedbu stambenih politika na nacionalnoj razini. Studiju je za potrebe Ministarstva prostornoga uređenja, graditeljstva i državne imovine izradio Ekonomski institut, Zagreb, a autorica studije je Maruška Vizek.

2. Ograničenja dostupnih pokazatelja o cijenama stanova i najmaninama

Informacijski sustav eNekretnine predstavlja ključni administrativni izvor podataka o kupoprodajnim transakcijama i najmu stambenih nekretnina u Republici Hrvatskoj te čini temelj za izračun pokazatelja kretanja cijena na tržištu nekretnina. Sustav obuhvaća podatke prikupljene od strane javnih bilježnika, Porezne uprave, jedinica lokalne samouprave i drugih nadležnih institucija te omogućuje detaljan uvid u ostvarene transakcijske cijene, karakteristike nekretnina i prostornu distribuciju tržišne aktivnosti. Međutim, unatoč svojoj važnosti i podatkovnoj iscrpnosti, eNekretnine imaju niz strukturnih ograničenja koja značajno otežavaju njihovu izravnu primjenu za potrebe izračuna pouzdanih tržišnih pokazatelja na razini svih jedinica lokalne samouprave.

Najvažnije ograničenje proizlazi iz izrazito neravnomjerne prostorne raspodjele transakcija. Dok veće urbane sredine bilježe velik broj kupoprodaja i relativno stabilne tržišne obrasce, većina jedinica lokalne samouprave – osobito u ruralnim, demografski slabije razvijenim i perifernim područjima – bilježi vrlo mali broj transakcija ili ih u pojedinim godinama uopće nema. Posljedično, izravni izračun prosječnih i medijalnih cijena u takvim jedinicama (ako je uopće moguć) temelji se na malim uzorcima koji su podložni snažnim utjecajima ekstremnih vrijednosti, strukturne heterogenosti nekretnina i slučajnih tržišnih fluktuacija, što dovodi u pitanje statističku pouzdanost dobivenih pokazatelja.

Ograničenja su još izraženija u slučaju podataka o najamninama. Sustav eNekretnine ne obuhvaća sustavno sve ugovore o najmu, već prvenstveno one koji su formalno prijavljeni i evidentirani, što rezultira vrlo ograničenim prostornim obuhvatom i izrazito malim brojem jedinica lokalne samouprave za koje je moguće izravno izračunati medijalne ili prosječne iznose najamnina. Ugovore o najmu u sustav eNekretnine upisuju jedinice lokalne samouprave i to ili veliki gradovi ili županije (za sve lokalne jedinice koje nisu veliki gradovi) te je očito da u velikom broju slučajeva ni županije, a ni pojedini veliki gradovi, ne upisuju podatke iz ugovora o najmu u sustav. Kada tome nadodamo činjenicu da u određenom dijelu manjih i slabije razvijenih lokalnih jedinica uopće nema tijekom jedne godine zabilježenih novih ugovora o najmu, postaje jasnije zašto u velikom broju jedinica lokalne samouprave podaci o najmu uopće ne postoje, dok su u onima gdje postoje često nedostatni za izračun reprezentativnih pokazatelja tržišnih uvjeta.

Dodatno ograničenje odnosi se na zakonski okvir objave podataka. Sukladno važećim propisima, izračuni prosječnih i medijalnih cijena nekretnina ne mogu se javno objavljivati za jedinice lokalne samouprave koje bilježe manje od deset transakcija godišnje. Iako

je ova odredba opravdana s aspekta zaštite privatnosti i sprječavanja identifikacije pojedinačnih transakcija, ona u praksi dodatno sužava skup jedinica za koje su službeni tržišni pokazatelji dostupni, bez obzira na to što i izračuni temeljeni na nešto većem, ali i dalje malom broju transakcija mogu biti statistički nestabilni i nepouzdana.

Sva navedena ograničenja upućuju na potrebu nadogradnje postojećeg sustava praćenja tržišta nekretnina metodološkim okvirom koji omogućuje objedinjavanje transakcijskih podataka s dodatnim izvorima informacija te primjenu statističkih i metoda strojnog učenja u svrhu dobivanja pouzdanih, konzistentnih i prostorno cjelovitih procjena tržišnih cijena stanova i najamnina. Upravo iz tih razloga u nastavku studije razvija se metodologija koja nadilazi ograničenja izravnih izračuna temeljenih isključivo na bazi eNekretnine.

Analiza podataka iz sustava eNekretnine za prvih jedanaest mjeseci 2025. godine kvantitativno potvrđuje navedena strukturna ograničenja u pogledu prostorne pokrivenosti i statističke pouzdanosti tržišnih pokazatelja. Republika Hrvatska ima ukupno 556 jedinica lokalne samouprave, no izravni izračun prosječnih i medijalnih tržišnih cijena stanova moguć je samo za one jedinice koje u promatranom razdoblju bilježe najmanje deset kupoprodajnih transakcija. Taj minimalni zakonski uvjet u 2025. godini zadovoljava 155 jedinica lokalne samouprave, što čini manje od 28 % ukupnog broja JLS. **Posljedično, za 401 jedinicu lokalne samouprave, odnosno za oko 72 % svih jedinica lokalne samouprave u Republici Hrvatskoj, ne postoji podatak o prosječnoj i/ili medijalnoj cijeni stana.**

Detaljniji uvid u raspodjelu broja kupoprodajnih transakcija dodatno ilustrira razmjere problema ograničene tržišne aktivnosti na lokalnoj razini (Tablica 1.). Od ukupno 556 jedinica lokalne samouprave, čak 414 jedinica (74,5 %) u 2025. godini ne bilježi niti jednu kupoprodajnu transakciju stana. Među jedinicama u kojima je tržišna aktivnost zabilježena, 44 jedinice ostvaruju između 10 i 19 kupoprodajnih transakcija, 24 jedinice između 20 i 29 transakcija, dok svega 9 jedinica bilježi između 30 i 39 transakcija godišnje. U rasponu od 40 do 49 transakcija nalazi se 14 jedinica lokalne samouprave, a dodatnih 16 jedinica bilježi između 50 i 99 transakcija. Tek 24 jedinice lokalne samouprave ostvaruju između 100 i 199 kupoprodajnih transakcija, dok je broj jedinica s vrlo intenzivnom tržišnom aktivnošću izrazito malen: 6 jedinica u rasponu od 200 do 499 transakcija, 3 jedinice između 500 i 999 te svega 2 jedinice s više od 1.000 kupoprodajnih transakcija u promatranom razdoblju. Ovakva distribucija jasno pokazuje da je kupoprodajno tržište stanova izrazito koncentrirano u malom broju, pretežito urbanih jedinica lokalne samouprave, dok se u velikom dijelu zemlje tržišna aktivnost uopće ne bilježi ili je svedena na zanemariv broj transakcija. Valja pritom napomenuti da je u prvih jedanaest mjeseci 2025. godine u Republici Hrvatskoj ukupno ostvareno 18.477 kupoprodajnih transakcija stanovima i apartmanima.

Čak i među jedinicama lokalne samouprave u kojima je evidentiran dovoljan broj transakcija da bi izravni izračun cijena bio metodološki dopušten, struktura raspodjele upućuje na izražene probleme statističke pouzdanosti. Velik dio tih jedinica nalazi se u rasponima s vrlo malim apsolutnim brojem transakcija, zbog čega se prosječne i medijalne cijene stanova temelje na iznimno ograničenim uzorcima. Takvi izračuni izrazito su osjetljivi na pojedinačne transakcije, heterogenost prodanih nekretnina i slučajne promjene u strukturi tržišta, te stoga ne osiguravaju stabilne i vremenski konzistentne pokazatelje lokalnih tržišnih uvjeta.

Kada je riječ o tržištu najma, prostorna pokrivenost i statistička pouzdanost administrativnih podataka još su znatno slabije nego kod kupoprodajnih transakcija (Tablica 1). U prvih jedanaest mjeseci 2025. godine tek 51 jedinica lokalne samouprave bilježi deset ili više evidentiranih ugovora o najmu, što je minimalni prag koji omogućuje izravni izračun medijalnih najamnina. Time su administrativnim podacima o najamninama obuhvaćene svega 9,2 % svih JLS u Republici Hrvatskoj, dok čak **505 lokalna jedinica (91,8 %) ne bilježi niti jedan ili bilježi manje od deset ugovora o najmu u promatranom razdoblju te stoga nema dostupan podatak o medijalnoj najamnini**. Među jedinicama u kojima je tržišna aktivnost najma evidentirana, 18 jedinica bilježi između 10 i 19 ugovora, 8 jedinica između 20 i 29 ugovora, dok svega 4 jedinice bilježe između 30 i 39 ugovora o najmu godišnje. U rasponu od 40 do 49 ugovora nalazi se samo 1 jedinica lokalne samouprave, dok 8 jedinica bilježi između 50 i 99 ugovora. Tek 3 jedinice ostvaruju između 100 i 199 ugovora o najmu, 6 jedinica između 200 i 499, 1 jedinica između 500 i 999, dok svega 2 jedinice lokalne samouprave bilježe više od 1.000 ugovora o najmu u promatranom razdoblju.

Posljedično, čak i unutar vrlo ograničenog skupa jedinica lokalne samouprave za koje je izravni izračun medijalnih najamnina uopće moguć, prevladavaju iznimno mali uzorci. U 26 od ukupno 51 jedinice lokalne samouprave s deset ili više ugovora o najmu, odnosno u više od polovice tog skupa, broj evidentiranih ugovora ne prelazi 30 godišnje. Takvi uzorci ne omogućuju stabilne ni robusne procjene lokalnih tržišnih najamnina, budući da su izrazito osjetljivi na pojedinačne ugovore, heterogenost stanova i kratkoročne promjene u strukturi tržišta. Valja napomenuti i da je u čitavoj Republici Hrvatskoj u prvih jedanaest mjeseci 2025. godine zabilježeno ukupno 2.555 ugovora o dugoročnom najmu stanova, što dodatno potvrđuje izrazitu plitkost i fragmentiranost administrativno vidljivog tržišta najma.

Ovi rezultati jasno pokazuju da su podaci o tržištu najma u administrativnim izvorima još znatno ograničeniji nego podaci o kupoprodajama, kako u pogledu prostorne pokrivenosti, tako i u pogledu statističke pouzdanosti. Valja pritom naglasiti da bi se vrlo slični obrasci pojavili i u analizi bilo koje druge godine, s obzirom na strukturnu koncentraciju tržišne aktivnosti u malom broju jedinica lokalne samouprave. Ovakva distribucija podataka

snažno opravdava potrebu za primjenom modeliranih procjena tržišnih cijena i najamnina u službene i regulatorne svrhe, osobito u kontekstu provedbe mjera stambene politike koje se izravno oslanjaju na lokalno specifične tržišne pokazatelje.

TABLICA 1. Frekvencijska raspodjela broja jedinica lokalne samouprave prema broju transakcija ostvarenih na tržištu stanova u 2025.

Raspon broja transakcija	Broj JLS prema registriranim kupoprodajnim ugovorima (broj i %)	Broj JLS prema registriranim ugovorima o najmu (broj i %)
10–19	44 (7,9 %)	18 (3,2 %)
20–29	24 (4,3 %)	8 (1,4 %)
30–39	9 (1,6 %)	4 (0,7 %)
40–49	14 (2,5 %)	1 (0,2 %)
50–99	16 (2,9 %)	8 (1,4 %)
100–199	24 (4,3 %)	3 (0,5 %)
200–499	6 (1,1 %)	6 (1,1 %)
500–999	3 (0,5 %)	1 (0,2 %)
1000 i više	2 (0,4 %)	2 (0,4 %)
Bez transakcija	414 (74,5 %)	505 (90,8 %)
Ukupno broj JLS	556 (100 %)	556 (100 %)

Izvor: izračun autorice na temelju podataka iz baze eNekretnine.

3. Metode strojnog učenja korištene u procjeni medijalnih i prosječnih cijena stanova te medijalnih najamnina

Ovo poglavlje opisuje metodološki okvir i konkretne postupke primjene metoda strojnog učenja za procjenu medijalnih i prosječnih cijena stanova u jedinicama lokalne samouprave koje nemaju zabilježene tržišne cijene u administrativnim izvorima podataka. Cilj poglavlja je detaljno prikazati kako su, na temelju jedinica s dovoljno velikim brojem transakcija, trenirani i evaluirani različiti prediktivni modeli te kako su njihove procjene iskorištene za dobivanje pouzdanih i prostorno konzistentnih procjena cijena za sve jedinice lokalne samouprave u Republici Hrvatskoj. Poglavlje započinje opisom formiranja trening i test uzorka, a zatim sustavno predstavlja korištene modele – od višestruke linearne regresije do penaliziranih regresijskih metoda, stablastih i ansambl modela te gradijentno pojačanih algoritama – uz poseban naglasak na njihovu ulogu u selekciji varijabli i evaluaciji prediktivne uspješnosti. U nastavku se razrađuje konstrukcija prostornog pomaka kao ključne varijable za hvatanje prostorne međuovisnosti tržišta nekretnina, postupak identifikacije najinformativnijeg skupa prediktora te proces odabira konačnog, parsimonioznog modela. Time se postavlja metodološka osnova za izradu službenih procjena cijena stanova u jedinicama lokalne samouprave bez izravno promatranih tržišnih podataka.

3.1. Podjela uzorka i treniranje modela

Procjena svih modela provedena je korištenjem podataka za jedinice lokalne samouprave koje u promatranoj godini bilježe dovoljan broj kupoprodajnih transakcija za izravni izračun medijalnih ili prosječnih cijena stanova. Kao ciljane varijable korištena je medijalna cijena stana s uključenim porezom na dodanu vrijednost za 2025. godinu te kasnije prosječna cijena stana bez poreza na dodanu vrijednost za 2025. godinu. U postupak modeliranja uključene su isključivo one jedinice lokalne samouprave koje bilježe najmanje 30 kupoprodajnih transakcija, čime se osigurava da su izravno izračunate medijalne cijene dovoljno stabilne i reprezentativne da mogu poslužiti kao pouzdana referentna vrijednost u postupku treniranja i evaluacije modela.

Na temelju tog kriterija formiran je analitički uzorak jedinica lokalne samouprave s poznatom ciljnom varijablom, koji je zatim nasumično podijeljen na trening i test (holdout) uzorak u omjeru 80 % prema 20 %. Podjela uzorka provedena je korištenjem funkcije `initial_split` iz paketa `rsample`, uz postavljanje sjemena generatora slučajnih brojeva radi osiguravanja potpune replikabilnosti rezultata. Trening uzorak korišten je

za procjenu parametara modela i optimizaciju hiperparametara, dok je test uzorak u cijelosti rezerviran za evaluaciju prediktivne uspješnosti. Podjela je provedena na razini jedinica lokalne samouprave, a ne pojedinačnih transakcija, čime je osigurano da se u fazi evaluacije procjenjuje sposobnost modela da generalizira na jedinice koje uopće nisu sudjelovale u procesu treniranja. Takav pristup predstavlja standard u primjeni metoda strojnog učenja te omogućuje realističnu procjenu prediktivne moći modela uz minimalan rizik preučenja (eng. overfitting).

U početni model u kojem je ciljana varijabla bila ili medijalna cijena stana s PDV-om ili prosječna cijena stana bez PDV-a, bilo je uključeno ukupno 20 objašnjavajućih varijabli koje su detaljno opisane u sljedećem poglavlju Studije, a koje obuhvaćaju demografska, stambena, infrastrukturna, prostorna i ekonomska obilježja jedinica lokalne samouprave, kao i konstruirani prostorni pomak ciljne varijable. Detaljan opis korištenih varijabli i njihova interpretacija dani su u sljedećem dijelu studije, nakon metodološkog prikaza primijenjenih modela. Taj početni model s 20 objašnjavajućih varijabli s medijalnom cijenom stanova kao ciljanom (zavisnom) varijablom, te kasnije s prosječnom cijenom stanova kao ciljanom varijablom, zatim je treniran na uzorku izabranom za trening primjenom metoda opisanih u nastavku teksta.

3.2. Modeli korišteni u strojnom učenju

Višestruka linearna regresija

Višestruka linearna regresija korištena je kao početni, referentni model zbog svoje jednostavnosti, transparentnosti i interpretabilnosti. Model pretpostavlja linearnu vezu između ciljne varijable (medijalne ili prosječne cijene stana) i skupa objašnjavajućih varijabli koje opisuju demografska, ekonomska, infrastrukturna i prostorna obilježja jedinica lokalne samouprave.

U procjeni modela korištena je metoda najmanjih kvadrata (OLS), bez dodatne regularizacije. Linearni model ne uključuje hiperparametre u užem smislu, no njegova je uloga u analizi dvostruka: prvo, služi kao osnovna usporedna točka za procjenu dobitaka u prediktivnoj točnosti koje nude složeniji modeli; drugo, omogućuje osnovni uvid u smjer i relativnu važnost pojedinih varijabli, što je korisno za ekonomsku interpretaciju rezultata.

Penalizirane regresijske metode: Ridge, LASSO i Elastic Net

Penalizirane regresijske metode korištene su kako bi se adresirali problemi multikolinearnosti i potencijalnog preučenja koji proizlaze iz relativno velikog broja demografskih, ekonomskih, infrastrukturnih i prostornih obilježja jedinica lokalne

samouprave. Sve tri metode zadržavaju linearnu strukturu odnosa između ciljne varijable i prediktora, ali uvode penalizaciju veličine regresijskih koeficijenata.

Ridge regresija smanjuje varijancu procjena dodavanjem L2 penalizacije na sumu kvadrata koeficijenata (Hoerl i Kennard, 1970). Ovaj pristup ne provodi selekciju varijabli, već smanjuje utjecaj onih varijabli koje doprinose nestabilnosti procjena. U implementaciji je korišten parametar $\alpha = 0$, dok je ključni hiperparametar penalizacije λ određen putem 10-struke unakrsne validacije, pri čemu je kao konačna vrijednost korištena ona koja minimizira pogrešku predikcije na trening uzorku ($\lambda_{\min}$).

LASSO regresija koristi L1 penalizaciju, koja uz regularizaciju omogućuje i eksplicitnu selekciju varijabli postavljanjem dijela koeficijenata točno na nulu (Tibshirani, 1996). Time se dobiva parsimonijski model, što je osobito važno u kontekstu analize determinanti cijena nekretnina i interpretabilnosti rezultata za potrebe javnih politika. U ovoj analizi LASSO regresija implementirana je s $\alpha = 1$, dok je optimalna vrijednost λ ponovno određena putem 10-struke unakrsne validacije.

Elastic Net regresija kombinira L1 i L2 penalizaciju te je osobito prikladna u situacijama u kojima postoji skup snažno koreliranih objašnjavajućih varijabli (Zou i Hastie, 2005), što je čest slučaj kod prostorno i socioekonomski uvjetovanih pokazatelja. U implementaciji je korišten $\alpha = 0.5$, čime su LASSO i Ridge komponenta ponderirane u jednakim omjerima, dok je parametar λ odabran empirijski na temelju minimizacije pogreške predikcije.

Regresijska stabla (Regression Tree)

Regresijska stabla predstavljaju neparametarski, nelinearni pristup modeliranju odnosa između obilježja jedinica lokalne samouprave i tržišnih cijena stanova (Breiman et al., 1984). Model funkcionira tako da podatkovni prostor dijeli u hijerarhijski strukturirane segmente, pri čemu se u svakoj podjeli bira ona varijabla i prag koji minimiziraju varijancu ciljne varijable unutar podskupova.

Prednost regresijskih stabala leži u njihovoj sposobnosti hvatanja nelinearnih odnosa i interakcija među varijablama bez potrebe za njihovim eksplicitnim specficiranjem. Međutim, pojedinačna stabla sklona su visokoj varijanci i preučanju. U ovoj analizi složenost stabala ograničena je korištenjem parametra $cp = 0.001$, čime je dopuštena veća dubina stabla uz zadržavanje kontrole nad preučanjem. Regresijska stabla korištena su prvenstveno kao interpretativni alat i kao temelj za složenije ansambl metode.

Slučajna šuma (Random Forest)

Metoda slučajnih šuma nadograđuje regresijska stabla kombiniranjem velikog broja stabala u ansambl (Breiman, 2001). Svako stablo trenira se na slučajnom poduzorku trening podataka, dok se pri svakoj podjeli razmatra samo slučajni podskup objašnjavajućih varijabli. Takav postupak smanjuje korelaciju među stablima i značajno smanjuje varijancu ukupne predikcije.

U ovoj analizi Random Forest modeli trenirani su korištenjem 1000 stabala, čime je osigurana stabilnost predikcija i konvergencija ansambla. Kao mjera važnosti varijabli korištena je mjera nečistoće (impurity), a kategorijske varijable tretirane su uz uvažavanje njihova poretka. Random Forest se pokazao posebno robusnim u prisutnosti nelinearnosti, šuma i ekstremnih vrijednosti u podacima.

Gradijentno pojačani modeli (XGBoost)

Gradijentno pojačani modeli predstavljaju najsloženiji pristup korišten u ovoj studiji. XGBoost algoritam iterativno gradi niz slabih regresijskih stabala, pri čemu svaki novi model uči na pogreškama prethodnih predikcija (Friedman, 2001; Chen i Guestrin, 2016). Time se postupno poboljšava ukupna prediktivna sposobnost modela.

Model je treniran s ciljem minimizacije kvadratne pogreške, uz sljedeće hiperparametre: $\eta = 0.05$, $\text{max_depth} = 5$, $\text{subsample} = 0.8$ i $\text{colsample_bytree} = 0.8$. Maksimalni broj iteracija postavljen je na 800, uz primjenu ranog zaustavljanja nakon 40 iteracija bez poboljšanja pogreške na testnom uzorku, čime se učinkovito sprječava preučenje.

Prostorni pomak i metoda K najbližih susjeda (KNN)

Važna metodološka inovacija ove studije je uvođenje prostornog pomaka kao dodatne objašnjavajuće varijable. Prostorni pomak odražava činjenicu da tržišta nekretnina nisu prostorno izolirana, već da se cijene stanova u jednoj jedinici lokalne samouprave u velikoj mjeri formiraju u odnosu na cijene u prostorno bliskim područjima.

Prostorni pomak konstruiran je u dva koraka. U prvom koraku, za sve jedinice lokalne samouprave izgrađena je mreža $K = 10$ najbližih susjeda na temelju geografskih koordinata centroida, pri čemu je prostorni pomak definiran kao aritmetička sredina ciljne varijable u susjednim jedinicama. U slučajevima u kojima ciljana varijabla nije dostupna u svim susjednim jedinicama, prostorni pomak ostaje neizračunat. U drugom koraku, kako bi se osigurala potpuna prostorna pokrivenost, nedostajuće vrijednosti prostornog pomaka imputirane su primjenom KNN regresije s do 50 najbližih opažanja s poznatom vrijednošću prostornog pomaka. Time je osigurano da prostorni pomak postoji za sve jedinice lokalne samouprave te da je informiran prostorno relevantnim informacijama. Činjenica da je

bilo potrebno i do 50 najbližih opažanja s poznatom vrijednošću prostornog pomaka da izračunamo prostorne pomake za sve jedinice lokalne samouprave u zemlji, najbolje govori o tome koliko je velik problem s nedostatnim podacima o cijenama stanova u zemlji. Naime, za pojedine jedinice lokalne samouprave tek 50. najbliža jedinica lokalne samouprave ima zabilježeni službeni podataka o medijalnoj i prosječnoj cijeni stana.

3.3. Evaluacija modela

Uspješnost svih modela evaluirana je na testnom uzorku korištenjem standardnih mjera prediktivne točnosti. Kao primarne mjere korištene su srednja apsolutna pogreška (MAE) i korijen srednje kvadratne pogreške (RMSE), dok su dodatno izračunati koeficijent determinacije (R^2) i prilagođeni koeficijent determinacije, pri čemu je broj prediktora eksplicitno uzet u obzir. Konačni odabir modela za službene procjene temeljen je na kombinaciji prediktivne točnosti, stabilnosti rezultata i interpretabilnosti, uz poseban naglasak na robusnost modela u jedinicama lokalne samouprave s malim brojem ili potpunim izostankom transakcija.

3.4. Izbor optimalnih objašnjavajućih varijabli za konačni prediktivni model

Kao što je već naglašeno, u osnovni model bilo je uključeno ukupno 20 objašnjavajućih varijabli i metode strojnog učenja su prvo procijenjene korištenjem tog modela. Nakon procjene tog osnovnog modela, sve metode osim višestruke linearne regresije iskorištene su za sustavnu identifikaciju varijabli koje su ključne za postizanje dobre prediktivne sposobnosti modela, s posebnim naglaskom na stabilnost, robusnost i generalizabilnost procjena medijalnih, a kasnije i prosječnih cijena stanova.

Kod penaliziranih regresijskih metoda (LASSO, Ridge i Elastic Net) važnost varijabli procijenjena je na temelju apsolutne vrijednosti procijenjenih regresijskih koeficijenata pri optimalnoj vrijednosti parametra penalizacije λ , koja je određena putem unakrsne validacije. U tom kontekstu, LASSO regresija ima posebno važnu ulogu jer, za razliku od Ridge regresije, omogućuje eksplicitnu selekciju varijabli tako što dio koeficijenata postavlja točno na nulu, čime se jasno identificiraju varijable koje ne doprinose prediktivnoj sposobnosti modela. Ridge regresija, iako ne eliminira varijable, pruža informaciju o relativnoj važnosti prediktora kroz stupanj njihova „skupljanja“, dok Elastic Net kombinira oba pristupa te je osobito informativan u prisutnosti skupina snažno koreliranih varijabli, jer omogućuje stabilniju raspodjelu važnosti unutar takvih skupina.

U slučaju regresijskih stabala, važnost varijabli nije izražena eksplicitnom numeričkom metrikom, već se implicitno očituje kroz učestalost korištenja varijabli u čvorovima

stabla, njihovu poziciju u hijerarhiji podjela te njihov doprinos smanjenju varijance ciljne varijable, pri čemu varijable koje se pojavljuju bliže korijenu stabla i češće u podjelama imaju veći prediktivni značaj.

Kod metode slučajnih šuma korištene su mjere važnosti temeljene na smanjenju nečistoće (impurity-based importance), koje kvantificiraju koliko pojedina varijabla u prosjeku doprinosi smanjenju pogreške predikcije kroz veliki broj stabala u ansamblu, čime se dobiva robusna procjena važnosti varijabli u nelinearnom i interakcijskom okruženju.

Kod XGBoost modela važnost varijabli procijenjena je pomoću mjere dobitka (gain), koja pokazuje koliki relativni doprinos pojedina varijabla ima u smanjenju funkcije gubitka kroz sve iteracije gradijentnog pojačavanja, pri čemu se veći gain interpretira kao veći marginalni doprinos poboljšanju predikcije.

Dodatno, kako bi se izbjegla ovisnost selekcije varijabli o specifičnostima pojedinih algoritama, za Random Forest model primijenjena je i permutacijska važnost varijabli (permutation importance), koja predstavlja model-agnostički pristup procjeni prediktivne važnosti. Ova metoda temelji se na mjerenju porasta pogreške predikcije nakon nasumične permutacije vrijednosti pojedine varijable, pri čemu se sve ostale varijable zadržavaju nepromijenjene; veći porast pogreške ukazuje na veću stvarnu važnost varijable za prediktivnu sposobnost modela. Upravo zbog svoje neovisnosti o unutarnjoj strukturi modela, permutacijska važnost korištena je kao ključni korektiv u odnosu na model-specifične metrike.

Kombiniranjem rezultata svih navedenih pristupa (penaliziranih regresija, stablastih modela, ansambl metoda i model-agnostičke permutacijske važnosti) dobiven je **konsenzusni rang važnosti varijabli**, koji je poslužio kao temelj za identifikaciju manjeg, informacijski najbogatijeg skupa varijabli te za izgradnju parsimonioznijih modela s usporedivom, a u pojedinim slučajevima i boljom prediktivnom sposobnošću u odnosu na puni model.

3.5. Izbor zaključnog modela i metode za procjenu cijene stanova

U nastavku metodološkog postupka, nakon što je na temelju konsenzusnog ranga važnosti identificirano **pet najinformativnijih varijabli** iz početnog skupa od dvadeset, provedena je dodatna faza modeliranja usmjerena na sustavno ispitivanje **parsimonioznih specifikacija modela**. Cilj ove faze nije bio samo dodatno poboljšati prediktivnu točnost, već i razumjeti u kojoj je mjeri moguće postići usporedive ili bolje rezultate uz manji broj objašnjavajućih varijabli, što je osobito važno u kontekstu stabilnosti procjena i njihove primjenjivosti u javnim politikama. Iz odabranih pet varijabli generirani su svi mogući podskupovi veličine dvije, tri, četiri i pet varijabli, čime je dobiveno ukupno

26 različitih kombinacija prediktora. Za svaki od tih podskupova procijenjen je puni skup od **sedam modela** – višestruka linearna regresija, LASSO, Ridge, Elastic Net, regresijsko stablo, Random Forest i XGBoost – pri čemu su za svaku kombinaciju varijabli i za svaki algoritam izračunate standardne mjere uspješnosti (MAE, RMSE, R^2 i prilagođeni R^2) na testnom uzorku.

Ovakav iscrpan “grid” pristup omogućio je izravnu usporedbu ne samo različitih algoritama, već i različitih **razina složenosti modela**, pri čemu se moglo precizno identificirati kako se prediktivna sposobnost mijenja s dodavanjem ili uklanjanjem pojedinih varijabli. Posebna pažnja posvećena je robusnosti penaliziranih regresijskih metoda, pri čemu su za svaki podskup varijabli parametri regularizacije ponovno određivani unakrsnom validacijom, dok su kod nelinearnih modela procjene temeljene na standardnim postavkama koje osiguravaju stabilne rezultate i sprječavaju preučenje.

Nakon što su rezultati objedinjeni u jedinstvenu bazu, za svaki algoritam identificirane su **najbolje specifikacije** prema tri kriterija: minimalnoj srednjoj apsolutnoj pogrešci, minimalnoj RMSE i maksimalnom prilagođenom koeficijentu determinacije. Time je omogućena jasna identifikacija optimalnih kombinacija varijabli i modela, ali i usporedba između algoritama u pogledu njihove osjetljivosti na redukciju broja prediktora. Završno, na temelju tih rezultata odabran je **zaključni model s malim brojem varijabli**, koji je ponovno procijenjen svim relevantnim metodama te korišten za izradu konačnih predikcija za sve jedinice lokalne samouprave, uključujući i one bez izravno promatranih tržišnih cijena. Ovim postupkom osigurano je da konačne procjene ne proizlaze iz jednog arbitrarnog izbora modela, već iz sustavne, višekriterijske evaluacije velikog broja mogućih specifikacija, čime se dodatno povećava pouzdanost i transparentnost cjelokupne metodologije.

4. Podaci korišteni u procjeni medijalnih i prosječnih cijena stanova te medijalnih najamnina primjenom metoda strojnog učenja

Analiza i modeliranje tržišnih cijena stanova i najamnina na razini jedinica lokalne samouprave temelje se na objedinjavanju administrativnih, demografskih, prostornih, infrastrukturnih i ekonomskih podataka iz više službenih izvora. Jedinica opažanja u svim modelima je jedinica lokalne samouprave (JLS), dok su objašnjavajuće varijable izabrane tako da odražavaju strukturne karakteristike lokalnih tržišta nekretnina koje teorijski i empirijski utječu na razinu cijena stanova. U osnovni prediktivni model uključeno je ukupno 20 varijabli, koje se mogu grupirati u pet tematskih skupina: (i) demografska obilježja, (ii) stambene karakteristike, (iii) prostorne i geografske značajke, (iv) infrastrukturna dostupnost te (v) ekonomski i tržišni pokazatelji, uključujući konstruirani prostorni pomak ciljne varijable. U nastavku se detaljno opisuje svaka od korištenih varijabli, pri čemu se u zagradama navodi naziv varijable korišten u modeliranju.

Postupak procjene proveden je kroz dva odvojena, ali metodološki identična ciklusa strojnog učenja opisana u prethodnom poglavlju, pri čemu je svaki ciklus započeo s istim početnim modelom koji uključuje identičan skup od 20 objašnjavajućih varijabli. U prvom ciklusu **ciljana varijabla bila je medijalna cijena stanova s uključenim porezom na dodanu vrijednost** (za onaj dio transakcija koje predstavljaju novogradnju) u jedinici lokalne samouprave u 2025. godini. Na temelju tog ciklusa izrađene su predikcije medijalnih cijena stanova za sve jedinice lokalne samouprave, uključujući i one koje ne bilježe dovoljan broj transakcija za izravni izračun. Dobivene procjene medijalnih cijena stanova zatim su korištene kao temelj za procjenu medijalnih najamnina u svim jedinicama lokalne samouprave, što predstavlja ključni ulazni podatak za operativnu provedbu Programa priuštivog najma.

U drugom ciklusu strojnog učenja korišten je isti početni model s istih dvadeset objašnjavajućih varijabli, ali s drugačije definiranom ciljanom varijablom. **Ciljana varijabla u ovom ciklusu bila je prosječna cijena stanova bez uključenog poreza na dodanu vrijednost** u 2025. godini. Kod izračuna ovog prosjeka PDV je isključen u svim slučajevima u kojima je uz transakciju u bazi eNekretnine postojala oznaka da je PDV bio uključen u izračun prosječne cijene. Ova definicija ciljne varijable odabrana je na izričit zahtjev naručitelja studije, Ministarstva prostornoga uređenja, graditeljstva i državne imovine, budući da prosječna cijena stana bez PDV-a predstavlja osnovicu za utvrđivanje opravdanosti zahtjeva za povrat dijela poreza na dodanu vrijednost pri kupnji prve stambene nekretnine u okviru Programa potpora za stjecanje prve stambene nekretnine.

Važno je naglasiti da definicije i medijalne cijene stanova s PDV-om i prosječne cijene stanova bez PDV-a u oba ciklusa obuhvaćaju i novogradnju i starogradnju. Za one jedinice lokalne samouprave u kojima postoji dovoljan broj kupoprodajnih transakcija, obje ciljne varijable izračunate su izravno iz pojedinačnih podataka o kupoprodajnim transakcijama dostupnih u informacijskom sustavu eNekretnine. U izračun medijalnih i prosječnih cijena nisu uključene cijene obiteljskih kuća, niti su one zasebno razmatrane u modeliranju, zbog činjenice da za obiteljske kuće u bazi eNekretnine u pravilu nije dostupan pouzdan podatak o korisnoj površini, što onemogućuje izračun usporedivih i točnih cijena po četvornom metru korisne površine. Također, kako bi se smanjila kolebljivost i povećala pouzdanost ciljane varijable, prilikom strojnog učenja u obzir su uzimane samo one vrijednosti ciljanih varijabli koje su nastale na temelju minimalno 30 dostupnih individualnih transakcijskih podataka.

Kao što je već navedeno, u osnovni prediktivni model koji je podvrgnut strojnom učenju u svrhu definiranja zaključnog prediktivnog modela medijalnih (a kasnije i prosječnih) cijena stanova uključeno je ukupno 20 varijabli, koje se mogu grupirati u pet tematskih skupina: (i) demografska obilježja, (ii) stambene karakteristike, (iii) prostorne i geografske značajke, (iv) infrastrukturna dostupnost te (v) ekonomski i tržišni pokazatelji, uključujući konstruirani prostorni pomak ciljne varijable. Za razliku od vrijednosti ciljanih varijabli (cijena stanova), vrijednosti svih dvadeset varijabli dostupne su za sve 556 jedinica lokalne samouprave. Upravo ta karakteristika skupa objašnjavajućih varijabli omogućuje nam da pomoću modela strojnog učenja procijenimo cijene stanova u jedinicama lokalne samouprave koje ne bilježe taj pokazatelj. U nastavku se detaljno opisuje svaka od korištenih varijabli te se u zgradama navodi ime opisane varijable korišteno u modeliranju.

Demografska obilježja

Stopa totalne fertiliteta (tfr)

Stopa totalne fertiliteta mjeri prosječan broj djece koje bi žena rodila tijekom reproduktivnog razdoblja pod pretpostavkom nepromijenjenih dobno-specifičnih stopa fertilitnosti. Ova varijabla služi kao indikator demografske vitalnosti lokalne zajednice i dugoročnog potencijala potražnje za stanovanjem. Više vrijednosti TFR-a upućuju na povoljniju demografsku strukturu i potencijalno stabilniju ili rastuću potražnju za stambenim prostorom. Izvor podataka je Ekonomski institut, Zagreb (2024).

Indeks vitalnosti (vitindex)

Indeks vitalnosti predstavlja složeni demografski pokazatelj koji kombinira informacije o prirodnom kretanju stanovništva i dobnoj strukturi. Varijabla odražava opću demografsku

održivost jedinice lokalne samouprave te je snažno povezana s dugoročnim trendovima potražnje za stanovanjem i atraktivnošću lokacije za mlade i radno aktivno stanovništvo. Izvor podataka je Državni zavod za statistiku.

Stambene karakteristike

Kumulativni broj dovršenih stanova 2012.–2022. (comapart_kum12_22)

Ova varijabla mjeri ukupnu stambenu izgradnju u jedinici lokalne samouprave u prethodnom desetljeću te predstavlja indikator ponude novih stanova. Veća vrijednost upućuje na intenzivniju građevinsku aktivnost i potencijalno veći pritisak na tržište nekretnina, osobito u rastućim urbanim i turističkim sredinama. Izvor podataka je Državni zavod za statistiku.

Udio stalno nastanjenih stanova (permunocushare)

Udio stanova koji se koriste za stalno stanovanje u ukupnom stambenom fondu. Niže vrijednosti ove varijable često su karakteristične za turističke i sekundarne rezidencijalne destinacije, gdje velik dio stanova služi za povremeni boravak ili kratkoročni najam, što može povećavati cijene i smanjivati dostupnost stanovanja lokalnom stanovništvu. Izvor podataka je Popis stanovništva u 2021. godini Državnog zavoda za statistiku.

Udio praznih stanova (vacapart)

Ova varijabla mjeri udio stanova koji se ne koriste ni za stalno stanovanje ni za registrirani najam. Visok udio praznih stanova može upućivati na strukturne neravnoteže na tržištu nekretnina, spekulativnu potražnju ili demografski pad, ovisno o lokalnom kontekstu. Izvor podataka je Popis stanovništva u 2021. godini Državnog zavoda za statistiku.

Gustoća kućanstava (densityhousehold)

Gustoća kućanstava po jedinici površine koristi se kao indikator urbanizacije i intenziteta korištenja prostora. Veće vrijednosti tipične su za urbana područja i povezane su s višim razinama cijena stanova. Izvor podataka je Popis stanovništva u 2021. godini Državnog zavoda za statistiku, a izvor podataka za površinu jedinica lokalne samouprave je također Državni zavod za statistiku.

Prostorne i geografske značajke

Duljina obalne linije (coast_length)

Ova varijabla mjeri duljinu obalne linije unutar granica jedinice lokalne samouprave. Predstavlja ključni pokazatelj turističke i rezidencijalne atraktivnosti, osobito u priobalnim područjima, gdje prisutnost mora snažno utječe na razinu cijena nekretnina. Izvor podataka je Europska okolišna agencija, odnosno EU-Hydro: Copernicus Land Monitoring Service.

Udaljenost od mora (dist_sea)

Minimalna udaljenost centroida jedinice lokalne samouprave od morske obale. Manja udaljenost u pravilu je povezana s višim cijenama stanova zbog turističke potražnje i veće atraktivnosti lokacije. Izvor podataka je Europska okolišna agencija, odnosno EU-Hydro: Copernicus Land Monitoring Service.

Udio urbanih površina (urban_perc)

Udio zemljišta klasificiranog kao urbano. Ova varijabla izravno odražava razinu urbanizacije i intenzitet izgrađenosti prostora. Izvor podataka je GIS izračun temeljen na CORINE Land Cover podacima Europske okolišne agencije.

Udio šumskih površina (forest_perc)

Udio površine pod šumama koristi se kao indikator prirodnih obilježja i potencijalne atraktivnosti za stanovanje, ali i kao proxy za ograničenja daljnje urbanizacije. Izvor podataka je GIS izračun temeljen na CORINE Land Cover podacima Europske okolišne agencije.

Udio vodnih površina (water_perc)

Uključuje rijeke, jezera i ostale vodne površine te može imati dvosmjernan učinak na cijene, ovisno o lokalnom kontekstu (atraktivnost nasuprot ograničenjima gradnje). Izvor podataka je GIS izračun temeljen na CORINE Land Cover podacima Europske okolišne agencije.

Otočni status (island)

Binarna varijabla koja označava nalazi li se jedinica lokalne samouprave na otoku. Varijabla poprima vrijednost 1 za sve lokalne jedinice koje se nalaze na otocima, te 0 za sve lokalne jedinice koje se ne nalaze na otocima. Otočne JLS često imaju specifične tržišne dinamike, višu turističku potražnju i ograničenu ponudu prostora.

Metropolitanski status (metro)

Binarni indikator pripadnosti metropolitanskom području. Varijabla poprima vrijednost 1 za Zagreb, Split, Rijeku i Osijek, te 0 za sve ostale lokalne jedinice. Metropolitanske JLS karakterizira veća potražnja, viši dohoci i razvijenija infrastruktura, što se u pravilu odražava u višim cijenama stanova.

Infrastrukturna dostupnost

Udaljenost od zračne luke (dist_airport)

Ova varijabla mjeri prometnu dostupnost putem zračnog prijevoza. Kraće udaljenosti povećavaju atraktivnost lokacije za turizam, poslovne aktivnosti i međunarodnu mobilnost, što se često reflektira u višim cijenama nekretnina. Izvor podataka je GIS izračun temeljen na Open Street Maps podacima.

Udaljenost od marine (dist_marina)

Indikator dostupnosti nautičke infrastrukture, osobito relevantan za obalne i otočne JLS. Prisutnost marina povezana je s luksuznim segmentom tržišta nekretnina. Izvor podataka je GIS izračun temeljen na Open Street Maps podacima.

Udaljenost od autoceste (dist_highway)

Mjeri cestovnu dostupnost i povezanost s regionalnim i nacionalnim prometnim mrežama. Bolja prometna povezanost smanjuje troškove mobilnosti i povećava tržišnu atraktivnost lokacije. Izvor podataka je GIS izračun temeljen na Open Street Maps podacima.

Ukupna duljina cesta (length_roads)

Predstavlja mjeru razvijenosti lokalne cestovne infrastrukture i prostorne dostupnosti unutar jedinice lokalne samouprave. Izvor podataka je GIS izračun temeljen na Open Street Maps podacima.

Ekonomski i tržišni pokazatelji

Indeks razvijenosti (ir)

Složeni pokazatelj socioekonomske razvijenosti jedinice lokalne samouprave koji obuhvaća dohodovne, fiskalne i razvojne komponente. Više vrijednosti indeksa povezane su s većom platežnom moći i višim cijenama nekretnina. Izvor podataka je Ministarstvo regionalnog razvoja i fondova Europske Unije Republike Hrvatske (2024).

Indeks turističke razvijenosti (itr)

Ova varijabla mjeri intenzitet turističke aktivnosti i važnost turizma u lokalnom gospodarstvu. Turistički snažne jedinice često bilježe veće cijene stanova i najamnina zbog dodatne potražnje za sekundarnim stanovanjem i kratkoročnim najmom. Izvor podataka je Institut za turizam (2025).

Prostorni pomak ciljne varijable

Prostorni pomak medijalne cijene stana (lag_target)

Prostorni pomak konstruiran je kao prosječna vrijednost ciljne varijable u prostorno najbližim jedinicama lokalne samouprave. Ova varijabla eksplicitno uvodi prostornu međuovisnost u model, odražavajući činjenicu da se cijene stanova u jednoj JLS formiraju u odnosu na cijene u okolnim područjima. Prostorni pomak značajno poboljšava prediktivnu sposobnost modela, osobito u jedinicama s malim brojem ili potpunim izostankom transakcija. Prostorni pomak izračunat je primjenom metode K najbližih susjeda (KNN) na temelju geografskih koordinata centroida svake jedinice lokalne samouprave, pri čemu je za svaku jedinicu definiran kao prosječna vrijednost ciljne varijable u prostorno najbližim jedinicama. Izvor podataka je izračun autorice za potrebe ove studije.

Osim navedenih podataka, za potrebe treniranja, testiranja i evaluacije modela korišteni su još i podaci o **razvojnim grupama kojima pripadaju jedinice lokalne samouprave s obzirom na ostvarenu vrijednost indeksa razvijenosti**. Ta varijabla naziva se **gir** i može poprimiti vrijednosti od 1 (što označava najnerazvijeniju skupinu) do 8 (što označava najrazvijeniju skupinu). Izvor podataka za tu varijablu je Ministarstvo regionalnog razvoja i fondova Europske Unije Republike Hrvatske (2024). Kao što je već naglašeno, za potrebe procjene KNN modela te za potrebe izrade prostornih karata, korištene su **geografske koordinate centroida jedinice lokalne samouprave** Državne geodetske uprave.

5. Procjena medijalnih cijena stanova za jedinice lokalne samouprave primjenom metoda strojnog učenja

U ovom poglavlju provodi se strojno učenje modela medijalnih cijena stanova s PDV-om za sve jedinice lokalne samouprave u Republici Hrvatskoj s ciljem dobivanja prostorno konzistentnih i metodološki utemeljenih predikcija medijalnih cijena stanova koje se u sljedećem poglavlju koriste kao ulazna informacija za izračun medijalnih najamnina. Procjene medijalnih cijena stanova te zatim i medijalnih najamnina predstavljaju ključnu analitičku osnovu za provedbu **Programa priuštivog najma**, koji provodi Ministarstvo prostornoga uređenja, graditeljstva i državne imovine, budući da one omogućuju određivanje referentnih razina najamnina i cijena stanovanja i u jedinicama lokalne samouprave u kojima ne postoji dovoljan broj tržišnih transakcija. Analiza je strukturirana kao višefazni postupak koji započinje konstrukcijom prostornog pomaka ciljane varijable radi eksplicitnog uvođenja prostorne međuovisnosti, nastavlja se definiranjem uzorka za treniranje modela i testiranje, sustavnom usporedbom različitih modela strojnog učenja radi identifikacije ključnih objašnjavajućih varijabli, te završava izgradnjom parsimonioznog i interpretabilnog zaključnog modela prikladnog za operativnu primjenu u kontekstu primjene Programa priuštivog najma.

5.1. Konstrukcija prostornog pomaka ciljane varijable

Prostorni pomak ciljane varijable konstruiran je kao objašnjavajuća varijabla koja eksplicitno uvodi prostornu međuovisnost u modele procjene cijena stanova. Polazišna ideja je da se cijene stanova u jednoj jedinici lokalne samouprave ne formiraju izolirano, već su snažno povezane s cijenama u prostorno bliskim jedinicama koje čine isto ili susjedno lokalno tržište nekretnina.

U prvom koraku prostorni pomak izračunat je kao aritmetička sredina vrijednosti ciljane varijable u deset prostorno najbližih jedinica lokalne samouprave. Prostorna bliskost definirana je isključivo na temelju geografskih koordinata centroida svake jedinice lokalne samouprave. Za svaku JLS izračunate su euklidske udaljenosti do svih ostalih jedinica, nakon čega je identificirano deset jedinica s najmanjom udaljenošću, neovisno o administrativnim granicama ili pripadnosti županiji.

Važno je naglasiti da u ovom koraku nije bilo nužno da svih deset najbližih jedinica imaju dostupne podatke o cijenama stanova. Aritmetička sredina izračunata je koristeći isključivo one susjedne jedinice među deset najbližih za koje je ciljane varijable

(medijalna cijena stana) bila dostupna. Drugim riječima, prostorni pomak je u ovom koraku definiran kao:

- prosjek poznatih cijena u prostorno najbližem okruženju,
- uz ignoriranje onih susjeda koji nemaju zabilježene cijene stanova.

Ako barem jedna od deset najbližih jedinica ima poznatu vrijednost ciljane varijable, prostorni pomak za promatranu jedinicu može se izračunati. Međutim, u slučajevima gdje nijedna od deset najbližih jedinica nema dostupne podatke o cijenama, prostorni pomak ostaje neizračunat. Upravo se to dogodilo kod 75 jedinica lokalne samouprave koje se nalaze u prostorno i tržišno izrazito neatraktivnim područjima.

Ovakav konzervativan pristup u prvom koraku namjerno je odabran kako se ne bi umjetno konstruirala prostorna informacija tamo gdje ona lokalno ne postoji, odnosno kako bi se prostorni pomak temeljio isključivo na stvarno opaženim tržišnim cijenama u neposrednom okruženju.

Nakon prvog koraka utvrđeno je da za 75 jedinica lokalne samouprave prostorni pomak nije bilo moguće izravno izračunati, jer niti jedna od njihovih deset najbližih susjednih jedinica nije imala poznatu ciljanu varijablu. Riječ je o oko 13,5 % svih lokalnih jedinica, koje se u pravilu nalaze u područjima s vrlo slabom tržišnom aktivnošću i izrazitim nedostatkom transakcijskih podataka.

Kako bi se osigurala potpuna prostorna pokrivenost i omogućilo korištenje prostornog pomaka kao objašnjavajuće varijable u svim modelima strojnog učenja, u drugom je koraku primijenjena KNN regresija (K-Nearest Neighbors).

U ovom koraku KNN metoda nije korištena za procjenu cijena stanova, već isključivo za imputaciju nedostajućih vrijednosti prostornog pomaka. Postupak je proveden na sljedeći način:

1. Kao ulazne varijable korištene su geografske koordinate centroida svih jedinica lokalne samouprave.
2. Skup opažanja podijeljen je na:
 - jedinice s poznatom vrijednošću prostornog pomaka (iz prvog koraka),
 - jedinice bez izračunatog prostornog pomaka.
3. Za svaku jedinicu bez prostornog pomaka identificiran je skup prostorno najbližih jedinica s poznatom vrijednošću prostornog pomaka, ponovno na temelju euklidske udaljenosti centroida.
4. Vrijednost prostornog pomaka imputirana je kao aritmetička sredina vrijednosti prostornog pomaka u K najbližih opažanja.

Ključno metodološko pitanje u ovom koraku bilo je koliko susjeda (K) koristiti. Za razliku od prvog koraka, gdje je $K = 10$ odabrano kako bi se zadržala stroga lokalna interpretacija prostorne bliskosti, u drugom koraku cilj nije bio zadržati isključivo vrlo lokalnu informaciju, već osigurati da svaka jedinica dobije stabilnu i informativnu procjenu prostornog konteksta.

Zbog toga je K postavljen na do 20 najbližih jedinica s poznatom vrijednošću prostornog pomaka, pri čemu se u praksi koristio maksimalni mogući broj dostupnih opažanja (ako ih je bilo manje od 20). Ovakav izbor ima nekoliko važnih implikacija:

- omogućuje imputaciju prostornog pomaka i u izrazito izoliranim područjima,
- smanjuje osjetljivost imputacije na pojedinačne ekstremne vrijednosti,
- osigurava da imputirani prostorni pomak odražava širi regionalni tržišni kontekst, a ne slučajnu ili lokalno specifičnu vrijednost.

Činjenica da je u pojedinim slučajevima bilo potrebno uključiti i do 20 prostorno najbližih jedinica kako bi se imputirao prostorni pomak jasno ilustrira razmjere problema nedostatka podataka o cijenama stanova u Republici Hrvatskoj i snažno potvrđuje opravdanost primjene modeliranih procjena i prostorno informiranih metoda strojnog učenja.

5.2. Formiranje analitičkog uzorka i podjela na trening i test skup

Proces treniranja i evaluacije modela strojnog učenja započeo je definiranjem analitičkog uzorka jedinica lokalne samouprave za koje je moguće konstruirati statistički dovoljno pouzdanu ciljnu varijablu. Kao što je ranije obrazloženo, u ovoj studiji ciljana varijabla (medijalna, a kasnije i prosječna cijena stana) izravno je izračunata isključivo za one jedinice lokalne samouprave koje bilježe najmanje 30 kupoprodajnih transakcija u 2025. godini. Ovaj prag odabran je kako bi se osiguralo da su izravno izračunate cijene dovoljno stabilne, reprezentativne i robusne za potrebe treniranja i evaluacije prediktivnih modela.

Primjenom tog kriterija, iz ukupnog skupa od 556 jedinica lokalne samouprave formiran je poduzorak od 74 jedinice za koje su istovremeno dostupni i izravno izračunata ciljana varijabla (medijalna cijena stana u 2025. godini) i svih 20 objašnjavajućih varijabli uključenih u početni prediktivni model.

Distribucija broja kupoprodajnih transakcija u cijelom uzorku dodatno potvrđuje opravdanost ovakvog restriktivnog pristupa. Dok je medijan broja transakcija u svim jedinicama lokalne samouprave svega 30,5, čak 414 jedinica nema niti jednu izravno zabilježenu vrijednost ciljne varijable, a znatan dio preostalih jedinica bilježi vrlo mali

broj transakcija. Time se jasno pokazuje da bi uključivanje jedinica s manjim brojem transakcija u proces treniranja dovelo do učenja modela na statistički nestabilnim i potencijalno pristranim ciljnim vrijednostima.

Nakon definiranja analitičkog uzorka, provedena je nasumična podjela podataka na trening i test (holdout) uzorak u omjeru 80 % prema 20 %. Konkretno, od ukupno 74 jedinice lokalne samouprave uključene u analizu, 59 jedinica korišteno je za treniranje modela, dok je preostalih 15 jedinica rezervirano isključivo za evaluaciju prediktivne uspješnosti. Podjela je provedena korištenjem funkcije `initial_split` iz paketa *rsample*, uz fiksirano sjeme generatora slučajnih brojeva, čime je osigurana potpuna replikabilnost rezultata.

5.3. Usporedba performansi modela strojnog učenja sa proširenim skupom objašnjavajućih varijabli radi identifikacije ključnih varijabli za predikciju medijalnih cijena stanova

U inicijalnom dijelu provedbe strojnog učenja proveli smo usporednu procjenu sedam različitih modela s ciljem identifikacije objašnjavajućih varijabli koje najviše doprinose preciznoj predikciji medijalne cijene stanova, a ne s ciljem odabira konačnog, optimalnog prediktivnog modela. Svi modeli trenirani su na istom prostorno proširenom početnom modelu koji uključuje 20 objašnjavajućih varijabli definiranih u prethodnom poglavlju, među kojima je i konstruirani prostorni pomak ciljne varijable, a sam trening modela proveden je na testnom uzorku od 59 lokalnih jedinica lokalne samouprave koje su sve u 2025. ostvarile najmanje 30 kupoprodajnih transakcija stanovima. Ovaj je testni uzorak korišten isključivo kao zajednički referentni okvir za izračun standardnih metrika uspješnosti za svih sedam modela korištenih u strojnom učenju. U ovoj fazi, mjere poput MAE, RMSE, R^2 i prilagođenog R^2 ne interpretiraju se kao kriteriji za izbor „najboljeg“ modela, već kao pomoćni alati za procjenu stabilnosti, robusnosti i informativnosti pojedinih varijabli unutar različitih modelskih struktura. Na trening uzorku su trenirani modeli opisani u trećem poglavlju studije; višestruka linearna regresija, penalizirane regresijske metode (LASSO, Ridge i Elastic Net s 10-strukom unakrsnom validacijom), Regression Tree, Random Forest (1000 stabala) i XGBoost model, pri čemu je svaki od tih pristupa pružio komplementarne informacije o relativnoj važnosti pojedinih prediktora.

Dobiveni rezultati (Tablica 2.) pokazuju da penalizirane linearne metode, osobito LASSO i Elastic Net, ostvaruju stabilne i konzistentne odnose između objašnjavajućih varijabli i ciljne varijable, što ih čini osobito pogodnima za redukciju dimenzionalnosti i selekciju varijabli. Istodobno, stablasti i ansambl modeli pružaju dodatni uvid u nelinearne obrasce i međudjelovanja varijabli, što je važno za provjeru robusnosti selekcije. Zaključno, svrha ove faze nije identifikacija konačnog modela za predikciju cijena stanova, već sustavno

smanjenje početnog skupa od 20 varijabli na manji, informacijski najbogatiji podskup, koji će u sljedećoj fazi poslužiti kao temelj za izgradnju zaključnog prediktivnog modela.

TABLICA 2. Usporedba performansi modela s punim skupom od 20 objašnjavajućih varijabli

Model	MAE	RMSE	R ²	Prilagođeni R ²
Višestruka linearna regresija	358,5	535,7	0,606	1,919
LASSO	291,4	487,3	0,674	1,761
Ridge	346,4	515,2	0,635	1,851
Elastic Net	282,0	465,6	0,702	1,694
Regression Tree	430,9	566,8	0,559	2,029
Random Forest	429,0	567,3	0,558	2,031
XGBoost	421,8	561,8	0,567	2,011

Napomena: ciljana varijabla medijalna cijena stanova s PDV-om po m²; MAE = srednja apsolutna pogreška; RMSE = korijen srednje kvadratne pogreške; R² – koeficijent determinacije; prilagođeni R² – prilagođeni koeficijent determinacije.

Izvor: izračun autorice.

Kao što je već prethodno napomenuto, cilj ovog dijela analize nije identifikacija najboljeg pojedinačnog prediktivnog modela, već **sustavna i višemodelska procjena prediktivne važnosti objašnjavajućih varijabli**, s isključivim ciljem **redukcije početnog skupa od 20 objašnjavajućih varijabli na manji, informacijski najbogatiji podskup**. U tu svrhu je na temelju procjena modela čiji su rezultati prikazani u Tablici 2., pokrenut postupak određivanja važnosti pojedinih objašnjavajućih varijabli za prediktivnu moć modela.

U tom je postupku važnost objašnjavajućih varijabli procijenjena korištenjem više komplementarnih pristupa temeljenih na različitim modelima strojnog učenja. **Permutation importance** izračunata je **isključivo na temelju Random Forest modela**, pri čemu se važnost pojedine varijable definira kao porast pogreške predikcije (RMSE) nakon nasumične permutacije njezinih vrijednosti, uz zadržavanje svih ostalih varijabli nepromijenjenima. Veći porast pogreške upućuje na veći stvarni doprinos varijable prediktivnoj sposobnosti modela. Paralelno s tim, iz **penalized regression** modela (LASSO, Ridge i Elastic Net) korištene su **apsolutne vrijednosti procijenjenih koeficijenata** pri optimalnoj vrijednosti parametra penalizacije (λ), što omogućuje usporedbu relativne važnosti varijabli unutar linearnih, ali regulariziranih specifikacija. Dodatno, iz **Random Forest** modela preuzeta je i **impurity-based importance**, koja mjeri prosječno smanjenje varijance ciljne varijable uzrokovano pojedinom varijablom kroz cijeli ansambl stabala, dok je iz **XGBoost** modela korištena mjera **gain**, koja kvantificira doprinos pojedine varijable smanjenju funkcije gubitka tijekom procesa gradijentnog pojačavanja.

Važno je naglasiti da za višestruku linearnu regresiju i za regresijsko stablo (Regression Tree) u ovom okviru nije korištena niti jedna eksplicitna metrika važnosti pojedinih varijabli. Kod višestruke linearne regresije procijenjeni koeficijenti nisu usporediva mjera važnosti zbog osjetljivosti na skalu varijabli i prisutnost multikolinearnosti, dok kod regresijskog stabla stabilne i interpretabilne mjere globalne važnosti varijabli nisu dostupne u obliku usporedivom s ansambl metodama. Iz tog razloga, ove dvije metode korištene su isključivo za evaluaciju prediktivne uspješnosti modela, ali ne i za konstrukciju rangiranja važnosti objašnjavajućih varijabli. Također je važno napomenuti da Ridge regression, za razliku od LASSO i Elastic Net modela, ne provodi selekciju varijabli, odnosno zadržava sve objašnjavajuće varijable u modelu, pri čemu se njihovi koeficijenti kontinuirano smanjuju prema nuli, ali nikada ne postaju točno jednaki nuli. Zbog toga se u Ridge regresiji važnost varijabli ne može određivati na temelju njihove statističke značajnosti, budući da standardni testovi značajnosti nisu dobro definirani u prisutnosti penalizacije. Umjesto toga, relativna važnost varijabli u Ridge modelu procjenjuje se na temelju apsolutne veličine standardiziranih regresijskih koeficijenata, pri čemu veća apsolutna vrijednost koeficijenta upućuje na veći doprinos varijable prediktivnoj sposobnosti modela.

Svi ovi pokazatelji objedinjeni su u jedinstvenu tablicu važnosti varijabli (Tablica 3.), pri čemu su nedostajuće vrijednosti kod penalizirajućih regresijskih modela postavljene na nulu kako bi se osigurala usporedivost među metodama. Budući da su dobivene mjere po svojoj skali i interpretaciji međusobno neusporedive, u sljedećem koraku svaka je varijabla **rangirana unutar pojedinog modela**, pri čemu su varijable s nultom važnošću (primjerice regresijski koeficijent točno jednak nuli u LASSO regresiji) izuzete iz rangiranja za tu metodu, što je napravljeno tako da je toj varijabli nije dodijeljen rang. Rangiranje je provedeno tako da niži rang označava veću važnost varijable.

Nakon toga je za svaku varijablu izračunat **konsenzusni rang** kao aritmetička sredina rangova dobivenih iz svih metoda u kojima je varijabla imala nenultu važnost. Time je dobivena jedinstvena, model-agnostička mjera relativne važnosti varijabli, koja ublažava pristranosti pojedinih algoritama i naglašava one varijable koje su stabilno visoko rangirane kroz različite metodološke pristupe. Rangovi važnosti varijabli za predikciju medijalnih cijena stanova prikazani su na Slici 1, a numeričke vrijednosti svih rangova te završni konsenzusni rang prikazani su u Tablici 4.

TABLICA 3. Višemodelska procjena važnosti objašnjavajućih varijabli

Objašnjavajuća varijabla	LASSO	Ridge	Elastic Net	Random Forest (impurity)	Random Forest (permutation)	XGBoost (gain)
dist_sea	0,000	0,001	0,001	3.595.791	96,013	0,315
lag_target	0,486	0,271	0,445	2.861.330	91,482	0,199
vacapart	0,000	0,086	0,001	2.802.917	69,625	0,105
ir	32,157	30,349	35,241	1.922.296	71,380	0,087
itr	23,733	15,557	24,313	2.326.330	50,323	0,075
comapart_kum12_22	0,000	0,050	0,035	817.388	22,582	0,057
vitindex	0,000	0,700	0,000	642.544	16,475	0,022
coast_length	0,000	0,000	0,000	2.245.193	56,072	0,021
permunocushare	2,470	5,751	6,583	559.753	15,938	0,020
dist_highway	0,000	0,002	0,000	426.966	11,003	0,018
urban_perc	0,000	0,036	0,000	428.919	11,485	0,016
tfr	0,000	56,743	0,000	441.697	13,920	0,016
length_roads	0,000	0,001	0,000	630.682	15,982	0,014
forest_perc	0,000	2,387	0,914	913.568	29,500	0,012
densityhousehold	0,000	0,021	0,000	520.471	12,020	0,010
dist_airport	0,000	0,001	0,000	438.204	10,006	0,008
water_perc	0,000	29,315	1,296	467.645	7,066	0,004
dist_marina	0,000	0,001	0,000	1.396.447	21,930	0,000
island	0,000	5,453	0,000	50.108	1,309	0,000
metro	0,000	0,000	0,000	0	0,000	0,000

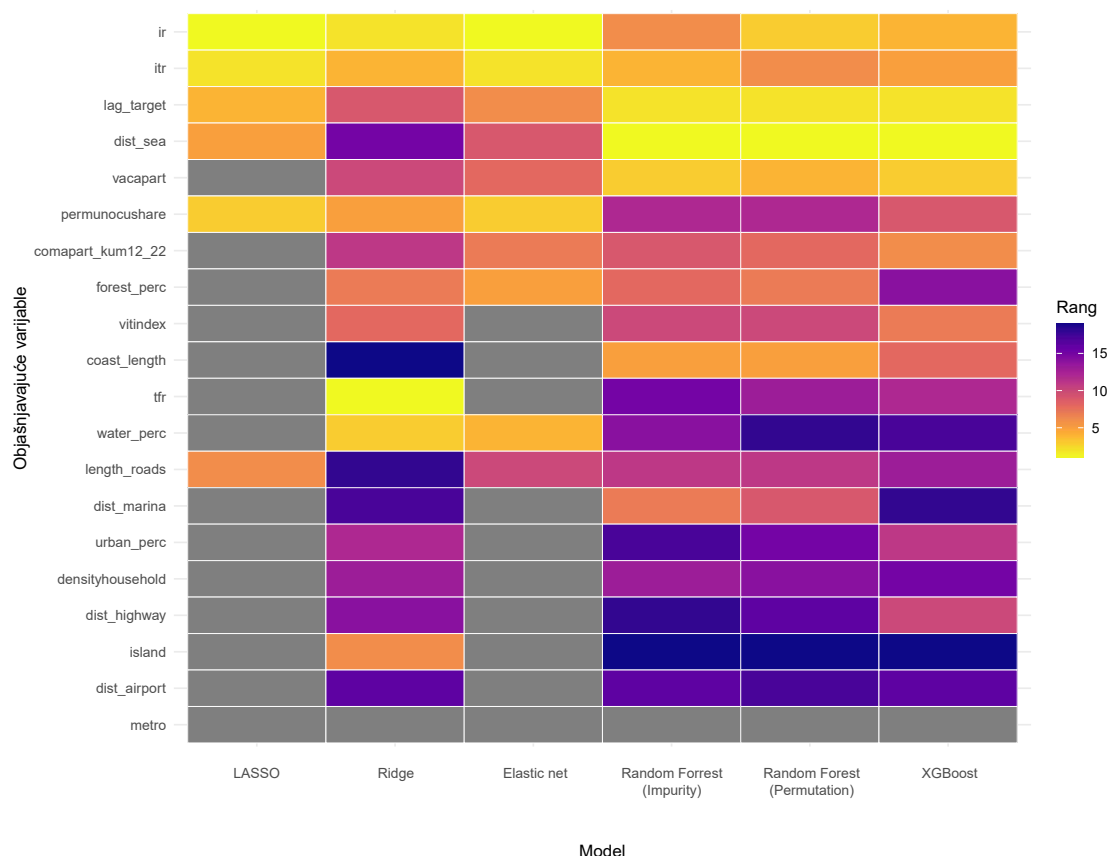
Napomena: ciljane varijabla medijalna cijena stanova s PDV-om po m².

Izvor: izračun autorice.

Rezultati tog postupka jasno pokazuju da se pet varijabli konzistentno izdvaja kao najinformativnije, poredanih prema važnosti prema konsenzusnom rankingu. Na prvom mjestu nalazi se varijabla **udaljenost od mora (dist_sea)**, koja mjeri minimalnu udaljenost centroida jedinice lokalne samouprave od morske obale i hvata snažan prostorni gradijent cijena povezan s obalnom i turističkom atraktivnošću. Druga po važnosti je varijabla **prostorni pomak medijalne cijene stana (lag_target)**, definirana kao prosječna vrijednost ciljane varijable u prostorno najbližim jedinicama lokalne samouprave, kojom se u model eksplicitno uvodi prostorna međuovisnost i prenosi informacija o regionalnim cjenovnim obrascima. Treće mjesto zauzima **udio praznih stanova (vacapart)**, pokazatelj strukture i iskorištenosti stambenog fonda, koji signalizira neravnoteže između ponude i efektivne potražnje te je povezan kako sa spekulativnim ponašanjem, tako i s demografskim i prostornim obilježjima pojedinih lokalnih tržišta. **Indeks razvijenosti (ir)** je četvrta najvažnija varijabla za predikciju medijalne cijene

stana. To je složeni pokazatelj socioekonomske razvijenosti jedinice lokalne samouprave koji obuhvaća dohodovne, fiskalne i razvojne komponente te izravno odražava platežnu moć i tržišni potencijal lokalnog stanovništva. Na petom mjestu je **indeks turističke razvijenosti (itr)**, koji mjeri intenzitet turističke aktivnosti i važnost turizma u lokalnom gospodarstvu, a koji je snažno povezan s dodatnom potražnjom za stanovima, osobito u segmentu sekundarnog stanovanja i kratkoročnog najma.

SLIKA 1. Rang objašnjavajućih varijabli u izabranim modelima strojnog učenja



Napomena: ciljana varijabla medijalna cijena stanova s PDV-om po m²; objašnjavajuće varijable za koje nije izračunat rang jer su kod penalizirajućih regresijskih modela isključene iz procjene su označene sivom bojom. Isto vrijedi i za indikator varijablu metro kod stablastih ansambl modela.

Izvor: izračun autorice.

Upravo činjenica da se ove varijable pojavljuju među najvažnijima u različitim vrstama modela – od linearnih penaliziranih regresija do nelinearnih ansambl metoda i model-agnostičke permutacijske analize – daje dodatnu težinu njihovoj interpretaciji i opravdava njihovo izdvajanje. Na temelju tog konsenzusnog rangiranja donesena je odluka da se daljnja faza modeliranja fokusira na pet navedenih varijabli i njihove kombinacije, čime se osigurava da redukcija modela bude vođena empirijskom stabilnošću i robusnošću rezultata, a ne performansama jednog specifičnog algoritma. Pri tome je posebno

važno za napomenuti da od navedenih pet objašnjavajućih varijabli, važnost varijabli prostorni pomak medijalne cijene stana (lag_target), indeks razvijenosti (ir) i indeks turističke razvijenosti (itr) za određivanje medijalnih cijena stanova podržana i od strane penalizirajućih regresijskih modela, ali i od strane ansambl stablastih modela (Random Forrest i XGBoost). Preostale dvije varijable; udaljenost od mora (dist_sea) i udio praznih stanova (vacapart) su snažnije podržane od strane ansambl stablastih u odnosu na penalizirajuće regresijske modele.

TABLICA 4. Rangiranje važnosti varijabli prema pojedinim modelima i konsenzusnom rang

Objašnjavajuća varijabla	LASSO	Ridge	Elastic Net	Random Forest (impurity)	Random Forest (permutation)	XGBoost	Konsenzusni rang
dist_sea	17	16	15	1	1	1	1,5
lag_target	2	2	2	2	2	2	2,0
vacapart	-	7	14	3	4	3	3,5
ir	1	1	1	5	3	4	3,8
itr	3	3	3	4	5	5	4,0
permunocushare	4	4	4	15	13	9	8,2
comapart_kum12_22	-	12	10	12	9	6	9,8
forest_perc	-	8	8	10	8	14	11,3
vitindex	-	6	-	14	12	7	11,5
coast_length	-	18	-	6	6	8	11,8
water_perc	-	5	7	11	19	17	13,2
length_roads	18	17	16	13	11	13	14,7
dist_marina	-	19	-	7	10	18	15,7
urban_perc	-	11	-	17	16	11	15,8
tfr	-	0	-	16	14	12	16,0
dist_highway	-	14	-	18	17	10	16,5
densityhousehold	-	13	-	19	15	15	17,0
dist_airport	-	15	-	20	18	16	18,2
island	-	9	-	20	20	19	18,8
metro	-	-	-	-	-	-	-

Napomena: ciljane varijabla medijalna cijena stanova s PDV-om po m².

Izvor: izračun autorice.

5.4. Identifikacija ključnih varijabli za predikciju medijalnih cijena stanova

Nakon što je u prethodnom koraku, na temelju kombinacije model-specifičnih i model-agnostičkih mjera važnosti varijabli, utvrđen konsenzualni rang prediktivne važnosti objašnjavajućih varijabli, daljnja analiza usmjerena je na sustavno ispitivanje mogućnosti redukcije dimenzionalnosti početnog modela. Cilj ove faze nije bio identificirati jedan “najbolji” prediktivni model u užem smislu, već **procijeniti u kojoj mjeri se prediktivna sposobnost modela može očuvati uz znatno manji broj objašnjavajućih varijabli te time identificirati parsimonioznu strukturu pogodnu za zaključni model procjene cijena stanova.**

Kao što je već napomenuto, polazeći od konsenzualnog rangiranja, izdvojeno je pet varijabli koje su se pokazale najvažnijima u većini primijenjenih metoda strojnog učenja: indeks razvijenosti jedinice lokalne samouprave (ir), indeks turističke razvijenosti (itr), prostorni pomak medijalne cijene stana (lag_target), udaljenost od mora (dist_sea) te udio praznih stanova u stambenom fondu (vacapart). Ove varijable obuhvaćaju ključne ekonomske, turističke, prostorne i stambene mehanizme formiranja cijena nekretnina te stoga predstavljaju prirodan kandidatni skup za izgradnju reduciranih modela.

Na temelju tog skupa generirane su sve moguće kombinacije varijabli veličine dvije, tri, četiri i pet, čime je dobiveno ukupno 26 različitih podskupova objašnjavajućih varijabli. Za svaki od tih podskupova treniran je puni skup od sedam modela – višestruka linearna regresija, LASSO, Ridge, Elastic Net, regresijsko stablo, Random Forest i XGBoost – pri čemu su za svaki model izračunate standardne mjere uspješnosti na izdvojenom testnom uzorku: srednja apsolutna pogreška (MAE), korijen srednje kvadratne pogreške (RMSE), koeficijent determinacije (R^2) i prilagođeni koeficijent determinacije. U svim procjenama (a ukupno ih je 182, po 26 za svaku od sedam treniranih metoda strojnog učenja) korišteni su isti trening i testni uzorci, čime je osigurana potpuna usporedivost rezultata među različitim specifikacijama.

Rezultati¹ jasno pokazuju da se značajan dio prediktivne moći početnog modela s dvadeset varijabli može reproducirati već s tri ili četiri objašnjavajuće varijable. Posebno se ističe činjenica da kombinacije koje uključuju ekonomske pokazatelje razvijenosti (ir) i turističke aktivnosti (itr), zajedno s prostornim pomakom ciljne varijable (lag_target), ostvaruju znatno bolje rezultate od modela koji se oslanjaju isključivo na prostorne ili isključivo na ekonomske karakteristike. Dodavanje varijable udaljenosti od mora (dist_sea) dodatno poboljšava rezultate, osobito u linearnim i penaliziranim regresijskim

¹ Budući da je procijenjeno ukupno 182 modela, rezultati procjena su bili preveliki da bi bili prikazani u studiji, pa su radi uštede prostora prikazani u tablici P1. u Prilogu Studiji.

modelima, dok se marginalni doprinos pete varijable, udjela praznih stanova (vacapart), pokazuje manjim i manje stabilnim kroz različite algoritme.

Upravo ta stabilnost rezultata preko različitih modela predstavlja ključni nalaz ove faze analize. Linearni i penalizirani regresijski modeli dosljedno identificiraju kombinaciju varijabli *ir*, *itr*, *lag_target* i *dist_sea* kao optimalnu u smislu kompromisa između složenosti i prediktivne uspješnosti. Iako se kod nelinearnih modela, poput Random Foresta i XGBoosta, najbolji rezultati ponekad postižu s manjim brojem varijabli, takvi modeli pokazuju veću osjetljivost na specifikaciju i manju interpretabilnost, što ih čini manje pogodnima za potrebe službenih procjena i primjene u javnim politikama.

Ova faza analize stoga potvrđuje da je konsenzualni rang važnosti varijabli empirijski utemeljen te da identificirane ključne varijable nisu samo pojedinačno važne, već i zajedno čine stabilnu i informacijski bogatu osnovu za zaključni prediktivni model. Na temelju sustavne evaluacije svih kombinacija najvažnijih varijabli i različitih algoritama, dobiven je jasan empirijski argument za odabir parsimonioznog modela koji zadržava temeljne prostorne i ekonomske determinante cijena stanova, uz istodobno smanjenje rizika preučenja i povećanje transparentnosti procjena.

Tablica 5. sažima rezultate ove faze strojnog učenja u kojoj su, za svaki razmatrani algoritam strojnog učenja, identificirane najbolje specifikacije modela prema tri standardna kriterija uspješnosti: srednjoj apsolutnoj pogrešci (MAE), korijenu srednje kvadratne pogreške (RMSE) i prilagođenom koeficijentu determinacije (prilagođeni R^2). Tablica je konstruirana na temelju sustavnog pretraživanja svih kombinacija dviju do pet najvažnijih varijabli, prethodno identificiranih konsenzualnim rangiranjem važnosti, pri čemu je za svaku kombinaciju procijenjeno više modela (višestruka linearna regresija, LASSO, Ridge, Elastic Net, Random Forest, regresijsko stablo i XGBoost). Rezultati procjene sva 182 modela prikazana su u Tablici P1. u Prilogu studiji, dok su u Tablici 5. prikazani najbolji rezultati za svaku od sedam metoda strojnog učenja, odnosno rezultati za onu metodu strojnog učenja koja iskazuje najbolje prediktivne rezultate, pri čemu je RMSE korištena kao standardna mjera za izbor. Za svaki takav model u Tablici 5. prikazane su MAE, RMSE, i prilagođeni R^2 , te varijable uključene u model čija ciljana varijabla je medijalna cijena stana s uključenim PDV-om.

Rezultati prikazani u Tablici 5. dodatno potvrđuju visoku stabilnost linearnih i penaliziranih regresijskih modela. Kod višestruke linearne regresije, LASSO-a, Ridge-a i Elastic Neta, optimalne specifikacije prema kriterijima MAE i RMSE dosljedno uključuju istu kombinaciju četiri varijable: indeks razvijenosti (*ir*), indeks turističke razvijenosti (*itr*), prostorni pomak medijalne cijene stana (*lag_target*) i udaljenost od mora (*dist_sea*). Istodobno, najviši prilagođeni R^2 kod ovih modela postiže se s nešto parsimonioznijom specifikacijom koja izostavlja varijablu udaljenosti od mora, ali zadržava ključne

ekonomske i prostorne determinante. Nelinearni modeli, poput Random Foresta, regresijskog stabla i XGBoosta, u pravilu postižu najbolje rezultate s manjim brojem varijabli, no njihova prediktivna uspješnost mjerena RMSE-om ostaje slabija u odnosu na linearne i penalizirane regresijske modele, uz dodatno manju interpretabilnost.

Kao završni kriterij za odabir zaključnog modela za predikciju medijalnih (a kasnije i prosječnih) cijena stanova u ovoj studiji odabran je RMSE, budući da ta mjera jače penalizira velika odstupanja između stvarnih i predviđenih cijena te je posebno prikladna u kontekstu službenih procjena i primjene u javnim politikama. Prema tom kriteriju, najbolju ukupnu izvedbu postiže **model višestruke linearne regresije sa četiri objašnjavajuće varijable (ir, itr, lag_target i dist_sea)**, koji kombinira visoku prediktivnu točnost, stabilnost rezultata i jasnu ekonomsku interpretabilnost. Zbog toga je upravo ta specifikacija odabrana kao završni prediktivni model za procjenu cijena stanova u nastavku studije.

TABLICA 5. Najbolje specifikacije modela prema različitim kriterijima uspješnosti

Model stognog učenja	Najbolji MAE	Objašnjavajuće varijable	Najbolji RMSE	Objašnjavajuće varijable	Najbolji prilagođeni R ²	Objašnjavajuće varijable
Višestruka linearna regresija	292,84	ir, itr, lag_target, dist_sea	463,01	ir, itr, lag_target, dist_sea	0,619	ir, itr, lag_target
LASSO	292,15	ir, itr, lag_target, dist_sea	464,02	ir, itr, lag_target, dist_sea	0,618	ir, itr, lag_target
Ridge	281,54	ir, itr, lag_target, dist_sea	463,48	ir, itr, lag_target, dist_sea	0,619	ir, itr, lag_target
Elastic Net	286,56	ir, itr, lag_target, dist_sea	465,96	ir, itr, lag_target, dist_sea	0,619	ir, itr, lag_target
Random Forest	346,27	ir, dist_sea	518,58	ir, dist_sea	0,569	ir, dist_sea
Regression Tree	390,97	ir, itr, vacapart	538,96	ir, itr, vacapart	0,505	ir, vacapart
XGBoost	340,25	ir, dist_sea, vacapart	498,80	ir, dist_sea, vacapart	0,565	ir, dist_sea, vacapart

Napomena: ciljana varijabla medijalna cijena stanova s PDV-om po m².

Izvor: izračun autorice.

5.5. Evaluacija modela medijalnih cijena stanova s izabranim objašnjavajućim varijablama

U sljedećem koraku analize procijenjen je parsimoniozni model koji uključuje četiri objašnjavajuće varijable identificirane kao empirijski najstabilnije i najvažnije: indeks razvijenosti (ir), indeks turističke razvijenosti (itr), prostorni pomak medijalne cijene stana (lag_target) i udaljenost od mora (dist_sea). Iako su prethodni rezultati upućivali na to da višestruka linearna regresija s tom specifikacijom ostvaruje najbolje performanse, u ovoj je fazi dodatno provedena usporedba s ostalim metodama strojnog učenja uz identičan skup objašnjavajućih varijabli.

Najprije je procijenjen model višestruke linearne regresije na trening uzorku. Rezultati pokazuju da su varijable ir, itr i lag_target statistički značajno i pozitivno povezane s medijalnom cijenom stanova, dok varijabla dist_sea ima negativan, ali statistički neznačajan učinak, što upućuje na njezin posredni doprinos kroz zajedničko djelovanje s ostalim varijablama. Model objašnjava oko 73 % varijance u trening uzorku i statistički je visoko značajan u cjelini.

U nastavku su svi razmatrani modeli – višestruka linearna regresija, LASSO, Ridge, Elastic Net, regresijsko stablo, Random Forest i XGBoost – evaluirani na izdvojenom testnom uzorku primjenom standardnih mjera uspješnosti (MAE, RMSE, R^2 i prilagođeni R^2). Rezultati potvrđuju da linearni i penalizirani regresijski modeli ostvaruju najbolje prediktivne performanse, pri čemu Ridge regresija postiže najniži MAE, dok višestruka linearna regresija ima najniži RMSE. Nelinearni modeli, osobito Random Forest i regresijsko stablo, pokazuju slabije rezultate, što upućuje na veću osjetljivost na preučenje u uvjetima ograničenog uzorka.

TABLICA 6. Usporedba modela za završnu specifikaciju (ir, itr, lag_target, dist_sea)

Model	MAE	RMSE	R^2	Prilagođeni R^2	Objašnjavajuće varijable
Višestruka linearna regresija	292,84	463,01	0,71	0,59	ir, itr, lag_target, dist_sea
LASSO	292,01	464,27	0,70	0,59	ir, itr, lag_target, dist_sea
Ridge	280,03	466,21	0,70	0,58	ir, itr, lag_target, dist_sea
Elastic Net	284,74	467,59	0,70	0,58	ir, itr, lag_target, dist_sea
Random Forest	407,09	574,99	0,55	0,36	ir, itr, lag_target, dist_sea
Regression Tree	451,97	619,51	0,47	0,26	ir, itr, lag_target, dist_sea
XGBoost	403,13	554,02	0,58	0,41	ir, itr, lag_target, dist_sea

Napomena: ciljana varijabla medijalna cijena stanova s PDV-om po m².

Izvor: izračun autorice.

Tablica 6. prikazuje sažete rezultate evaluacije svih modela za istu četverovarijantnu specifikaciju. U skladu s unaprijed definiranim kriterijem, kao završni model odabrana je višestruka linearna regresija, budući da kombinira najniži RMSE s visokom interpretabilnošću i stabilnošću procjena, što je ključno za primjenu rezultata u kontekstu službenih procjena i javnih politika.

5.6. Zaključni model za procjenu medijalnih cijena stanova u jedinicama lokalne samouprave i rezultati procjene modela

U završnom koraku analize treniran je model koji se koristi za predikciju medijalnih cijena stanova za sve jedinice lokalne samouprave u Republici Hrvatskoj, uključujući i one JLS-ove za koje u promatranoj godini ne postoji dovoljan broj tržišnih transakcija. Time se analitički okvir pomiče s faze evaluacije i usporedbe modela prema njihovoj operativnoj primjeni u svrhu izrade sveobuhvatnih i konzistentnih procjena na nacionalnoj razini.

Najprije je formiran trening uzorak koji uključuje samo one JLS-ove za koje je raspoloživa opažena medijalna cijena stana (ciljna varijabla) i koje imaju najmanje 30 evidentiranih kupoprodajnih transakcija. Time se osigurava da se procjena modela temelji na dovoljno pouzdanim i reprezentativnim tržišnim signalima. Istodobno je zadržan i puni skup JLS-ova, koji se koristi isključivo u fazi predikcije kako bi se omogućila procjena medijalnih cijena i za jedinice s ograničenom ili nepostojećom tržišnom aktivnošću.

U ovoj završnoj fazi analize procijenjen je **linearni regresijski model**, čija je specifikacija temeljena na **skupu objašnjavajućih varijabli identificiranih u prethodnim fazama analize** kroz sustavna testiranja različitih kombinacija varijabli i algoritama strojnog učenja. Prethodni rezultati pokazali su da upravo određena kombinacija varijabli dosljedno daje stabilne i interpretabilne rezultate kroz različite modele, zbog čega je ona odabrana kao jezgra zaključnog prediktivnog okvira.

Kako bi se dodatno unaprijedile prediktivne sposobnosti modela te uvažila prostorna i institucionalna heterogenost jedinica lokalne samouprave, osnovni linearni regresijski model proširen je u **linearni mješoviti model (linear mixed-effects model)** (Gelman i Hill, 2007). Ovakva specifikacija omogućuje istodobno modeliranje utjecaja ključnih objašnjavajućih varijabli kroz fiksne efekte, uz dodatno uvažavanje razvojne hijerarhijske strukture podataka putem slučajnih efekata. Ova specifikacija se, nažalost, nije mogla direktno trenirati, testirati i uspoređivati uz ostale metode strojnog učenja jer ostale metode ne podržavaju treniranje modela koji se djelomično zasniva na slučajnom učinku, odnosno na uvođenju slučajnog odsječka u specifikaciju modela.

Ciljna varijabla u linearnom mješovitom modelu definirana je kao prirodni logaritam medijalne cijene stana, čime se stabilizira varijanca i omogućuje interpretacija koeficijenata u približno relativnim (postotnim) terminima. Model uključuje četiri fiksna efekta koji su u prethodnim fazama analize identificirani kao empirijski najstabilniji i najinformativniji determinanti cijena stanova:

- indeks razvijenosti (*ir*)
- indeks turističke razvijenosti (*itr*)
- prostorni pomak medijalne cijene stana (*lag_target*)
- udaljenost od mora, uključena u model kao logaritamska transformacija $\log(\text{dist_sea})$

Uz fiksne efekte, model uključuje i **slučajni odsječak (random intercept)** na razini varijable za grupu kojoj jedinica lokalne samouprave pripada prema indeksu razvijenosti (*gir*). Vrijednost ove varijable se kreće od 1 do 8 (1 = najnerazvijenija skupina lokalnih jedinica; 8 = najrazvijenija grupa lokalnih jedinica) i ona predstavlja šire skupine jedinica lokalne samouprave. Uvođenjem slučajnog odsječka dopušta se da se osnovna razina cijena razlikuje među razvojnim skupinama, neovisno o uključenim objašnjavajućim varijablama. Time se modelom eksplicitno uvažavaju neopažene regionalne i institucionalne karakteristike koje mogu utjecati na razinu cijena stanova.

Ovakav pristup osobito je važan s obzirom na činjenicu da se dobivene procjene koriste prvenstveno za potrebe provedbe javnih politika, u kojima stupanj razvijenosti lokalne zajednice ima ključnu ulogu u dizajnu, ciljanju i implementaciji programa. Uključivanjem slučajnog odsječka model dodatno veže procijenjene vrijednosti medijalnih cijena uz širi kontekst razvijenosti i regionalne pripadnosti JLS-a, čime se povećava institucionalna relevantnost i robusnost rezultata.

Formalno, procijenjeni model može se zapisati kao:

$$\log(mprice_i) = \beta_0 + \beta_1 ir_i + \beta_2 itr_i + \beta_3 lag_target_i + \beta_4 \log(dist_sea_i + 1) + u_{g(i)} + \varepsilon_i$$

gdje je:

- $mprice_i$ – medijalna cijena stana u JLS-u *i*
- β_0 – zajednički odsječak (fixed intercept)
- $\beta_1, \dots, \beta_4$ – koeficijenti fiksnih efekata
- $u_{g(i)} \sim N(0, \sigma_u^2)$ – slučajni odsječak za razvojnu skupinu *gir* kojoj JLS pripada
- $\varepsilon_i \sim N(0, \sigma^2)$ – slučajna pogreška

Model je procijenjen metodom REML (Restricted Maximum Likelihood), što je standardni i preporučeni pristup za procjenu varijanci slučajnih efekata u mješovitim modelima (Bates et al, 2015).

Rezultati procjene prikazani u Tablici 7. potvrđuju da su indeks razvijenosti, prostorni pomak cijena i udaljenost od mora statistički značajni determinanti medijalnih cijena stanova, dok indeks turističke razvijenosti ima pozitivan, ali slabiji učinak kada se kontrolira za ostale čimbenike. Procijenjena varijanca slučajnog odsječka po razvojnim skupinama prikazana u Tablici 8. upućuje na postojanje dodatne regionalne heterogenosti u razinama cijena, koja nije u potpunosti obuhvaćena fiksnim dijelom modela.

Rezultati procjene potvrđuju da ovakva mješovita specifikacija linearnog regresijskog modela ima empirijsko opravdanje. Procijenjena varijanca slučajnog odsječka iznosi 0,0017, dok rezidualna varijanca iznosi 0,0212. Iz tih vrijednosti izračunat je intraklasni koeficijent korelacije (ICC) od približno 7,4 %, što znači da se oko sedam posto ukupne varijance logaritmira medijalne cijene stanova može pripisati razlikama među razvojnim skupinama aproksimiranim varijablom *gir*. Iako većina varijabilnosti proizlazi iz razlika među pojedinim JLS-ovima koje su određene objašnjavajućim varijablama, ovaj udio jasno upućuje na postojanje sistematske regionalne heterogenosti koju čista višestruka linearna regresija i četiri korištene objašnjavajuće varijable ne mogu eksplicitno modelirati.

TABLICA 7. Procjena fiksnih efekata u linearnom mješovitom modelu

Objašnjavajuća varijabla	Procjena koeficijenta	Standardna pogreška	t-vrijednost
Slučajni odsječak	4,398	0,575	7,645
Indeks razvijenosti (ir)	0,0269	0,0060	4,479
Indeks turističke razvijenosti (itr)	0,0057	0,0041	1,392
Prostorni pomak cijena (lag_target)	0,00017	0,00005	3,510
Udaljenost od mora (dist_sea)	-0,0186	0,0071	-2,636

Napomena: ciljana varijabla medijalna cijena stanova s PDV-om po m².

Izvor: izračun autorice.

TABLICA 8. Varijanca slučajnih i rezidualnih komponenti

Komponenta varijance	Varijanca	Standardna devijacija
Slučajni odsječak (gir)	0,0017	0,0413
Rezidualna varijanca	0,0212	0,1455

Napomena: ciljana varijabla medijalna cijena stanova s PDV-om po m².

Izvor: izračun autorice.

Uključivanje slučajnog presretanja omogućuje modelu da “upije” dio neopaženih regionalnih čimbenika – poput povijesnih razvojnih putanja, institucionalnih razlika ili šireg prostornog konteksta, čime se poboljšava stabilnost koeficijenata fiksnih efekata i povećava robusnost predikcija. To je osobito važno u kontekstu primjene rezultata u javnim politikama, gdje je stupanj razvijenosti lokalne zajednice ključna dimenzija u dizajnu i provedbi različitih programa, a procjene cijena stanova moraju biti konzistentne i usporedive među teritorijalnim cjelinama. Dodatno, procijenjeni fiksni efekti u mješovitom modelu zadržavaju očekivane znakove i statističku značajnost, što potvrđuje da uvođenje slučajnog presretanja ne narušava interpretabilnost osnovnih ekonomskih i prostornih determinanti cijena stanova, već ih nadopunjuje. Distribucija skaliranih reziduala pokazuje približno simetričan oblik oko nule, bez izraženih sistematskih odstupanja, što dodatno upućuje na adekvatnu specifikaciju modela.

Kako bi bili apsolutno sigurni da je linearni mješoviti model unaprijeđenja u odnosu na model višestruke linearne regresije koji je u prethodnim fazama analize odabran kao najbolji model za predikciju s četiri ključne objašnjavajuće varijable, u Tablici 9. prikazana je usporedba prediktivnih rezultata višestrukog linearnog regresijskog modela i linearnog mješovitog modela medijalnih cijena stanova sa četiri odabrane objašnjavajuće varijable. Ta usporedba pokazuje da uvođenje hijerarhijske strukture putem slučajnog odsječka formiranog na osnovi pripadnosti jedinice lokalne samouprave razvojnoj skupini (*gir*) dovodi do sustavnog poboljšanja prediktivne uspješnosti modela. Linearni mješoviti model ostvaruje nižu prosječnu apsolutnu pogrešku (MAE) u odnosu na standardnu regresiju (270 €/m² naspram 293 €/m²), što znači da su predikcije u prosjeku bliže stvarnim vrijednostima medijalnih cijena stanova. Istodobno, RMSE je također niži kod mješovitog modela (453 €/m² u odnosu na 463 €/m²), što upućuje na smanjenje velikih pogrešaka i bolju kontrolu ekstremnih odstupanja u predikcijama.

Koeficijent determinacije dodatno potvrđuje prednost mješovitog modela: R^2 se povećava s 0,71 na 0,72, što znači da model s hijerarhijskom komponentom objašnjava nešto veći udio ukupne varijabilnosti medijalnih cijena stanova. Iako je razlika u R^2 numerički umjerena, ona je metodološki značajna jer se postiže bez povećanja broja objašnjavajućih varijabli, već isključivo boljim modeliranjem prostorne i institucionalne strukture podataka.

Ključna prednost linearnog mješovitog modela u odnosu na čistu linearnu regresiju stoga proizlazi iz njegove sposobnosti da eksplicitno uvaži neopažene regionalne razlike u razinama cijena stanova. Uvođenjem slučajnog presretanja dopušta se da osnovna razina cijena varira među razvojnim skupinama, čime se smanjuje pristranost procjena i povećava stabilnost koeficijenata fiksnih efekata. To je osobito važno u kontekstu primjene rezultata u javnim politikama, gdje se procjene koriste za usporedbu jedinica lokalne samouprave različitog stupnja razvijenosti i institucionalnog okruženja.

TABLICA 9. Usporedba višestruke linearne regresije i linearnog mješovitog modela

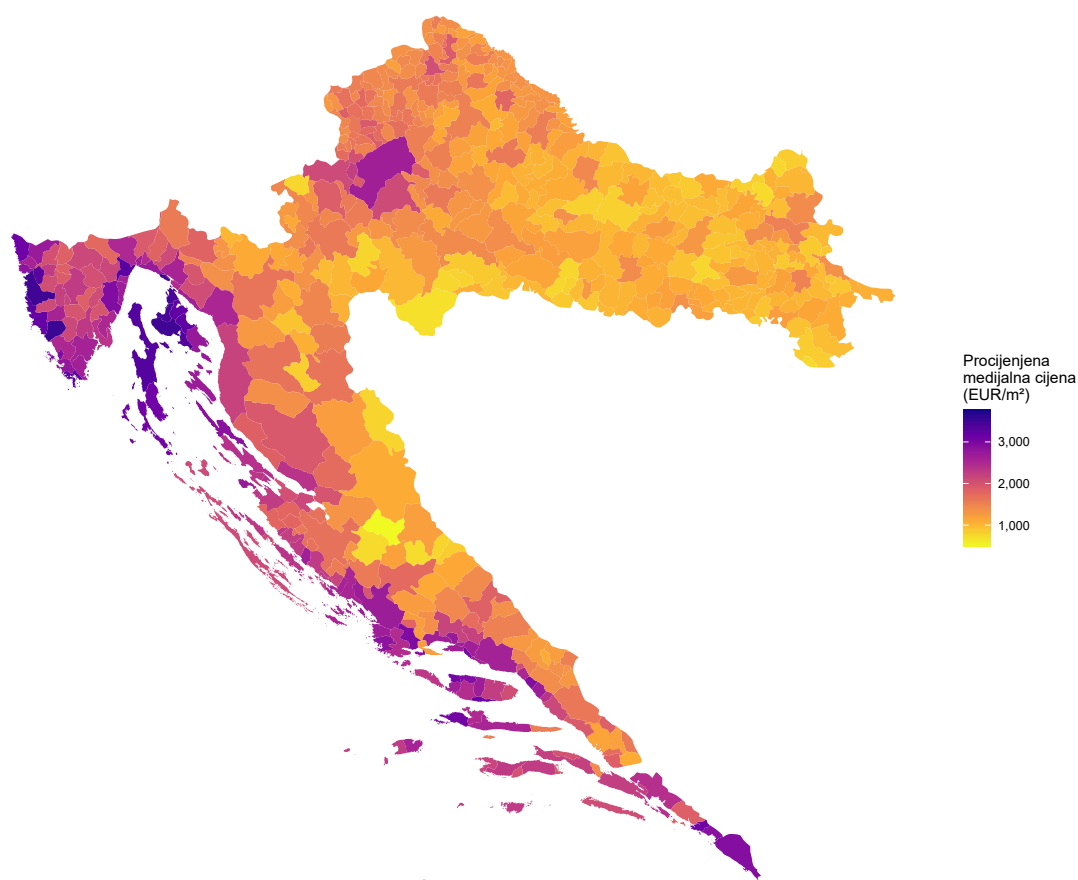
Model	MAE	RMSE	R ²
Višestruka linearna regresija	292,84	463,01	0,71
Linearni mješoviti model	270,21	453,47	0,72

Napomena: ciljane varijabla medijalna cijena stanova s PDV-om po m², MAE, RMSE i R² za linearni mješoviti model izračunate su na originalnim (nelogaritmiranim) podacima za ciljane varijablu kako bi mjere bile usporedive s onima koje ostvaruje model višestruke linearne regresije.

Izvor: izračun autorice.

Na temelju navedenih rezultata, linearni mješoviti model o odabranim ključnim objašnjavajućim varijablama na osnovi rezultata strojnog učenja predstavlja metodološki i empirijski utemeljen unaprjeđenje u odnosu na model višestruke linearne regresije. On također uvažava potrebu za uvođenjem hijerarhijske strukture podataka. U odnosu na čistu linearnu regresiju, ovakav pristup omogućuje realističnije modeliranje regionalnih razlika te pouzdanije i policy-relevantne procjene medijalnih cijena stanova za sve jedinice lokalne samouprave u Republici Hrvatskoj.

SLIKA 2. Procijenjene medijalne cijene stanova s uključenim PDV-om



Izvor: izračun autorice.

Na temelju procijenjenog modela generirane su predikcije medijalnih cijena stanova za sve jedinice lokalne samouprave u Hrvatskoj, pri čemu su predikcije transformirane natrag iz logaritamskog u izvornu cjenovnu skalu. Dobiveni rezultati kartografski su prikazani na Slici 2, koja prikazuje prostornu distribuciju procijenjenih medijalnih cijena stanova za sve JLS-ove u Republici Hrvatskoj. Oni jasno ilustriraju ulogu prostornih, ekonomskih i tržišnih čimbenika u formiranju medijalne cijene stanova. Medijalna cijena stanova nam je u kontekstu ove studije važna jer će se ona koristiti za procjenu medijalne najamnine, o čemu će više biti riječi u sljedećem poglavlju studije. Ovdje je samo važno napomenuti da je u svrhu izračuna medijalnih mjesečnih najamnina korištena stvarna medijalna najmnina za jedinicu lokalne samouprave ukoliko je ta jedinica ostvarila 20 i više ugovora o najmu stanova u 2025. godini. Ukoliko nije, korištene su procjene medijalne cijene stanova za lokalne jedinice prikazane na Slici 2.

Valja napomenuti da je u prilogu studiji, u Tablici P3., prikazana i procjena medijalnih cijena stanova s PDV-om za katastarske općine Grada Zagreba. Procjene su izrađene kako bi se omogućila detaljnija primjena rezultata u kontekstu nacionalnih stambenih politika na području grada Zagreba. Međutim, procjena medijalnih cijena stanova na razini katastarskih općina se, nažalost, ne može provesti primjenom istog postupka strojnog učenja koji je korišten za jedinice lokalne samouprave.

Razlog tome je strukturna neusklađenost dostupnih podataka: dok se transakcijski podaci o kupoprodaji stanova u bazi eNekretnine administrativno vode po katastarskim općinama, svi ostali relevantni statistički podaci o ekonomskoj, demografskoj, turističkoj i građevinskoj aktivnosti te o stambenom fondu koji su potrebni za strojno učenje i statističko modeliranje dostupni su isključivo na razini gradskih naselja. Ovakva razlika u prostornoj razini podataka onemogućuje izravnu primjenu strojnog učenja i posljedičnu procjenu medijalnih cijena stanova po katastarskim općinama. Da bi se takav pristup mogao metodološki ispravno provesti, bilo bi nužno sve transakcije iz baze eNekretnine, korištenjem detaljnih adresnih podataka, jednoznačno povezati s pripadajućim gradskim naseljima, što nadilazi opseg i vremenski okvir ovog projekta.

S obzirom na navedena ograničenja, medijalne cijene stanova za katastarske općine Grada Zagreba procijenjene su primjenom hijerarhijskog i pragom uvjetovanog pristupa koji u najvećoj mogućoj mjeri koristi stvarno ostvarene transakcijske podatke. U katastarskim općinama koje su u promatranom razdoblju bilježile 20 ili više zabilježenih kupoprodajnih ugovora, kao medijalna cijena korištena je stvarno ostvarena medijalna vrijednost izračunata iz podataka dostupnih u bazi eNekretnine.² Za katastarske općine koje nisu imale niti dovoljan broj kupoprodajnih transakcija, medijalna cijena stana

² Podrobnije objašnjenje razloga zbog kojeg je kao prag broja transakcija izabran 20 može se naći u osmom poglavlju Studije.

procijenjena je imputacijom vrijednosti iz prostorno najbliže katastarske općine koja raspolaže stvarno zabilježenom ili procijenjenom medijalnom cijenom.

Opisanim postupcima stoga je osigurana potpuna prostorna pokrivenost Republike Hrvatske i Grada Zagreba procijenjenim medijalnim vrijednostima cijena stanova, što nam u sljedećem koraku omogućava izračun mjesečnih medijalnih najamnina za sve jedinice lokalne samouprave i katastarske općine Grada Zagreba uz najveći mogući stupanj minimiziranja proizvoljnosti tih procjena.

6. Procjena medijalnih najamnina za jedinice lokalne samouprave

U ovom poglavlju razmatra se metodološki pristup procjeni medijalnih najamnina po jedinicama lokalne samouprave u Republici Hrvatskoj, pri čemu se polazi od prethodno procijenjenih medijalnih cijena stanova dobivenih primjenom metoda strojnog učenja. Za razliku od tržišta kupoprodaje, gdje je unatoč ograničenjima moguće izravno izračunati cijene za određeni broj lokalnih jedinica, tržište najma obilježeno je još izraženijim problemima prostorne pokrivenosti i statističke pouzdanosti administrativnih podataka. Upravo ta činjenica ključni je razlog zbog kojeg metode strojnog učenja nisu izravno primijenjene na modeliranje medijalnih najamnina, već je usvojen dvofazni pristup koji najprije modelira medijalne cijene stanova opisan u prethodnom poglavlju Studije, a zatim na temelju empirijski uočenog odnosa između cijena i najamnina rekonstruira procjene najamnina za sve jedinice lokalne samouprave. Druga faza tog pristupa opisuje se u ovom poglavlju Studije.

6.1. Opis podataka i metodološki pristup procjeni medijalnih najamnina

Analiza podataka iz sustava eNekretnine za prvih jedanaest mjeseci 2025. godine jasno pokazuje razmjere ograničenja u pogledu dostupnosti podataka o najamninama. Dok Republika Hrvatska ima ukupno 556 jedinica lokalne samouprave, izravni izračun medijalnih najamnina moguć je samo za one jedinice koje u promatranom razdoblju bilježe najmanje deset evidentiranih ugovora o najmu. Taj uvjet u 2025. godini ispunjava svega 51 jedinica lokalne samouprave, što čini oko 9 % ukupnog broja JLS. Za preostalih 496 jedinica, odnosno za gotovo 91 % svih lokalnih jedinica u Republici Hrvatskoj, administrativni podaci ne omogućuju nikakav statistički smislen izračun najamnina (Tablica 1).

Dodatni problem predstavlja i struktura dostupnih podataka unutar tog ionako vrlo ograničenog skupa jedinica. Od ukupno 51 JLS-ova za koje je izravni izračun najamnina moguć, čak 26 jedinica, odnosno 51 %, jedinica koje bilježe podatak o medijalnoj najamnini imaju registrirano do 30 ugovora o najmu godišnje. Medijalne najmnine u tim jedinicama temelje se stoga na izrazito malim uzorcima, zbog čega su izračuni osjetljivi na pojedinačne ugovore, heterogenost nekretnina i slučajne promjene u strukturi tržišta. U cijeloj Republici Hrvatskoj u prvih jedanaest mjeseci 2025. godine evidentirano je ukupno svega 2.555 ugovora o dugoročnom najmu stanova, što dodatno potvrđuje izrazitu koncentraciju tržišne aktivnosti u malom broju, pretežito urbanih lokalnih jedinica.

U takvim okolnostima izravna primjena metoda strojnog učenja na podatke o najamninama ne bi bila metodološki opravdana. Strojno učenje zahtijeva relativno velik broj opažanja, uravnoteženu prostornu pokrivenost i stabilan odnos između objašnjavajućih varijabli i ciljne varijable kako bi rezultati bili robusni i generalizabilni. U slučaju najamnina, izrazito mali broj opažanja i snažna prostorna koncentracija podataka doveli bi do modela koji bi u najvećoj mjeri odražavali specifičnosti nekolicine velikih gradova, uz vrlo slabu prediktivnu vrijednost za preostali dio zemlje. Takav pristup ne bi osigurao konzistentne procjene, niti bi te procjene bile relevantne za primjenu stambenih politika na nacionalnoj i lokalnoj razini.

Nasuprot tome, podaci o kupoprodajama, iako također ograničeni, ipak su dostupni za znatno veći broj jedinica lokalne samouprave. U 2025. godini najmanje deset kupoprodajnih transakcija bilježi 142 JLS-ova, dok 74 jedinice bilježe 30 i više transakcija, što u konačnici omogućuje procjenu i validaciju modela medijalnih cijena stanova na relativno širem i heterogenijem uzorku. Upravo zbog toga metode strojnog učenja primijenjene su na modeliranje medijalnih cijena stanova, koje su potom korištene kao temeljna i informacijski bogatija varijabla za daljnju rekonstrukciju najamnina.

Polazeći od jedinica lokalne samouprave za koje postoje pouzdani podaci i o medijalnim cijenama stanova i o medijalnim najamninama, u nastavku poglavlja empirijski se analizira odnos između ta dva tržišna pokazatelja. Na temelju tog odnosa rekonstruirati se procijenjena medijalna najamnina za sve jedinice lokalne samouprave, uključujući i one za koje izravni podaci o najmu ne postoje. Time se osigurava potpuna prostorna pokrivenost i konzistentnost procjena, što je nužan preduvjet za primjenu rezultata u kontekstu provedbe stambenih politika, osobito programa koji se oslanjaju na lokalno specifične, ali nacionalno usporedive pokazatelje tržišta najma.

Ovakav dvofazni pristup – prvo modeliranje cijena stanova primjenom metoda strojnog učenja, a zatim rekonstrukcija najamnina na temelju empirijski utemeljenog odnosa između cijena i najamnina – predstavlja metodološki opravdan kompromis između ograničenja raspoloživih podataka i potrebe za robusnim, transparentnim i za potrebe javnih politika relevantnim procjenama tržišnih najamnina za cijelo područje Republike Hrvatske. Valja napomenuti i da se podaci o medijalnim cijenama odnose na medijalne cijene s uključenim PDV-om., a isto se odnosi i na najmnine, iako podaci iz sustava eNekretnine sugeriraju da najamnina u koje je uključen PDV, gotovo i nema.

Za potrebe ovog dvofaznog pristupa, prvo je potrebno analizirati podatke o medijalnim cijenama stanova i medijalnim najamninama za stanove prikazane u Tablici 10. Ta Tablica, osim medijalnih cijena stanova i medijalnih najamnina za stanove, prikazuje i broj kupoprodajnih i najamnih ugovora za 2025. godinu za one jedinice lokalne samouprave koje istodobno bilježe najmanje 20 ugovora o najmu i dovoljan broj kupoprodajnih

transakcija da bi se medijalna cijena stana mogla izračunati i smatrati pouzdanom. Na temelju ta dva podatka izračunat je omjer medijalne najamnine i medijalne cijene stana, koji u nastavku studije služi kao ključni parametar za rekonstrukciju najamnina u jedinicama lokalne samouprave za koje izravni podaci o najmu ne postoje.

Valja napomenuti da bi, sukladno zaključcima osmoga poglavlja Studije u kojem se ocjenjuje pouzdanost izračuna medijalnih cijena za lokalne jedinice s malim brojem transakcija, metodološki opravdani pristup bio ograničiti analizu na jedinice lokalne samouprave s minimalno 20, a poželjnije 30 zabilježenih ugovora o najmu stanova. Međutim, zbog izrazito ograničene dostupnosti podataka o najamninama, u ovom je slučaju bilo nužno postići ravnotežu između zahtjeva za statističkom pouzdanošću medijalnih najamnina i potrebe za obuhvaćanjem dovoljnog broja jedinica lokalne samouprave kako bismo mogli procijeniti medijalne najamnine za sve lokalne jedinice. Stoga je, umjesto praga od 30, primijenjen minimalni prag od 20 zabilježenih ugovora o najmu i nakon primjene tog praga ostaje nam na raspolaganju svega 25 gradova koji bilježe podatak o medijalnoj najamnini.

Valja naglasiti i da su prikazani podaci za 2025. godinu nepotpuni, budući da u trenutku izrade studije nisu bili dostupni transakcijski podaci iz baze eNekretnine za prosinac 2025. godine. Ipak, s obzirom na visoku koncentraciju tržišne aktivnosti u većim i urbanim jedinicama lokalne samouprave, osobito na tržištu najma, izostanak jednog mjeseca ne mijenja bitno uočene obrasce odnosa između cijena stanova i najamnina. Također valja naglasiti i da su svi gradovi za koje se bilježe podaci o medijalnim najamninama veći i razvijeniji gradovi (uz izuzetak Knina koji pripada trećoj razvojnoj skupini po indeksu razvijenosti), što znači da su nam podaci za male, te posebice za nerazvijene gradove kada su medijalne najamnine u pitanju posve odsutni. To svojstvo podataka je ujedno i najveće ograničenje i nedostatak ove analize.

Medijalne cijene stanova po četvornom metru s PDV-om pokazuju snažne prostorne razlike. U primorskim jedinicama lokalne samouprave cijene su u pravilu znatno više nego u kontinentalnom dijelu zemlje. Najviše medijalne cijene zabilježene su u Splitu (4.077,5 EUR/m²), Dubrovniku (3.797,0 EUR/m²), Opatiji (3.719,5 EUR/m²) i Rovinju (3.483,0 EUR/m²), dok se u većini kontinentalnih gradova i općina medijalne cijene stanova kreću između približno 1.300 i 2.200 EUR/m². Grad Zagreb, kao najveće i najlikvidnije tržište u zemlji, zauzima srednju poziciju između ta dva prostora s medijalnom cijenom od 2.905,0 EUR/m².

Slični, ali ne identični obrasci vidljivi su i kod mjesečnih medijalnih najamnina po četvornom metru. U primorskim jedinicama najamnine su u apsolutnom iznosu više, osobito u velikim turističkim i urbanim središtima poput Splita (12,2 EUR/m²), Dubrovnika (15,0 EUR/m²), Rijeke (10,45 EUR/m²) i Opatije (10,0 EUR/m²). U kontinentalnim jedinicama

lokalne samouprave mjesečne medijalne najamnine uglavnom se kreću u rasponu od 6,0 do 9,5 EUR/m², pri čemu su najviše vrijednosti zabilježene u Gradu Zagrebu (12,7 EUR/m²) i Velikoj Gorici (9,5 EUR/m²).

Posebno je informativan omjer godišnje najamnine i cijene stana, koji se može interpretirati kao pokazatelj koliki se udio kupoprodajne vrijednosti stana "otplati" kroz jednu godinu plaćanja najamnine. Primjerice, omjer od 0,05 znači da godišnja najamnina iznosi približno 5 % vrijednosti stana po četvornom metru. Analiza tog omjera otkriva jasne razlike između primorskih i kontinentalnih krajeva Hrvatske. U primorskim jedinicama lokalne samouprave omjeri su u prosjeku niži i najčešće se kreću između 0,028 i 0,047. Najniže vrijednosti zabilježene su u izrazito turističkim sredinama s visokim cijenama stanova, poput Vodica (0,028), Opatije (0,032) i Splita (0,036). To upućuje na činjenicu da su cijene stanova u tim sredinama snažno pogonjene investicijskom i turističkom potražnjom, dok dugoročne najamnine rastu sporije, što rezultira nižim implicitnim godišnjim prinosima od najma.

Suprotno tome, u kontinentalnim jedinicama lokalne samouprave omjeri godišnje najamnine i cijene stana u pravilu su viši i stabilniji, najčešće između 0,043 i 0,057. Najviše vrijednosti bilježe Slavonski Brod i Osijek (0,057), Kutina (0,056) te Karlovac (0,054). To upućuje na relativno višu važnost tržišta dugoročnog najma u odnosu na kupoprodajno tržište, ali i na znatno niže cijene stanova, zbog čega i umjerenije najmnine čine veći relativni udio vrijednosti nekretnine. Kao što je već navedeno, među 25 lokalnih jedinica za koje bilježimo podatke o mjesečnim medijalnim najamninama, vidno nedostaju podaci za slabije razvijene i brojem stanovnika manje gradove u kontinentalnom dijelu zemlje.

Podaci o odnosu godišnje medijalne najamnine i medijalne cijene stanova, koji su za 2025. godinu dostupni za 25 gradova s dovoljnim brojem kupoprodajnih i najamnih transakcija, korišteni su zatim kao empirijska osnovica za procjenu prosječnog odnosa najamnine i cijene stana na razini hrvatskih županija, dviju makroregija (Primorske i Kontinentalne Hrvatske) te na razini čitave Republike Hrvatske.

Prosječna vrijednost omjera za pojedinu županiju izračunata je korištenjem podataka za sve gradove koji se nalaze unutar te županije i za koje je bilo moguće izračunati omjer najamnine i cijene stana. U slučajevima kada je u pojedinoj županiji postojao samo jedan grad s raspoloživim podatkom, upravo je ta vrijednost korištena kao reprezentativna procjena za županiju. Prosjeci za Primorsku i Kontinentalnu Hrvatsku izračunati su na temelju svih gradova koji pripadaju pojedinoj regiji, dok je prosjek za Republiku Hrvatsku dobiven korištenjem podataka za svih 25 gradova uključenih u analizu.

TABLICA 10. Ostvarene medijalne cijene stanova i najamnine te broj ugovora o kupoprodaji i najmu stanova u 2025.

Jedinica lokalne samouprave	Županija	Broj kupoprodajnih ugovora	Medijalna cijena stana (EUR/m ² , s PDV-om)	Broj ugovora o najmu stana	Medijalna mjesečna najamnina (EUR/m ²)	Omjer godišnje medijalne najamnine i cijene stana
Primorska Hrvatska						
Dubrovnik	Dubrovačko-neretvanska	123	3.797	25	15,0	0,047
Poreč	Istarska	188	2.762	63	10,4	0,045
Pula	Istarska	364	2.459	36	9,1	0,044
Rovinj	Istarska	122	3.483	148	10,9	0,037
Labin	Istarska	43	2.716	75	8,8	0,039
Pazin	Istarska	21	2.172	84	7,0	0,039
Šibenik	Šibensko-kninska	154	2.443	273	7,5	0,037
Vodice	Šibensko-kninska	100	2.880	90	6,7	0,028
Knin	Šibensko-kninska	48	1.272	31	3,7	0,035
Rijeka	Primorsko-goranska	1.035	2.789	206	10,5	0,045
Rab	Primorsko-goranska	46	2.267	27	7,3	0,039
Opatija	Primorsko-goranska	92	3.719	27	10,0	0,032
Kaštela	Splitsko-dalmatinska	112	2.575	39	10,0	0,047
Split	Splitsko-dalmatinska	704	4.078	1.45	12,2	0,036
Kontinentalna Hrvatska						
Karlovac	Karlovačka	247	1.771	335	7,9	0,054
Koprivnica	Koprivničko-križevačka	120	2.098	120	7,5	0,043
Zabok	Krapinsko-zagorska	21	1.851	50	8,0	0,052
Slavonski Brod	Brodsko-posavska	131	1.486	204	7,0	0,057
Čakovec	Međimurska	171	1.771	129	6,5	0,044
Osijek	Osječko-baranjska	639	1.660	813	7,9	0,057
Kutina	Sisačko-moslavačka	26	1.283	73	6,0	0,056
Varaždin	Varaždinska	147	2.144	453	8,5	0,048
Velika Gorica	Zagrebačka	391	2.181	416	9,5	0,052
Grad Zagreb						
Grad Zagreb	Grad Zagreb	6.778	2.905	3.044	12,7	0,052

Napomena: medijalne cijene i najamnine uključuju PDV.

Izvor: izračun autorice na temelju podataka iz baze eNekretnine.

Dobiveni rezultati ukazuju na jasnu i sustavnu razliku u odnosu najamnina i cijena stanova između primorskih i kontinentalnih dijelova Hrvatske. Prosječni omjer za Primorsku Hrvatsku iznosi 0,039, što znači da se godišnje najamnine u prosjeku kreću oko 3,9 % vrijednosti stana, dok je u Kontinentalnoj Hrvatskoj taj omjer znatno viši i iznosi 0,051, odnosno oko 5,1 % vrijednosti stana godišnje. Na razini cijele Republike Hrvatske prosječni omjer iznosi 0,044. Ostvareni omjer za Republiku Hrvatsku jasno indicira da on predstavlja prosjek između nižih vrijednosti koje ostvaruje primorske županije i viših vrijednosti omjera medijalnog najma i cijene stana kojeg ostvaruju kontinentalne županije. Baš iz tog razloga, ne bi bilo uputno koristiti nacionalni prosjek omjera za one lokalne jedinice koje ne bilježe medijalnu najamninu, već je uputnije u tu svrhu koristiti ili prosječni županijski omjer (ako postoji) ili prosječni omjer za primorsku ili kontinentalnu makroregiju.

TABLICA 11. Omjer medijalne najamnine i medijalne cijene stana za županije i regije

Regija/županija	Omjer godišnje medijalne najamnine i cijene stana
Primorska Hrvatska	0,039
Dubrovačko-neretvanska	0,047
Istarska	0,040
Šibensko-kninska	0,033
Primorsko-goranska	0,039
Splitsko-dalmatinska	0,041
Kontinentalna Hrvatska	0,051
Karlovačka	0,054
Koprivničko-križevačka	0,043
Krapinsko-zagorska	0,052
Brodsko-posavska	0,057
Međimurska	0,044
Osječko-baranjska	0,057
Sisačko-moslavačka	0,056
Varaždinska	0,048
Zagrebačka	0,052
Grad Zagreb	0,052
Republika Hrvatska	0,044

Izvor: izračun autorice na temelju podataka iz baze eNekretnine.

Na županijskoj razini, najniži omjeri zabilježeni su u Šibensko-kninskoj županiji (0,033) te u Primorsko-goranskoj županiji (0,039), što upućuje na relativno visoke kupoprodajne cijene u odnosu na ostvarene najamnine. Suprotno tome, najviši omjeri zabilježeni su u Brodsko-posavskoj i Osječko-baranjskoj županiji (po 0,057), kao i u Sisačko-moslavačkoj županiji (0,056), gdje najamnine čine znatno veći udio u vrijednosti stanova.

Ovakva prostorna distribucija omjera može se tumačiti kroz strukturne razlike u funkcioniranju tržišta stanovanja. U primorskim područjima visoke kupoprodajne cijene, koje su u velikoj mjeri pod utjecajem turističke potražnje i investicijskih motiva, nisu praćene proporcionalno višim dugoročnim najamninama, što rezultira nižim omjerima najamnine i cijene. Nasuprot tome, u kontinentalnim županijama, gdje su kupoprodajne cijene relativno niže, a potražnja za dugoročnim najmom stabilnija i dominantno vezana uz lokalno stanovništvo, ostvareni omjeri su viši.

Interpretacijski, omjer godišnje najamnine i cijene stana može se promatrati i kao približan pokazatelj godišnjeg "povrata" na stambenu nekretninu iz perspektive dugoročnog najma, odnosno udjela kupoprodajne vrijednosti stana koji se godišnje „otplaćuje“ kroz plaćanje najamnine. U tom smislu, dobiveni rezultati dodatno potvrđuju da je relativna isplativost dugoročnog najma, mjerena kroz ovaj omjer, izraženija u kontinentalnim dijelovima zemlje nego u primorskim, turistički atraktivnim područjima.

Ovi empirijski obrasci čine ključnu osnovu za sljedeću fazu analize, u kojoj se procijenjeni županijski i regionalni omjeri koriste za rekonstrukciju medijalnih najamnina u jedinicama lokalne samouprave koje ne raspolažu izravno opaženim tržišnim podacima o najmu.

6.2. Rezultati procjene medijalnih najamnina

U drugoj fazi postupka izračuna mjesečnih medijalnih najamnina za stanove za sve jedinice lokalne samouprave u Republici Hrvatskoj koristi se stoga kombinacija raspoloživih administrativnih podataka i rezultata modeliranja medijalnih cijena stanova iz prethodnog poglavlja. Polazišna točka ovog postupka bila je činjenica da su izravni podaci o tržištu najma dostupni za izrazito ograničen broj lokalnih jedinica, dok su procijenjene medijalne cijene stanova, dobivene primjenom metoda strojnog učenja, raspoložive za sve jedinice lokalne samouprave.

Kako bi se rezultate procjene medijalnih najamnina čim više približilo stvarnim odnosima na tržištu stambenih nekretnina, za jedinice lokalne samouprave koje su tijekom 2025. godine bilježile 20 ili više kupoprodajnih ugovora stanova korišteni su stvarni, opaženi podaci o medijalnim cijenama stanova, izravno izračunati iz transakcijskih podataka. Za preostale jedinice lokalne samouprave, koje ne zadovoljavaju navedeni prag ili za koje ne

postoje nikakvi transakcijski podaci, korištene su procijenjene medijalne cijene stanova dobivene primjenom modela strojnog učenja opisanih u prethodnom poglavlju.

Nadalje, kako bi se dodatno osigurala konzistentnost i stabilnost procjena, procijenjene medijalne cijene stanova za lokalne jedinice u pojedinim županijama dalje su razvrstane prema razvojnim skupinama temeljenima na indeksu razvijenosti. Za svaku županijsku razvojnu skupinu izračunata je prosječna medijalna cijena stana na osnovi procijenjenih medijalnih cijena za sve jedinice lokalne samouprave koje unutar iste županije pripadaju jednoj razvojnoj skupini, koja je potom korištena kao reprezentativna vrijednost za sve jedinice iste razvojne skupine unutar te županije koje ne bilježe stvarnu medijalnu cijenu stanova ili čija se ostvarena cijena temelji na manje od 20 kupoprodajnih ugovora. Ovakav postupak primijenjen je s ciljem da se procijenjene cijene stanova dobivene strojnim učenjem za jedinice lokalne samouprave iste razvojne skupine što je više moguće ujednače, te da se smanji utjecaj potencijalnih ekstremnih ili nestabilnih procjena na razini pojedinačnih jedinica. Kao što je već naglašeno, za jedinice lokalne samouprave koje ostvaruju 20 ili više kupoprodajnih ugovora zadržana je stvarna, ostvarena medijalna cijena stanova.

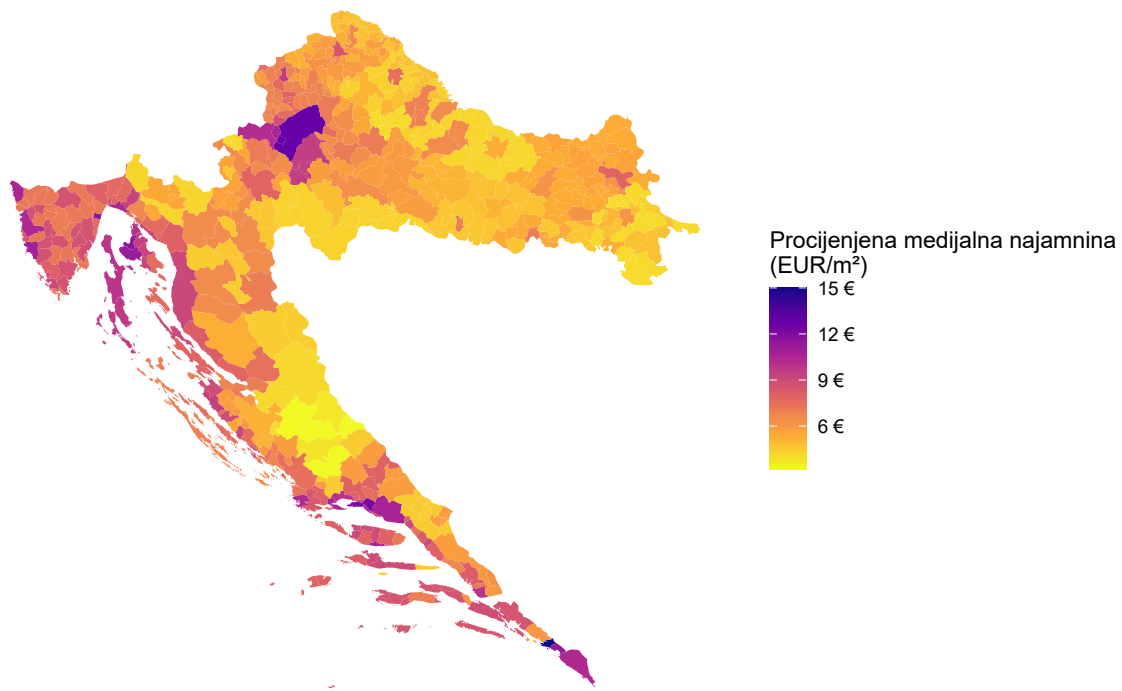
Nakon što je navedenim postupcima dobiven finalni izračun medijalne cijene stanova za sve jedinice lokalne samouprave, pristupilo se procjeni mjesečnih medijalnih najamnina. Procijenjene (ili ostvarene) medijalne cijene stanova podijeljene su s odgovarajućim omjerom godišnje medijalne najamnine i medijalne cijene stana, čime je dobiven procijenjeni godišnji iznos najamnine po metru kvadratnom. Dobivena godišnja vrijednost zatim je podijeljena s 12 kako bi se izračunao procijenjeni mjesečni iznos medijalne najamnine po metru kvadratnom.

Pri tom se vodilo računa o hijerarhiji raspoloživih informacija o omjeru najamnine i cijene. Za gradove koji bilježe 20 ili više ugovora o najmu stanova te istodobno najmanje 20 kupoprodajnih ugovora, ovaj je postupak rezultirao stvarnim iznosom medijalne najamnine prikazanim u Tablici 10, a koji se mogao i izravno izračunati iz transakcijskih podataka. Ako za pojedini grad ili jedinicu lokalne samouprave nije postojao izračunati omjer godišnje medijalne najamnine i cijene, korišten je prosječni omjer za županiju kojoj lokalna jedinica pripada, a koji je prikazan u Tablici 11. U slučajevima kada ni županijski omjer nije bio raspoloživ, primijenjen je prosječni omjer za odgovarajuću makroregiju prikazan također u Tablici 11., pri čemu je lokalna jedinica mogla pripadati ili Primorskoj ili Kontinentalnoj Hrvatska.

Primjenom opisanog postupka utvrđena je medijalna mjesečna najamnina za sve jedinice lokalne samouprave u Republici Hrvatskoj. U 25 lokalnih jedinica prikazanih u Tablici 10. procijenjena medijalna najamnina u potpunosti odgovara stvarnom iznosu medijalne najamnine izračunatom na temelju transakcijskih podataka, dok su za preostale

jedinice najamnine procijenjene na temelju kombinacije modeliranih cijena stanova i empirijski utemeljenih regionalnih i županijskih omjera godišnje medijalne najamnine i medijalne cijene stana. Time je osigurana potpuna prostorna pokrivenost i metodološka konzistentnost procjena što predstavlja ključnu pretpostavku za daljnju analizu i primjenu rezultata u kontekstu provedbe stambenih politika. Procijenjene vrijednosti medijalnih najamnina i stvarne vrijednosti medijalnih najamnina za 25 gradova za koje imamo pouzdane podatke su u potpunosti konzistentne, a sam iznos medijalnih najamnina dobro korespondira i sa stupnjem razvoja lokalne jedinice. Vrijednosti mjesečnih medijalnih najamnina za sve jedinice lokalne samouprave prikazane su na Slici 3, dok se numerički prikaz procijenjenih mjesečnih medijalnih najamnina za sve jedinice lokalne samouprave nalazi u Tablici P3 u Prilogu studiji.

SLIKA 3. Procijenjene mjesečne medijalne najamnine stanova



Izvor: izračun autorice.

U Tablici su dodatno prikazane i vrijednosti mjesečne medijalne najamnine izražene u obliku indeksnih brojeva, pri čemu je bazna vrijednost indeksa definirana kao mjesečna medijalna najamnina u Gradu Zagrebu. Na taj se način razine najamnina u svim ostalim jedinicama lokalne samouprave izražavaju relativno u odnosu na Zagreb, što omogućuje jednostavnu i transparentnu usporedbu prostornih razlika u cijenama najma.

Ovakav indeksni pristup otvara mogućnost da se u sljedeće dvije ili više godina vrijednosti medijalnih najamnina ne procjenjuju iznova primjenom metodološki zahtjevnog i

vremenski intenzivnog postupka strojnog učenja, već da se, polazeći od ažurirane vrijednosti mjesečne medijalne najamnine za Grad Zagreb, procijene najamnine za sve ostale jedinice lokalne samouprave. Drugim riječima, promjene u razini najamnina u Zagrebu mogle bi se koristiti kao referentna osnova za proporcionalno prilagođavanje procijenjenih najamnina u ostatku zemlje, uz zadržavanje prethodno utvrđenih relativnih odnosa među lokalnim tržištima.

Prednost ovakvog pristupa leži u njegovoj operativnoj jednostavnosti, brzini provedbe i stabilnosti rezultata, što je osobito važno u kontekstu redovite primjene procjena u javnim politikama i administrativnim postupcima. Time se smanjuje potreba za čestim ponovnim modeliranjem, a istodobno se osigurava konzistentnost procjena kroz vrijeme.

Istodobno, valja jasno naglasiti i ograničenja ovakvog indeksnog pristupa. Polazišna pretpostavka da se relativni odnosi najamnina između Zagreba i ostalih jedinica lokalne samouprave kratkoročno ne mijenjaju nužno ne mora u potpunosti vrijediti u svim okolnostima. Razlike u dinamici lokalnih tržišta, osobito u turistički izloženim primorskim područjima ili u jedinicama s naglim promjenama potražnje, mogu dovesti do toga da se najamnine u pojedinim sredinama kreću brže ili sporije od zagrebačkog tržišta. U takvim slučajevima indeksna metoda može privremeno podcijeniti ili precijeniti stvarnu razinu najamnina. Stoga se predloženi indeksni pristup može smatrati primjerenim rješenjem za kratkoročno razdoblje, primjerice u sljedeće dvije godine, uz pretpostavku relativne stabilnosti prostorne strukture tržišta najma. U srednjem roku preporučljivo je ponovno provesti cjelovitu procjenu medijalnih najamnina primjenom metoda opisanih u ovoj studiji, kako bi se uvažile moguće strukturne promjene u odnosima među lokalnim tržištima najma i osigurala dugoročna točnost i relevantnost procjena.

Valja napomenuti da je u Prilogu studiji, u Tablici P4., prikazana i procjena mjesečnih medijalnih najamnina za katastarske općine Grada Zagreba. Iako takva razina prostorne razrade nije bila predviđena projektnim zadatkom, procjene su izrađene kako bi se omogućila detaljnija primjena rezultata u kontekstu javnih politika i operativnih mjera na području Grada Zagreba. Međutim, procjena medijalnih mjesečnih najamnina na razini katastarskih općina se, nažalost, ne može provesti primjenom istog postupka strojnog učenja koji je korišten za jedinice lokalne samouprave.

Kao što je već navedeno u prethodnom poglavlju, razlog tome je strukturna neusklađenost dostupnih podataka: dok se transakcijski podaci o kupoprodaji i najmu stanova u bazi eNekretnine administrativno vode po katastarskim općinama, svi ostali relevantni statistički podaci o ekonomskoj, demografskoj, turističkoj i građevinskoj aktivnosti te o stambenom fondu dostupni su isključivo na razini gradskih naselja. Ovakva razlika u prostornoj razini podataka onemogućuje izravnu primjenu strojnog učenja i posljedičnu procjenu medijalnih cijena i najamnina stanova po katastarskim općinama. Da bi se

takav pristup mogao metodološki ispravno provesti, bilo bi nužno sve transakcije iz baze eNekretnine, korištenjem detaljnih adresnih podataka, jednoznačno povezati s pripadajućim gradskim naseljima, što nadilazi opseg i vremenski okvir ovog projekta.

TABLICA 12. Medijalne cijene stanova i najamnina po katastarskim općinama Grada Zagreba

Katastarska općina	Broj kupoprodajnih ugovora	Medijalna cijena stana (EUR/m ²)	Broj ugovora o najmu	Medijalna najamnina (EUR/m ²)
Blato	711	2.769	375	14,20
Centar	332	3.379	71	14,20
Centar novi	417	3.246	142	14,55
Črnomerec	288	3.125	72	12,65
Dubrava	172	2.694	79	12,00
Donje Vrapče	321	3.158	132	12,75
Jakuševac*	-	-	109	13,80
Klara	452	2.496	350	12,50
Klara nova	87	2.581	117	13,10
Lučko	17	2.181	29	11,00
Odra	106	2.367	82	10,70
Peščenica	642	3.090	87	12,70
Rudeš	141	3.049	77	13,30
Sesvetski Kraljevec	149	1.964	22	9,90
Stenjevec	265	2.950	61	10,80
Stenjevec jug	258	3.201	35	13,20
Trešnjevka	166	3.066	100	13,90
Trešnjevka nova	185	3.029	99	12,40
Trnje	321	3.263	126	12,90
Vrapče	137	2.980	41	10,70
Zaprudski otok	419	2.928	636	12,25
Žitnjak	38	2.819	30	13,55

Napomena: * - ne postoji dovoljan broj kupoprodajnih ugovora za izračun medijalne cijene stana.

Izvor: izračun autorice na temelju podataka iz baze eNekretnine.

S obzirom na navedena ograničenja, mjesečne medijalne najamnine za katastarske općine Grada Zagreba procijenjene su primjenom hijerarhijskog i pragom uvjetovanog pristupa koji u najvećoj mogućoj mjeri koristi stvarno ostvarene transakcijske podatke. U katastarskim općinama koje su u 2025. godini bilježile 20 ili više zabilježenih ugovora o najmu stanova, kao mjesečna medijalna najamnina korištena je stvarno ostvarena mjesečna medijalna vrijednost izračunata iz podataka dostupnih u bazi eNekretnine. Ti podaci su ujedno i prikazani u Tablici 12. U slučajevima kada katastarska općina nije

imala dovoljan broj ugovora o najmu, ali je bilježila 20 ili više kupoprodajnih transakcija stanova, mjesečna medijalna najamnina procijenjena je na temelju kvocijenta stvarno ostvarene medijalne cijene stana u toj katastarskoj općini i udjela godišnje medijalne najamnine u medijalnoj cijeni stana za Grad Zagreb prikazane u Tablici 10. Konačno, za katastarske općine koje nisu imale niti dovoljan broj ugovora o najmu niti dovoljan broj kupoprodajnih transakcija, mjesečna medijalna najamnina procijenjena je imputacijom vrijednosti iz prostorno najbliže katastarske općine koja raspolaže stvarno zabilježenom ili procijenjenom medijalnom najamninom. Ovakvim postupkom osigurana je potpuna prostorna pokrivenost Grada Zagreba procijenjenim vrijednostima najamnina, uz dosljedno uvažavanje hijerarhije pouzdanosti dostupnih podataka i minimiziranje proizvoljnosti procjena.

7. Procjena prosječnih cijena stanova za jedinice lokalne samouprave primjenom metoda strojnog učenja

U ovom poglavlju provodi se strojno učenje modela prosječnih cijena stanova bez poreza na dodanu vrijednost (PDV) za sve jedinice lokalne samouprave u Republici Hrvatskoj, s ciljem dobivanja prostorno konzistentnih i metodološki utemeljenih predikcija prosječnih cijena stanova. Dobivene procjene predstavljaju ključnu analitičku osnovu za provedbu Programa potpora za stjecanje prve stambene nekretnine, i to konkretno kao osnovica, odnosno referentna tržišna cijena stambene nekretnine, za utvrđivanje iznosa povrata poreza na promet nekretnina i PDV-a plaćenih prilikom stjecanja prve nekretnine.

Sukladno važećim propisima i uputama Ministarstva prostornoga uređenja, graditeljstva i državne imovine koje provodi Program, u navedenu osnovicu uključene su cijene starijih stanova te cijene novoizgrađenih stanova, pri čemu se iz cijene novoizgrađenih stanova izuzima PDV. U izračun prosječne cijene uključeni su isključivo stanovi, dok obiteljske kuće nisu obuhvaćene analizom, budući da je kupoprodajna cijena obiteljskih kuća dostupna u bazi eNekretnine nepouzdana zbog nepostojanja konzistentnih i usporedivih podataka o korisnoj površini u prostornim i vlasničkim registrima. Kao i kod medijalnih cijena, za izračun prosječnih cijena korišteni su podaci iz baze eNekretnine za prvih 11 mjeseci 2025. godine.

Kao i u poglavlju u kojem se strojno učenje koristi za procjenu medijalnih cijena stanova s PDV-om, i u ovom je poglavlju analiza strukturirana kao višefazni postupak koji započinje konstrukcijom prostornog pomaka ciljne varijable radi eksplicitnog uvođenja prostorne međuovisnosti, nastavlja se definiranjem uzorka za treniranje i testiranje modela, sustavnom usporedbom različitih modela strojnog učenja radi identifikacije ključnih objašnjavajućih varijabli, te završava izgradnjom parsimonioznog i interpretabilnog zaključnog modela prikladnog za operativnu primjenu u kontekstu provedbe Programa potpora za stjecanje prve stambene nekretnine.

Važno je pritom naglasiti da medijalne cijene stanova s uključenim PDV-om i prosječne cijene stanova bez PDV-a koje su dostupne za 2025. godinu imaju identične strukturne karakteristike u pogledu prostorne dostupnosti podataka po jedinicama lokalne samouprave, budući da se u oba slučaja temelje na istom skupu kupoprodajnih transakcija evidentiranih u sustavu eNekretnine, uz primjenu jednakih pragova minimalnog broja transakcija za izravni izračun pokazatelja. Istodobno, deskriptivna analiza pokazuje da su obrasci varijabilnosti cijena među jedinicama lokalne samouprave vrlo slični u slučaju medijalnih cijena s PDV-om i prosječnih cijena bez PDV-a, kako u pogledu raspona vrijednosti, tako i u pogledu njihove prostorne distribucije.

S obzirom na navedeno, opravdano je očekivati da će i rezultati strojnog učenja u ovom poglavlju – osobito kada je riječ o konstrukciji prostornog pomaka ciljane varijable, identifikaciji ključnih objašnjavajućih varijabli te izboru zaključnog prediktivnog modela za projekciju prosječnih cijena stanova – biti u velikoj mjeri usporedivi s rezultatima dobivenima u analizi medijalnih cijena stanova. Unatoč tome, svi koraci modeliranja i evaluacije provedeni su neovisno i u cijelosti, kako bi se osigurala metodološka dosljednost i empirijska utemeljenost konačnih procjena prosječnih cijena stanova bez PDV-a.

7.1. Konstrukcija prostornog pomaka ciljane varijable

Konstrukcija prostornog pomaka prosječne cijene stana bez PDV-a provedena je na identičan način kao i u poglavlju o procjeni medijalnih cijena stanova, te se u potpunosti oslanja na tamo detaljno opisanu metodologiju. Prostorni pomak uveden je kao objašnjavajuća varijabla s ciljem eksplicitnog uvođenja prostorne međuovisnosti u modele procjene cijena stanova, polazeći od pretpostavke da se lokalna tržišta nekretnina formiraju u snažnoj interakciji s prostorno bliskim jedinicama lokalne samouprave.

U prvom koraku prostorni pomak izračunat je kao aritmetička sredina prosječnih cijena stanova bez PDV-a u deset prostorno najbližih jedinica lokalne samouprave, pri čemu je prostorna bliskost definirana na temelju geografskih koordinata centroida JLS. Kao i u slučaju medijalnih cijena, u izračun su uključene isključivo one susjedne jedinice za koje je ciljane varijabla bila dostupna, dok su susjedi bez zabilježenih cijena zanemareni. Ako barem jedna od deset najbližih jedinica ima poznatu vrijednost prosječne cijene stana, prostorni pomak može se izravno izračunati.

Primjena ovog postupka rezultirala je time da za 75 jedinica lokalne samouprave prostorni pomak nije bilo moguće izravno izračunati, budući da niti jedna od njihovih deset prostorno najbližih susjednih jedinica nije imala dostupne podatke o prosječnoj cijeni stana. Riječ je o istom skupu jedinica kao i u analizi medijalnih cijena stanova, što dodatno potvrđuje strukturnu sličnost dostupnosti podataka za obje ciljane varijable.

Kako bi se osigurala potpuna prostorna pokrivenost i omogućilo korištenje prostornog pomaka u svim modelima strojnog učenja, za navedene jedinice provedena je imputacija prostornog pomaka primjenom KNN regresije, ponovno u skladu s metodološkim postupkom opisanim u petom poglavlju. Imputacija je temeljena na geografskim koordinatama centroida JLS, pri čemu je vrijednost prostornog pomaka određena kao aritmetička sredina vrijednosti prostornog pomaka u do 20 prostorno najbližih jedinica s poznatom vrijednošću prostornog pomaka.

Dobiveni rezultati u potpunosti su usporedivi s onima iz analize medijalnih cijena stanova, kako u pogledu broja jedinica za koje je bila potrebna imputacija, tako i u pogledu prostorne strukture i raspona vrijednosti prostornog pomaka. To potvrđuje stabilnost korištenog postupka te dodatno opravdava njegovu primjenu u procjeni prosječnih cijena stanova bez PDV-a.

7.2. Formiranje analitičkog uzorka i podjela na trening i test skup

Proces treniranja i evaluacije modela strojnog učenja započeo je definiranjem analitičkog uzorka jedinica lokalne samouprave za koje je moguće konstruirati statistički dovoljno pouzdanu ciljanu varijablu – prosječnu cijenu stana bez PDV-a. Kao i u analizi medijalnih cijena stanova, ciljana varijabla u ovom je poglavlju izravno izračunata isključivo za one jedinice lokalne samouprave koje u 2025. godini bilježe najmanje 30 kupoprodajnih transakcija stanovima. Odabir ovog praga ima za cilj osigurati da su izravno izračunate prosječne cijene dovoljno stabilne, reprezentativne i robusne da mogu poslužiti kao pouzdana osnova za treniranje i evaluaciju prediktivnih modela.

Primjenom navedenog kriterija, iz ukupnog skupa od 556 jedinica lokalne samouprave formiran je analitički poduzorak od 74 jedinice za koje su istodobno dostupni izravno izračunata prosječna cijena stana bez PDV-a u 2025. godini i sve objašnjavajuće varijable uključene u početni prediktivni model. Riječ je o istom skupu jedinica koji je korišten i u analizi medijalnih cijena stanova, što dodatno potvrđuje strukturnu podudarnost dostupnosti podataka za obje ciljane varijable.

Distribucija broja kupoprodajnih transakcija u cjelokupnom skupu jedinica lokalne samouprave dodatno opravdava primjenu ovako restriktivnog kriterija uključivanja u analitički uzorak. Medijan broja transakcija među svim jedinicama lokalne samouprave iznosi svega 30,5, dok čak **414 jedinica ne bilježi niti jednu kupoprodajnu transakciju stanovima u promatranoj godini**, a velik dio preostalih jedinica bilježi vrlo mali broj transakcija. Uključivanje takvih jedinica u proces treniranja dovelo bi do učenja modela na statistički nestabilnim i potencijalno pristranim vrijednostima ciljane varijable, čime bi se značajno narušila pouzdanost dobivenih predikcija.

Nakon definiranja analitičkog uzorka, provedena je nasumična podjela podataka na trening i test (holdout) skup u omjeru 80 % prema 20 %. Od ukupno 74 jedinice lokalne samouprave uključene u analizu, 59 jedinica korišteno je za treniranje modela, dok je 15 jedinica rezervirano isključivo za evaluaciju prediktivne uspješnosti. Podjela je provedena korištenjem funkcije `initial_split` iz paketa `rsample`, uz fiksirano sjeme generatora slučajnih brojeva, čime je osigurana potpuna replikabilnost rezultata.

7.3. Usporedba performansi modela strojnog učenja sa proširenim skupom objašnjavajućih varijabli radi identifikacije ključnih varijabli za predikciju medijalnih cijena stanova

U inicijalnoj fazi primjene metoda strojnog učenja provedena je usporedna procjena sedam različitih modela s ciljem identifikacije objašnjavajućih varijabli koje najviše doprinose preciznoj predikciji prosječne cijene stanova bez PDV-a, a ne s ciljem neposrednog odabira konačnog, optimalnog prediktivnog modela. Svi modeli trenirani su na istom prostorno proširenom početnom modelu koji uključuje 20 objašnjavajućih varijabli definiranih u prethodnom poglavlju, uključujući i konstruirani prostorni pomak ciljane varijable. Treniranje modela provedeno je na trening uzorku od 59 jedinica lokalne samouprave koje su u 2025. godini ostvarile najmanje 30 kupoprodajnih transakcija stanovima, dok je testni (holdout) uzorak od 15 jedinica korišten kao zajednički referentni okvir za izračun standardnih mjera prediktivne uspješnosti svih modela.

U ovoj fazi analize mjere poput MAE, RMSE, R^2 i prilagođenog R^2 ne interpretiraju se kao kriteriji za izbor “najboljeg” modela, već kao pomoćni alati za procjenu stabilnosti, robusnosti i informativnosti pojedinih objašnjavajućih varijabli unutar različitih modelskih struktura. Na trening uzorku procijenjeni su modeli opisani u metodološkom poglavlju studije: višestruka linearna regresija, penalizirane regresijske metode (LASSO, Ridge i Elastic Net s 10-strukom unakrsnom validacijom), Regression Tree, Random Forest (s 1.000 stabala) te XGBoost model. Svaki od navedenih pristupa pruža komplementarne informacije o relativnoj važnosti pojedinih prediktora, osobito u kontekstu multikolinearnosti, nelinearnih odnosa i prostorne međuovisnosti.

Dobiveni rezultati (Tablica 13.) pokazuju da penalizirane linearne metode, osobito LASSO i Elastic Net, ostvaruju najpovoljniji odnos između pogreške predikcije i objašnjene varijance, što ih čini osobito pogodnima za redukciju dimenzionalnosti i selekciju varijabli u uvjetima relativno malog uzorka i velikog broja strukturno povezanih prediktora. Istodobno, stablasti i ansambl modeli potvrđuju postojanje nelinearnih obrazaca u podacima, ali u ovoj fazi ne ostvaruju prednost u pogledu ukupne prediktivne točnosti. Svrha ove faze analize stoga nije identifikacija konačnog modela za projekciju prosječnih cijena stanova, već sustavno sužavanje početnog skupa od 20 objašnjavajućih varijabli na manji, informacijski najbogatiji podskup, koji će u sljedećem koraku poslužiti kao temelj za izgradnju parsimonioznog zaključnog prediktivnog modela.

Kao i u analizi medijalnih cijena stanova, rezultati ove usporedne procjene potvrđuju da se ključne informacije o strukturi tržišta nekretnina koncentriraju u relativno malom broju varijabli, dok dodatne varijable u punom modelu prvenstveno doprinose multikolinearnosti i povećanju varijance procjena, bez proporcionalnog dobitka u prediktivnoj moći.

TABLICA 13. Usporedba performansi modela s punim skupom od 20 objašnjavajućih varijabli

Model	MAE	RMSE	R ²	Prilagođeni R ²
Višestruka linearna regresija	346,8	517,3	0,631	1,862
LASSO	336,7	499,0	0,657	1,801
Ridge	359,5	532,5	0,609	1,913
Elastic Net	332,7	497,9	0,658	1,798
Regression Tree	483,3	590,3	0,519	2,122
Random Forest	416,3	551,2	0,581	1,978
XGBoost	423,2	552,7	0,579	1,983

Napomena: ciljane varijabla prosječna cijena stanova bez PDV-a po m²; MAE = srednja apsolutna pogreška; RMSE = korijen srednje kvadratne pogreške; R² – koeficijent determinacije; prilagođeni R² – prilagođeni koeficijent determinacije.

Izvor: izračun autorice.

Kao što je već naglašeno u petom poglavlju studije, cilj ove faze analize nije identifikacija najboljeg pojedinačnog prediktivnog modela, već sustavna i višemodelska procjena prediktivne važnosti objašnjavajućih varijabli, s isključivim ciljem redukcije početnog skupa od 20 varijabli na manji, informacijski najbogatiji podskup. Isti analitički pristup primijenjen je i u procjeni prosječnih cijena stanova bez PDV-a, budući da su struktura dostupnosti podataka i obrasci prostorne varijabilnosti cijena vrlo slični onima zabilježenima kod medijalnih cijena stanova.

Procjena važnosti objašnjavajućih varijabli provedena je korištenjem istih komplementarnih metoda i kriterija kao u analizi medijalnih cijena, a detaljan opis metodologije izračuna pojedinih mjera važnosti (penalizirane regresije, Random Forest i XGBoost) dan je u petom poglavlju. Ovdje se stoga naglasak stavlja isključivo na rezultate i njihovu interpretaciju, a ne na ponovno izlaganje tehničkih detalja.

Dobivene mjere važnosti objedinjene su u jedinstvenu tablicu (Tablica 14.), pri čemu su, radi osiguravanja usporedivosti među metodama, nedostajuće vrijednosti u penaliziranim regresijskim modelima postavljene na nulu. Budući da su apsolutne vrijednosti pojedinih mjera po svojoj skali međusobno neusporedive, rezultati se u nastavku interpretiraju prvenstveno kroz relativni položaj i konzistentnost varijabli među različitim modelima, a ne kroz njihove apsolutne vrijednosti. Svrha ove faze analize je identifikacija varijabli koje se sustavno pojavljuju kao informativne kroz više modelskih pristupa te koje će u sljedećem koraku poslužiti kao temelj za izgradnju parsimonioznih prediktivnih modela prosječnih cijena stanova.

TABLICA 14. Višemodelska procjena važnosti objašnjavajućih varijabli

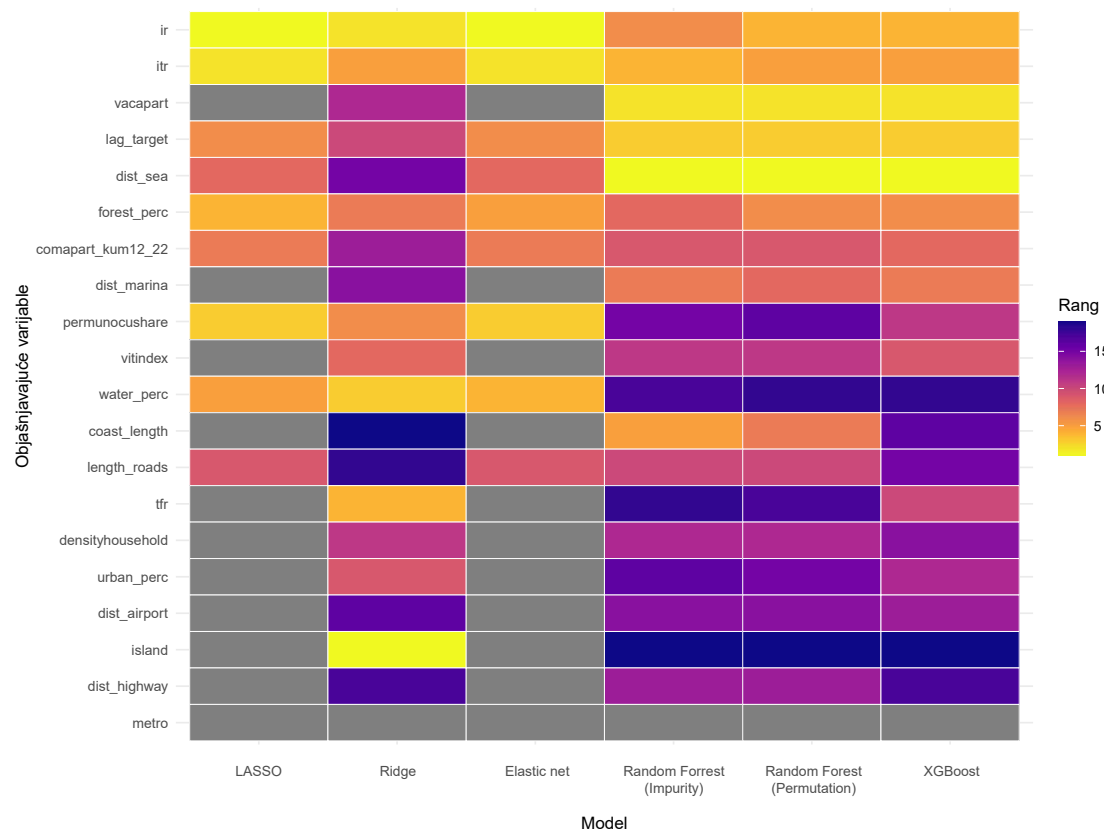
Objašnjavajuća varijabla	LASSO	Ridge	Elastic Net	Random Forest (impurity)	Random Forest (permutation)	XGBoost (gain)
dist_sea	0,001	0,001	0,001	3.797.194	104,621	0,292
vacapart	0,000	0,062	0,000	3.338.401	90,728	0,192
lag_target	0,235	0,202	0,242	2.961.621	80,565	0,132
ir	39,816	31,670	38,511	1.764.651	79,226	0,101
itr	34,023	18,835	32,338	2.618.281	63,242	0,063
forest_perc	5,015	4,140	4,687	1.465.838	53,251	0,060
dist_marina	0,000	0,001	0,000	1.678.091	31,214	0,037
comapart_kum12_22	0,038	0,051	0,037	769.878	19,776	0,027
vitindex	0,000	0,361	0,000	574.130	16,166	0,018
tfr	0,000	21,355	0,000	381.751	10,311	0,015
permunocushare	8,652	4,919	7,724	454.869	11,546	0,012
urban_perc	0,000	0,209	0,000	440.656	12,744	0,011
dist_airport	0,000	0,001	0,000	469.821	12,778	0,010
densityhousehold	0,000	0,110	0,000	550.542	13,166	0,007
length_roads	0,000	0,000	0,000	627.576	17,458	0,007
coast_length	0,000	0,000	0,000	2.603.595	46,087	0,007
dist_highway	0,000	0,001	0,000	494.068	12,817	0,005
water_perc	0,717	29,708	4,724	435.155	4,820	0,003
island	0,000	40,106	0,000	56.003	1,662	0,000
metro	0,000	0,000	0,000	0	0,000	0,000

Napomena: ciljane varijabla prosječna cijena stanova bez PDV-a po m².

Izvor: izračun autorice.

Nakon izračuna apsolutnih mjera važnosti za svaku pojedinu metodu, napravljeno je rangiranje varijabli te je izračunat konsenzusni rang kao aritmetička sredina rangova dobivenih iz svih metoda u kojima je varijabla imala nenultu važnost. Time je dobivena jedinstvena, model-agnostička mjera relativne važnosti objašnjavajućih varijabli, koja ublažava pristranosti pojedinih algoritama i naglašava varijable koje su stabilno visoko rangirane kroz različite metodološke pristupe. Redoslijed važnosti varijabli za objašnjavanje prosječnih cijena stanova bez PDV-a prikazan je na Slici 4., dok su numeričke vrijednosti pojedinačnih rangova i završni konsenzusni rang prikazani u Tablici 15.

SLIKA 4. Rang objašnjavajućih varijabli u izabranim modelima strojnog učenja



Napomena: objašnjavajuće varijable za koje nije izračunat rang jer su kod penalizirajućih regresijskih modela isključene iz procjene su označene sivom bojom. Isto vrijedi i za indikator varijablu metro kod stablastih ansambli modela.

Izvor: izračun autorice.

Rezultati pokazuju da se pet istih varijabli koje su identificirane kao ključne i u analizi medijalnih cijena stanova s PDV-om ponovno izdvajaju kao najinformativnije i u objašnjavanju prosječnih cijena stanova bez PDV-a. Razlika se pritom očituje isključivo u redoslijedu njihove važnosti prema konsenzusnom rangui, dok sam skup varijabli ostaje nepromijenjen.

Najvišu važnost u ovom slučaju ima **indeks razvijenosti (ir)**, koji odražava ukupnu socioekonomsku razvijenost i platežnu moć lokalnog stanovništva, a odmah iza njega slijedi **indeks turističke razvijenosti (itr)**, koji upućuje na snažan utjecaj turističke potražnje na razinu cijena stanova. Treće mjesto zauzima **udio praznih stanova (vacapart)**, pokazatelj strukture i iskorištenosti stambenog fonda, dok se na četvrtom mjestu nalazi **prostorni pomak prosječne cijene stana (lag_target)**, kojim se u model eksplicitno uvodi prostorna međuovisnost i prenosi informacija o regionalnim cjenovnim obrascima. Petu najvažniju varijablu čini **udaljenost od mora (dist_sea)**, koja hvata snažan prostorni gradijent cijena povezan s obalnom i turističkom atraktivnošću.

Činjenica da se isti skup varijabli pojavljuje među najvažnijima u oba modela – i za medijalne i za prosječne cijene stanova – dodatno potvrđuje stabilnost i robusnost identificiranih determinanti cijena. Na temelju tog konsenzusnog rangiranja donesena je odluka da se daljnja faza modeliranja fokusira upravo na ovih pet varijabli i njihove kombinacije, čime se osigurava da redukcija modela bude vođena empirijskom dosljednošću i robusnošću rezultata, a ne specifičnostima pojedinog algoritma.

TABLICA 15. Rangiranje važnosti varijabli prema pojedinim modelima i konsenzusnom rang

Objašnjavajuća varijabla	LASSO	Ridge	Elastic Net	Random Forest (impurity)	Random Forest (permutation)	XGBoost	Konsenzusni rang
ir	1	2	1	6	4	4	3,0
itr	2	5	2	4	5	5	3,8
vacapart	-	12	-	2	2	2	4,5
lag_target	6	10	6	3	3	3	5,2
dist_sea	8	15	8	1	1	1	5,7
forest_perc	4	7	5	8	6	6	6,0
comapart_kum12_22	7	13	7	9	9	8	8,8
dist_marina	-	14	-	7	8	7	9,0
permunocushare	3	6	3	15	16	11	9,0
vitindex	-	8	-	11	11	9	9,8
water_perc	5	3	4	17	18	18	10,8
coast_length	-	19	-	5	7	16	11,8
length_roads	9	18	9	10	10	15	11,8
tfr	-	4	-	18	17	10	12,3
densityhousehold	-	11	-	12	12	14	12,3
urban_perc	-	9	-	16	15	12	13,0
dist_airport	-	16	-	14	14	13	14,3
island	-	1	-	19	19	19	14,5
dist_highway	-	17	-	13	13	17	15,0
metro	-	-	-	-	-	-	-

Napomena: ciljana varijabla prosječna cijena stanova bez PDV-a po m².

Izvor: izračun autorice.

7.4. Identifikacija ključnih varijabli za predikciju medijalnih cijena stanova

Nakon utvrđivanja konsenzusnog rangiranja važnosti objašnjavajućih varijabli, provedena je sustavna analiza parsimonioznih specifikacija modela s ciljem ispitivanja u kojoj se mjeri prediktivna uspješnost može očuvati uz znatno manji broj varijabli. Polazeći od pet varijabli identificiranih kao ključne (ir, itr, vacapart, lag_target i dist_sea), generirane su sve kombinacije veličine dvije do pet varijabli te su za svaku kombinaciju procijenjeni svi razmatrani modeli strojnog učenja.

Ukupno je procijenjeno 182 modela (26 kombinacija varijabli za svaku od sedam metoda), pri čemu su u svim slučajevima korišteni isti trening i testni uzorci, čime je osigurana potpuna usporedivost rezultata. Detaljni rezultati svih 182 procjena dostupni su u Tablici P2 u Prilogu studiji, dok su u nastavku prikazani sažeti rezultati najboljih specifikacija po pojedinoj metodi (Tablica 16.).

TABLICA 16. Najbolje specifikacije modela prema različitim kriterijima uspješnosti

Model stojnog učenja	Najbolji MAE	Objašnjavajuće varijable	Najbolji RMSE	Objašnjavajuće varijable	Najbolji prilagođeni R ²	Objašnjavajuće varijable
Višestruka linearna regresija	282,7	ir, itr, vacapart, lag_target, dist_sea	438,4	ir, itr, vacapart, lag_target, dist_sea	0,645	ir, itr, lag_target
LASSO	276,2	ir, itr, vacapart, lag_target, dist_sea	442,6	ir, itr, lag_target, dist_sea	0,645	ir, itr, lag_target
Ridge	278,6	ir, itr, lag_target	448,9	ir, itr, vacapart, lag_target, dist_sea	0,641	ir, itr, lag_target
Elastic Net	277,1	ir, itr, vacapart, lag_target	446,7	ir, itr, lag_target, dist_sea	0,645	ir, itr, lag_target
Random Forest	350,0	ir, itr, lag_target	486,3	ir, vacapart	0,619	ir, vacapart
Regression Tree	403,3	ir, itr	516,3	ir, itr	0,571	ir, itr
XGBoost	352,8	ir, lag_target, dist_sea	465,5	ir, vacapart, lag_target	0,619	ir, vacapart, lag_target

Napomena: ciljana varijabla prosječna cijena stanova bez PDV-a po m².

Izvor: izračun autorice.

Rezultati potvrđuju da se visoka razina prediktivne uspješnosti može postići već s tri ili četiri objašnjavajuće varijable, osobito kod linearnih i penaliziranih regresijskih modela. U tim se modelima kao stabilna i robusna kombinacija izdvajaju varijable indeks razvijenosti (ir), indeks turističke razvijenosti (itr) i prostorni pomak prosječne cijene stana (lag_target), dok dodavanje varijabli udaljenosti od mora (dist_sea) i udjela

praznih stanova (vacapart) dodatno poboljšava rezultate prema kriterijima MAE i RMSE. Nelinearni modeli (Random Forest, regresijsko stablo i XGBoost) u pravilu postižu najbolje rezultate s manjim brojem varijabli, ali uz slabiju ukupnu prediktivnu točnost i manju interpretabilnost.

Kao i u analizi medijalnih cijena stanova, za izbor zaključnog modela korišten je RMSE kao ključni kriterij, budući da snažnije penalizira velika odstupanja između stvarnih i predviđenih cijena te je osobito prikladan u kontekstu službenih procjena i primjene u javnim politikama. Usporedba parsimonioznih specifikacija pokazuje da najnižu vrijednost RMSE-a ostvaruje višestruki linearni regresijski model s pet objašnjavajućih varijabli.

Prema tom kriteriju, kao zaključni model za predikciju **prosječnih cijena stanova bez PDV-a** odabran je linearni model koji uključuje sljedeće varijable: **indeks razvijenosti (ir), indeks turističke razvijenosti (itr), udio praznih stanova (vacapart), prostorni pomak prosječne cijene stana (lag_target) i udaljenost od mora (dist_sea)**. Ova specifikacija postiže najbolju prediktivnu točnost među svim razmatranim modelima, uz istodobno zadržavanje visoke interpretabilnosti i stabilnosti procjena.

Važno je naglasiti da se četiri od pet uključenih varijabli (ir, itr, lag_target i dist_sea) podudaraju s varijablama korištenima u zaključnom modelu za predikciju medijalnih cijena stanova s PDV-om, dok se udio praznih stanova (vacapart) pojavljuje kao dodatna varijabla koja poboljšava prediktivnu uspješnost modela za prosječne cijene, ali nije bila uključena u zaključnu specifikaciju za medijalne cijene. Ova razlika upućuje na nešto izraženiju osjetljivost prosječnih cijena na strukturu i iskorištenost stambenog fonda, dok temeljni ekonomski i prostorni mehanizmi formiranja cijena ostaju konzistentni između dvaju modela.

7.5. Evaluacija modela medijalnih cijena stanova s izabranim objašnjavajućim varijablama

U završnoj fazi analize evaluiran je zaključni skup objašnjavajućih varijabli identificiran u prethodnim koracima, koji uključuje pet varijabli: indeks razvijenosti (ir), indeks turističke razvijenosti (itr), prostorni pomak prosječne cijene stana (lag_target), udaljenost od mora (dist_sea) i udio praznih stanova (vacapart). Ova specifikacija odabrana je na temelju kriterija RMSE-a, budući da upravo ona postiže najnižu prediktivnu pogrešku među svim razmatranim parsimonioznim modelima.

Svi razmatrani modeli strojnog učenja – višestruka linearna regresija, LASSO, Ridge i Elastic Net, regresijsko stablo, Random Forest i XGBoost – procijenjeni su s identičnim

skupom objašnjavajućih varijabli te evaluirani na izdvojenom testnom uzorku primjenom standardnih mjera uspješnosti (MAE, RMSE, R^2 i prilagođeni R^2). Takav pristup omogućuje izravnu i transparentnu usporedbu performansi različitih algoritama.

Rezultati evaluacije (Tablica 17.) jasno pokazuju da linearni i penalizirani regresijski modeli ostvaruju znatno bolje prediktivne performanse u odnosu na nelinearne modele. Višestruka linearna regresija postiže najnižu vrijednost RMSE-a, uz istodobno visoku objašnjenu varijancu ($R^2 \approx 0,73$) i stabilan prilagođeni R^2 . Penalizirani modeli (LASSO i Elastic Net) ostvaruju vrlo slične rezultate, dok Ridge regresija pokazuje nešto slabije performanse prema svim kriterijima.

Nelinearni modeli, osobito regresijsko stablo, Random Forest i XGBoost, bilježe osjetno veće vrijednosti MAE-a i RMSE-a te niže vrijednosti R^2 , što upućuje na slabiju generalizaciju u uvjetima ograničenog broja opažanja. U skladu s unaprijed definiranim kriterijem i potrebom za visokom interpretabilnošću i stabilnošću procjena, višestruka linearna regresija s pet objašnjavajućih varijabli odabrana je kao završni model za projekciju prosječnih cijena stanova bez PDV-a.

TABLICA 17. Usporedba modela za završnu specifikaciju (ir, itr, lag_target, dist_sea, vacapart)

Model	MAE	RMSE	R^2	Prilagođeni R^2	Objašnjavajuće varijable
Višestruka linearna regresija	282,74	438,43	0,735	0,588	ir, itr, lag_target, dist_sea, vacapart
LASSO	275,76	445,92	0,726	0,573	ir, itr, lag_target, dist_sea, vacapart
Ridge	291,59	460,90	0,707	0,544	ir, itr, lag_target, dist_sea, vacapart
Elastic Net	279,43	450,39	0,720	0,565	ir, itr, lag_target, dist_sea, vacapart
Random Forest	398,78	524,66	0,620	0,409	ir, itr, lag_target, dist_sea, vacapart
Regression Tree	483,31	590,34	0,519	0,252	ir, itr, lag_target, dist_sea, vacapart
XGBoost	384,37	519,89	0,627	0,420	ir, itr, lag_target, dist_sea, vacapart

Napomena: ciljana varijabla prosječna cijena stanova bez PDV-a po m².

Izvor: izračun autorice.

7.6. Zaključni model za procjenu medijalnih cijena stanova u jedinicama lokalne samouprave i rezultati procjene modela

Kao i kod modela za medijalne cijene stanova s uključenim PDV-om, u završnom koraku ovog dijela analize za prosječne cijene stanova bez PDV-a treniran je model koji se koristi za predikciju prosječnih cijena stanova za sve jedinice lokalne samouprave u Republici Hrvatskoj, uključujući i one JLS-ove za koje u promatranoj godini ne postoji dovoljan broj tržišnih transakcija. Time se analitički okvir pomiče s faze evaluacije i usporedbe modela prema njihovoj operativnoj primjeni u svrhu izrade sveobuhvatnih i konzistentnih procjena na nacionalnoj razini.

Najprije je formiran trening uzorak koji uključuje samo one JLS-ove za koje je raspoloživa opažena prosječna cijena stana (ciljna varijabla) i koje imaju najmanje 30 evidentiranih kupoprodajnih transakcija. Time se osigurava da se procjena modela temelji na dovoljno pouzdanim i reprezentativnim tržišnim signalima. Istodobno je zadržan i puni skup JLS-ova, koji se koristi isključivo u fazi predikcije kako bi se omogućila procjena prosječnih cijena i za jedinice s ograničenom ili nepostojećom tržišnom aktivnošću.

U ovoj završnoj fazi analize procijenjen je **linearni regresijski model**, čija je specifikacija temeljena na **skupu objašnjavajućih varijabli identificiranih u prethodnim fazama analize** kroz sustavna testiranja različitih kombinacija varijabli i algoritama strojnog učenja. Prethodni rezultati pokazali su da upravo određena kombinacija varijabli dosljedno daje stabilne i interpretabilne rezultate kroz različite modele, zbog čega je ona odabrana kao jezgra zaključnog prediktivnog okvira.

Kako bi se dodatno unaprijedile prediktivne sposobnosti modela te uvažila prostorna i institucionalna heterogenost jedinica lokalne samouprave, osnovni linearni regresijski model proširen je u **linearni mješoviti model (linear mixed-effects model)** (Gelman i Hill, 2007). Kao što je već objašnjeno u petom poglavlju, ovakva specifikacija omogućuje istodobno modeliranje utjecaja ključnih objašnjavajućih varijabli kroz fiksne efekte, uz dodatno uvažavanje razvojne hijerarhijske strukture podataka putem slučajnih efekata. Ova specifikacija se, nažalost, nije mogla direktno trenirati, testirati i uspoređivati uz ostale metode strojnog učenja jer ostale metode ne podržavaju treniranje modela koji se djelomično zasniva na slučajnom učinku, odnosno na uvođenju slučajnog odsječka u specifikaciju modela.

Ciljna varijabla u linearnom mješovitom modelu definirana je kao prirodni logaritam prosječne cijene stana, čime se stabilizira varijanca i omogućuje interpretacija koeficijenata u približno relativnim (postotnim) terminima. Model uključuje pet fiksnih učinaka koji su u prethodnim fazama analize identificirani kao empirijski najstabilniji i najinformativniji determinanti cijena stanova:

- indeks razvijenosti (*ir*)
- indeks turističke razvijenosti (*itr*)
- prostorni pomak medijalne cijene stana (*lag_target*)
- udaljenost od mora, uključena u model kao logaritamska transformacija $\log(\text{dist_sea})$
- udio broja nenastanjenih stanova (*vacapart*)

Uz fiksne efekte, model uključuje i **slučajni odsječak (random intercept)** na razini varijable za grupa kojoj jedinica lokalne samouprave pripada prema indeksu razvijenosti (*gir*). Vrijednost ove varijable se kreće od 1 do 8 (1 = najnerazvijenija skupina lokalnih jedinica; 8 = najrazvijenija grupa lokalnih jedinica) i ona predstavlja šire skupine jedinica lokalne samouprave. Uvođenjem slučajnog odsječka dopušta se da se osnovna razina cijena razlikuje među razvojnim skupinama, neovisno o uključenim objašnjavajućim varijablama. Time se modelom eksplicitno uvažavaju neopažene regionalne i institucionalne karakteristike koje mogu utjecati na razinu prosječnih cijena stanova.

Ovakav pristup osobito je važan s obzirom na činjenicu da se dobivene procjene koriste prvenstveno za potrebe provedbe javnih politika, u kojima stupanj razvijenosti lokalne zajednice ima ključnu ulogu u dizajnu, ciljanju i implementaciji programa. Uključivanjem slučajnog odsječka model dodatno veže procijenjene vrijednosti medijalnih cijena uz širi kontekst razvijenosti i regionalne pripadnosti JLS-a, čime se povećava institucionalna relevantnost i robusnost rezultata.

Formalno, procijenjeni model prosječnih cijena stanova bez PDV-a može se zapisati kao:

$$\log \log (\text{aprice}_i) = \beta_0 + \beta_1 \text{ir}_i + \beta_2 \text{itr}_i + \beta_3 \text{lag_target}_i + \beta_4 \log \log (\text{dist_sea}_i + 1) + \beta_5 \log (\text{vacapart}_i + 1) u_{g(i)} + \varepsilon_i$$

gdje je:

- mprice_i – medijalna cijena stana u JLS-u *i*
- β_0 – zajednički odsječak (fixed intercept)
- $\beta_1, \dots, \beta_5$ – koeficijenti fiksnih efekata
- $u_{g(i)} \sim N(0, \sigma_u^2)$ – slučajni odsječak za razvojnu skupinu *gir* kojoj JLS pripada
- $\varepsilon_i \sim N(0, \sigma^2)$ – slučajna pogreška

Model je procijenjen metodom REML (Restricted Maximum Likelihood), što je standardni i preporučeni pristup za procjenu varijanci slučajnih efekata u mješovitim modelima (Bates et al, 2015).

Rezultati procjene prikazani u Tablici 18 pokazuju da su indeks razvijenosti (ir) i udaljenost od mora statistički značajni determinanti prosječnih cijena stanova bez PDV-a. Procijenjeni koeficijent indeksa razvijenosti je pozitivan i visoko statistički značajan, što upućuje na snažnu povezanost razine lokalnog socioekonomskog razvoja i razine tržišnih cijena stanova. Udaljenost od mora ima očekivani negativan i statistički značajan učinak, potvrđujući postojanje snažnog prostornog gradijenta cijena povezanog s obalnom atraktivnošću.

Prostorni pomak prosječne cijene stana (lag_target) ima pozitivan učinak i graničnu statističku značajnost, što upućuje na prisutnost regionalnih cjenovnih obrazaca i prostorne međuovisnosti. Indeks turističke razvijenosti (itr) i udio praznih stanova (vacapart) imaju pozitivan, ali statistički slabiji samostalni učinak kada se u modelu istodobno kontroliraju ostali ekonomski i prostorni čimbenici. Njihova uloga u modelu stoga se primarno očituje kroz zajedničko djelovanje s ostalim varijablama, a ne kroz izolirani marginalni učinak.

TABLICA 18. Procjena fiksnih efekata u linearnom mješovitom modelu

Objašnjavajuća varijabla	Procjena koeficijenta	Standardna pogreška	t-vrijednost
Slučajni odsječak	4,143	0,580	7,142
Indeks razvijenosti (ir)	0,030	0,006	4,949
Indeks turističke razvijenosti (itr)	0,005	0,005	1,054
Prostorni pomak cijena (lag_target)	0,00011	0,00006	1,931
Udaljenost od mora (dist_sea)	-0,0245	0,0090	-2,709
Broj nenastanjenih stanova (vacapart)	0,0098	0,0153	0,641

Izvor: izračun autorice.

Rezultati procjene potvrđuju i da ovakva mješovita specifikacija linearnog regresijskog modela ima empirijsko opravdanje. Procijenjena varijanca slučajnog odsječka po razvojnim skupinama (gir), prikazana u Tablici 19, iznosi 0,00159, dok rezidualna varijanca iznosi 0,02200. Na temelju tih vrijednosti izračunat je intraklasni koeficijent korelacije (ICC) od približno 6,8 %, što znači da se oko sedam posto ukupne varijance logaritmirane prosječne cijene stanova može pripisati razlikama među razvojnim skupinama jedinica lokalne samouprave. Ovaj rezultat je vrlo sličan rezultatu mješovitog modela za medijalne cijene stanova za kojeg intraklasni koeficijent korelacije iznosi 7,5 %.

Iako se najveći dio varijabilnosti cijena objašnjava razlikama među pojedinim JLS-ovima i njihovim ekonomskim i prostornim obilježjima, ovaj rezultat jasno potvrđuje postojanje sistematske institucionalne i regionalne heterogenosti koja nije u potpunosti obuhvaćena fiksnim dijelom modela. Uvođenje slučajnog odsječka omogućuje modelu da "upije" dio

neopaženih regionalnih čimbenika, poput povijesnih razvojnih putanja, institucionalnih razlika ili šireg prostornog konteksta.

TABLICA 19. Varijanca slučajnih i rezidualnih komponenti

Komponenta varijance	Varijanca	Standardna devijacija
Slučajni odsječak (<i>gir</i>)	0,00159	0,0399
Rezidualna varijanca	0,02200	0,1483

Izvor: izračun autorice.

Uključivanje slučajnog presretanja omogućuje modelu da “upije” dio neopaženih regionalnih čimbenika – poput povijesnih razvojnih putanja, institucionalnih razlika ili šireg prostornog konteksta, čime se poboljšava stabilnost koeficijenata fiksnih efekata i povećava robusnost predikcija. To je osobito važno u kontekstu primjene rezultata u javnim politikama, gdje je stupanj razvijenosti lokalne zajednice ključna dimenzija u dizajnu i provedbi različitih programa, a procjene cijena stanova moraju biti konzistentne i usporedive među teritorijalnim cjelinama. Dodatno, procijenjeni fiksni efekti u mješovitom modelu zadržavaju očekivane znakove i statističku značajnost, što potvrđuje da uvođenje slučajnog presretanja ne narušava interpretabilnost osnovnih ekonomskih i prostornih determinanti cijena stanova, već ih nadopunjuje. Distribucija skaliranih reziduala pokazuje približno simetričan oblik oko nule, bez izraženih sistematskih odstupanja, što dodatno upućuje na adekvatnu specifikaciju modela.

Kako bi bili apsolutno sigurni da je linearni mješoviti model unaprijeđenja u odnosu na model višestruke linearne regresije koji je u prethodnim fazama analize odabran kao najbolji model za predikciju s četiri ključne objašnjavajuće varijable, u Tablici 20. prikazana je usporedba prediktivnih rezultata višestrukog linearnog regresijskog modela i linearnog mješovitog modela prosječnih cijena stanova s pet odabranih objašnjavajućih varijabli. Ta usporedba pokazuje da uvođenje hijerarhijske strukture putem slučajnog odsječka formiranog na osnovi pripadnosti jedinice lokalne samouprave razvojnoj skupini (*gir*) dovodi do sustavnog poboljšanja prediktivne uspješnosti modela.

Rezultati pokazuju da linearni mješoviti model ostvaruje osjetno nižu prosječnu apsolutnu pogrešku (MAE) u odnosu na standardnu linearnu regresiju (251 €/m² naspram 283 €/m²), što znači da su predikcije u prosjeku znatno bliže stvarnim vrijednostima prosječnih cijena stanova. Istodobno, RMSE se smanjuje s 438 €/m² na 405 €/m², što upućuje na bolju kontrolu velikih pogrešaka i stabilnije predikcije u cijelom rasponu cijena. Koeficijent determinacije dodatno potvrđuje prednost mješovitog modela, budući da se R² povećava s 0,735 na 0,773.

Iako se poboljšanje R^2 numerički može činiti umjerenim, ono je metodološki značajno jer se postiže bez povećanja broja objašnjavajućih varijabli, već isključivo boljim modeliranjem hijerarhijske strukture podataka.

TABLICA 20. Usporedba višestruke linearne regresije i linearnog mješovitog modela

Model	MAE	RMSE	R^2
Višestruka linearna regresija	282,74	438,43	0,735
Linearni mješoviti model	251,44	405,47	0,773

Napomena: MAE, RMSE i R^2 za linearni mješoviti model izračunate su na originalnim (nelogaritmiziranim) podacima za ciljane varijablu kako bi mjere bile usporedive s onima koje ostvaruje model višestruke linearne regresije.

Izvor: izračun autorice.

Ključna prednost linearnog mješovitog modela u odnosu na čistu linearnu regresiju proizlazi iz njegove sposobnosti da eksplicitno uvaži neopažene regionalne razlike u razinama cijena stanova. Uvođenjem slučajnog odsječka dopušta se da osnovna razina cijena varira među razvojnim skupinama, čime se smanjuje pristranost procjena i povećava stabilnost koeficijenata fiksnih efekata.

Takav pristup osobito je prikladan u kontekstu primjene rezultata u javnim politikama, gdje se procjene koriste za usporedbu jedinica lokalne samouprave različitog stupnja razvijenosti i institucionalnog okruženja. Linearni mješoviti model stoga predstavlja metodološki i empirijski utemeljeno unaprjeđenje u odnosu na višestruku linearnu regresiju te pruža pouzdan i policy-relevantan okvir za procjenu prosječnih cijena stanova bez PDV-a u svim jedinicama lokalne samouprave u Republici Hrvatskoj.

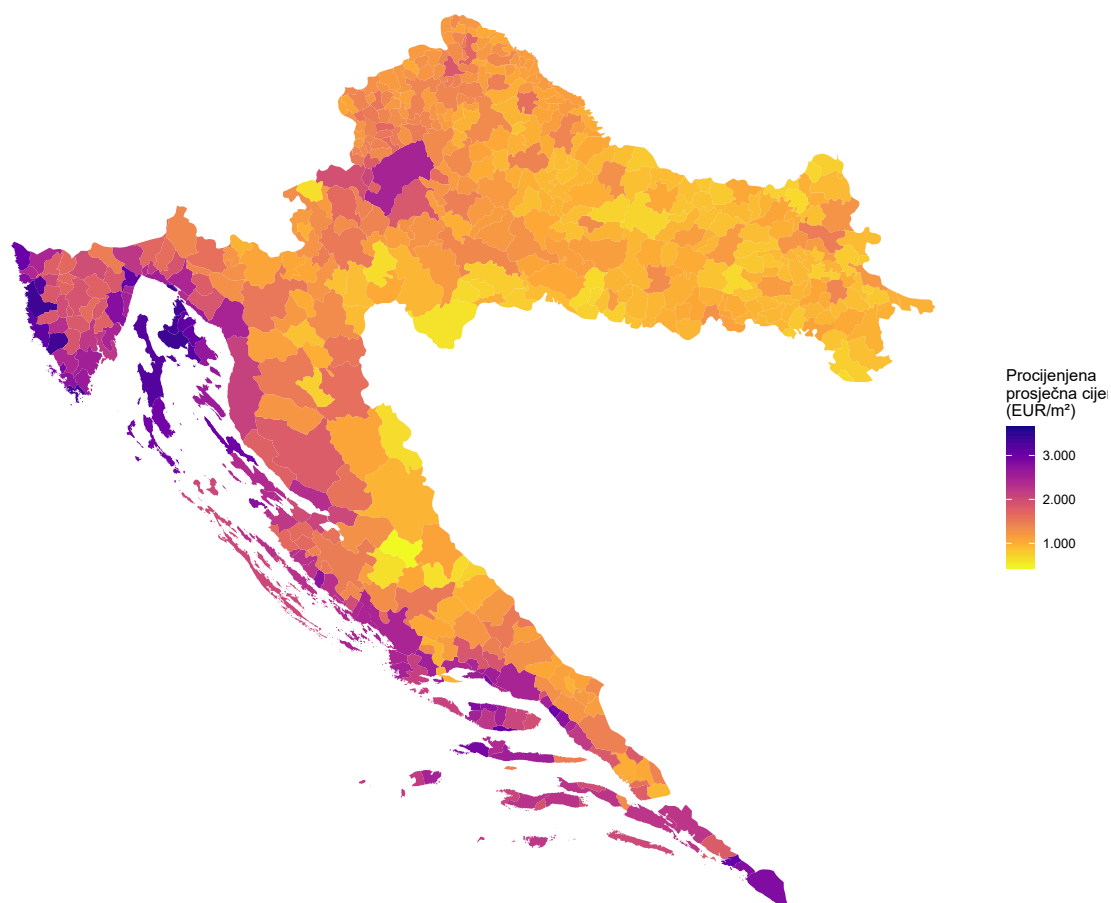
Na temelju procijenjenog modela generirane su predikcije prosječnih cijena stanova za sve JLS-ove u Hrvatskoj, pri čemu su predikcije transformirane natrag iz logaritamskog u izvornu cjenovnu skalu. Dobiveni rezultati kartografski su prikazani na Slici 5, koja ilustrira prostornu distribuciju procijenjenih prosječnih cijena stanova i jasno odražava ulogu ekonomskih, prostornih i tržišnih čimbenika u njihovu formiranju.

Navedeni rezultati su zatim nadopunjeni na način koji je identičan načinu na kojeg su nadopunjeni podaci o medijalnim cijenama i najamninama. Drugim riječima, ako je jedinica lokalne samouprave ostvarila 20 ili više kupoprodajnih transakcija u 2025. godini, umjesto procjene, korištena je stvarna (odnosno ostvarena) prosječna cijena stana bez PDV-a. Ukoliko takav podatak nije bio dostupan, korištena je vrijednost dobivena procjenom mješovitog modela opisanog u ovom poglavlju. *Procjenu prosječne vrijednosti stanova zasnovanu na tom modelu možemo pri tome protumačiti kao očekivana prosječnu vrijednost stana koju bi jedinica lokalne samouprave trebala ostvariti obzirom na svoj stupanj ekonomske i turističke razvijenosti, udaljenosti od mora, prosječne cijene*

stanova u susjednim jedinicama te s obzirom na razvojnu skupinu kojoj jedinica lokalne samouprave pripada.

Nakon što smo dobili završne prosječne vrijednosti stanova bez PDV-a za sve lokalne jedinice, iste su pretvorene u indeks u kojem bazu predstavlja prosječnu cijenu stanova u Zagrebu. Ovakav indeksni pristup otvara mogućnost da se u sljedeće dvije ili više godina vrijednosti prosječnih cijena stanova ne procjenjuju iznova primjenom metodološki zahtjevnog i vremenski intenzivnog postupka strojnog učenja, već da se, polazeći od ažurirane vrijednosti prosječne cijene stana bez PDV-a za Grad Zagreb, procijene prosječne cijene stanova i za sve ostale jedinice lokalne samouprave. Drugim riječima, promjene u razini cijena stanova u Zagrebu mogle bi se koristiti kao referentna osnova za proporcionalno prilagođavanje procijenjenih prosječnih cijena stanova u ostatku zemlje, uz zadržavanje prethodno utvrđenih relativnih odnosa među lokalnim tržištima.

SLIKA 5. Procijenjene prosječne cijene stanova bez PDV-a



Izvor: izračun autorice.

Prednost ovakvog pristupa leži u njegovoj operativnoj jednostavnosti, brzini provedbe i stabilnosti rezultata, što je osobito važno u kontekstu redovite primjene procjena u javnim politikama i administrativnim postupcima. Time se smanjuje potreba za čestim ponovnim modeliranjem, a istodobno se osigurava konzistentnost procjena kroz vrijeme.

Istodobno, valja jasno naglasiti i ograničenja ovakvog indeksnog pristupa. Polazišna pretpostavka da se relativni odnosi prosječnih cijena stanova između Zagreba i ostalih jedinica lokalne samouprave kratkoročno ne mijenjaju nužno ne mora u potpunosti vrijediti u svim okolnostima. Razlike u dinamici lokalnih tržišta, osobito u turistički izloženim primorskim područjima ili u jedinicama s naglim promjenama potražnje, mogu dovesti do toga da se prosječne cijene stanova u pojedinim sredinama kreću brže ili sporije od zagrebačkog tržišta. U takvim slučajevima indeksna metoda može privremeno podcijeniti ili precijeniti stvarnu vrijednost cijena stanova. Stoga se predloženi indeksni pristup može smatrati primjerenim rješenjem za kratkoročno razdoblje, primjerice u sljedeće dvije godine, uz pretpostavku relativne stabilnosti prostorne strukture tržišta stanova. U srednjem roku preporučljivo je ponovno provesti cjelovitu procjenu prosječnih cijena stanova primjenom metoda opisanih u ovoj studiji, kako bi se uvažile moguće strukturne promjene u odnosima među lokalnim tržištima stanova za prodaju i osigurala dugoročna točnost i relevantnost procjena.

8. Određivanje minimalnog praga broja transakcija za robustnu procjenu lokalnih tržišnih vrijednosti stambenih nekretnina

Osnovni preduvjet za izradu robustnog modela procjene tržišnih vrijednosti stambenih nekretnina i najamnina jest kvaliteta i reprezentativnost ulaznih podataka. S obzirom na to da se predložena metodologija oslanja na transakcijske podatke iz informacijskog sustava tržišta nekretnina (eNekretnine), nužno je utvrditi prag niske frekvencije transakcija ispod kojeg izračunati tržišni pokazatelji gube svoju analitičku vrijednost. Dosadašnja praksa i statistički standardi nalažu da je za izračun bilo kakvog relevantnog tržišnog pokazatelja (poput medijalne ili prosječne cijene) na razini pojedine jedinice lokalne samouprave (JLS) potrebno postojanje minimalno deset individualnih transakcija u promatranom razdoblju.

Međutim, podaci iz Tablice 1. ukazuju na značajne strukturne izazove u prostornoj pokrivenosti tržišta nekretnina u Republici Hrvatskoj. Od ukupno 556 jedinica lokalne samouprave, zakonski minimum od deset transakcija za kupoprodaju stanova/apartmana zadovoljava svega 142 JLS, dok je situacija na tržištu najma još nepovoljnija – taj kriterij ispunjava tek 51 jedinica. Ovakva distribucija ostavlja velik dio teritorija RH bez ikakvih službenih tržišnih pokazatelja, što izravno ograničava mogućnost donošenja informiranih odluka u sferi stambenih politika na tim područjima.

Dodatni metodološki problem leži u činjenici da čak i u onim jedinicama koje zadovoljavaju formalni minimum od deset transakcija, procijenjene tržišne vrijednosti poput medijalne ili aritmetičke sredine često pokazuju visoku razinu statističke nestabilnosti. Na malim uzorcima, jedna ekstremna transakcija (npr. prodaja luksuzne nekretnine ili derutnog objekta) može nerazmjerno iskriviti sliku cijelog lokalnog tržišta, čineći izračunate vrijednosti nepouzdanima za procese strojnog učenja. Ovaj problem se može djelomično ublažiti korištenjem medijana za izračun prosječne vrijednosti, no čak i uz primjenu medijana umjesto aritmetičke sredine problem nažalost ne nestaje, što u konačnici znači da je broj lokalnih jedinica za koje imamo dostupne pouzdane informacije o tržišnim vrijednostima stambenih nekretnina još manji, a jaz u prostornom obuhvatu podataka još dublji.

Slijedom navedenog, ključni izazov ove studije jest pronalaženje optimalne ravnoteže između beskompromisne pouzdanosti izračunatih tržišnih vrijednosti i postizanja što šire prostorne pokrivenosti tržišnim podacima. Cilj je ovog poglavlja, primjenom bootstrap metode simulacije, empirijski utvrditi točku u kojoj broj transakcija osigurava stabilnost medijalne cijene, čime se definira metodološki utemeljen prag za uključivanje lokalnih jedinica u finalne modele procjene.

Izabrani metodološki pristup i korišteni podaci

U analizi tržišta nekretnina, odabir adekvatne mjere srednje vrijednosti presudan je za preciznost modela. U ovoj studiji kao primarni indikator odabran je **medijan**, a ne aritmetička sredina. Takav odabir temelji se na činjenici da su distribucije cijena nekretnina u pravilu asimetrične i sadrže ekstremne vrijednosti koje mogu značajno iskriviti aritmetičku sredinu. Medijan, kao pozicijska srednja vrijednost, znatno je robusniji na ekstremno visoke ili niske cijene pojedinačnih transakcija, čime pruža realniju sliku tipične tržišne vrijednosti na određenom lokalnom području (Geltner i sur., 2001). Medijalna cijena nekretnina je ujedno i vrijednost koja se objavljuje u Pregledu tržišta nekretnina, zajedničkoj publikaciji Ekonomskog instituta, Zagreb i Ministarstva prostornoga uređenja, graditeljstva i državne imovine Republike Hrvatske, pa ima najviše smisla da se u ovoj analizi testira stabilnost i preciznost upravo te mjere.

Kako bi se empirijski utvrdio minimalni broj transakcija potreban za stabilnu procjenu medijana, primijenjena je metoda **ponovnog uzorkovanja s povratom** (engl. *bootstrapping*). Metoda, koju je inicijalno razvio Efron (1979), omogućuje procjenu distribucije uzorkovanja statističkog pokazatelja bez pretpostavke o normalnosti distribucije izvornih podataka. S obzirom na to da bootstrap zahtijeva “bazni skup” podataka koji vjerno reprezentira lokalno tržište, simulacija nije provedena na svim JLS-ovima. Za analizu su odabrane isključivo one jedinice lokalne samouprave koje u promatranom razdoblju unutar baze eNekretnine sadrže više od 100 ostvarenih kupoprodajnih transakcija stanovima u prvih 11 mjeseci 2025. godine. Kao rezultat primjene ovog kriterija, u analizu je uključeno ukupno 33 lokalne jedinice koje predstavljaju 33 lokalnih tržišta – od velikih urbanih centara poput Zagreba, Splita i Rijeke, do specifičnih obalnih i kontinentalnih gradova. Ovako definiran skup služi kao “zlatni standard” (referentna populacija) iz kojeg se generiraju manji uzorci kako bi se promatrala brzina konvergencije prema stvarnom medijanu.

Sam postupak simulacije stabilnosti medijana proveden je kroz sljedeće precizno definirane korake:

1. **Iterativno uzorkovanje:** Za svaku odabranu JLS, algoritam je nasumično izvlačio poduzorke različitih veličina, u rasponu od 2 do $n=60$ transakcija. Za svaku pojedinu veličinu uzorka, postupak izvlačenja ponovljen je **1.000 puta** uz primjenu metode s povratom (engl. *sampling with replacement*).
2. **Izračun distribucije medijana:** U svakoj od 1.000 iteracija izračunat je medijan cijene. Rezultat je distribucija od 1.000 različitih medijana za svaku promatranu razinu broja transakcija.

3. **Kvantifikacija nestabilnosti:** Za svaku razinu broja transakcija od 2 do 60 izračunata je standardna devijacija, a zatim i koeficijent varijacije tih 1.000 dobivenih medijana, kako bi se omogućila usporedba između gradova s različitim razinama cijena.
4. **Identifikacija točke stabilizacije:** Analizom dobivenih krivulja stabilnosti (engl. stability plots) za svaku promatranu jedinicu lokalne samouprave, sustavno je praćen pad koeficijenta varijacije s povećanjem broja transakcija, pri čemu je poseban fokus stavljen na identifikaciju točaka u kojima se nagib krivulje stabilnosti ublažava, odnosno na tzv. „točku lakta“ (engl. elbow point). Sam proces identifikacije započinje analizom dinamike promjene nagiba krivulje kroz tri karakteristične zone pouzdanosti. U početnoj fazi, koja obuhvaća zonu visoke nestabilnosti s koeficijentom varijacije iznad 10 %, nagib krivulje je izrazito strm, što ukazuje na to da su izračunati medijani ekstremno osjetljivi na svaku pojedinačnu transakciju te da u podacima prevladava statistički šum koji takve procjene čini neprihvatljivima za potrebe bilo kakve analize. Kako se broj transakcija povećava, krivulja ulazi u zonu tranzicije između 5 % i 10 % gdje se njezin nagib počinje ublažavati, signalizirajući postupan prijelaz prema umjerenoj pouzdanosti, iako rizik od slučajne pogreške i dalje ostaje statistički značajan. Ključni trenutak u analizi predstavlja dosezanje razine broja transakcija nakon koje koeficijent varijacije pada ispod kritičnog praga od 5 %, a nagib krivulje postaje gotovo vodoravan. Taj prag, sukladno statističkim standardima (Brown et al., 2001), označava postizanje visoke razine pouzdanosti i robusnosti procjene. U ovoj zoni nastupa zakon opadajućih prinosa (engl. diminishing returns), gdje daljnje povećanje uzorka više ne rezultira značajnim dobitkom na preciznosti, dok se istovremeno potvrđuje dobivanje „čistog“ tržišnog signala neophodnog za izradu pouzdanih modela procjene vrijednosti nekretnina. Na taj način, točka lakta služi kao objektivni metodološki kriterij koji razdvaja nestabilne aproksimacije od znanstveno utemeljenih tržišnih pokazatelja.

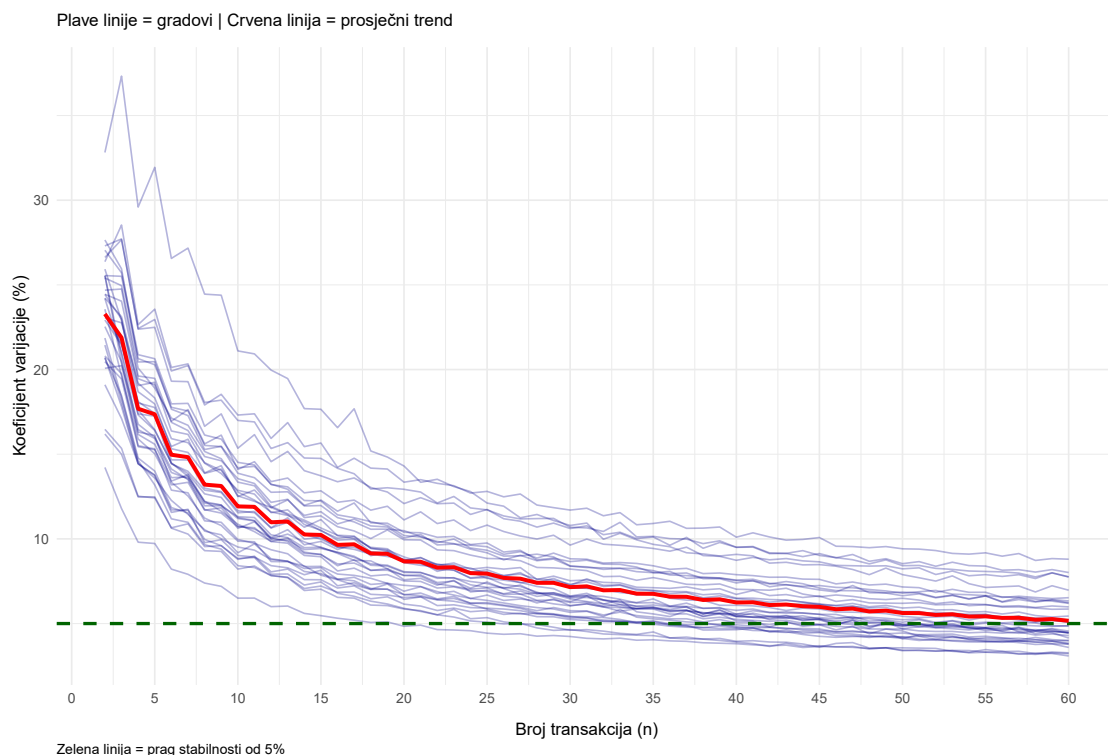
Rezultati bootstrap simulacija i utvrđivanje optimalnog praga broja transakcija

Rezultati provedene analize za sve gradove zajedno prikazani su na slici 6. u Tablici 21. Slika 6. predstavlja sintetički vizualni prikaz rezultata bootstrap simulacije provedene na svim jedinicama lokalne samouprave koje zadovoljavaju kriterij od preko 100 ostvarenih transakcija, strukturiran kao usporedni prikaz krivulja stabilnosti koji omogućuje uvid u dinamiku statističke pogreške na razini cijelog tržišta. Pojedinačne plave linije, iscrtane s određenom razinom prozirnosti, prikazuju kretanje koeficijenta varijacije za svih 33 lokalnih jedinica u analizi, a čime se zorno ilustrira heterogenost hrvatskog tržišta nekretnina i različite razine disperzije cijena u ovisnosti o likvidnosti lokalnih tržišta. Slike za pojedinačne gradove dostupne su u Prilogu Studiji. Ključni analitički element grafikona je crvena podebljana linija koja predstavlja prosječni trend stabilizacije svih promatranih jedinica, služeći kao univerzalna krivulja učenja koja neutralizira lokalne

specifičnosti i pruža empirijski utemeljenu preporuku za opći minimalni broj transakcija. Horizontalna isprekidana zelena linija, fiksirana na kritičnom pragu od 5 %, djeluje kao objektivna metodološka granica koji razdvaja zonu statističke nestabilnosti od područja robusne procjene.

Cjelokupni prikaz kretanja linija za gradove i zajedničke linije stabilizacije koje obuhvaćaju raspon od 2 do 60 transakcija, jasno demonstrira zakon opadajućih prinosa u statističkom uzorkovanju, gdje strmi pad koeficijenta varijacije u početnom dijelu sugerira da se **najveći dobiti u preciznosti ostvaruju ako je u izračun medijalne cijene stana uključeno barem 20 kupoprodajnih transakcija, a još nešto veća preciznost se ostvaruje ako je u izračun medijalne cijene stana uključeno minimalno 30 transakcija**. Uz pomoć Tablice 21. očito je i da crvena krivulja koja predstavlja prosječni trend stabilizacije kritični prag od 5% doseže tek na 62 transakcije uključene u izračun medijana, što znači da se u hrvatskim uvjetima moramo pomiriti sa nešto nepreciznijim izračunom medijana ako želimo imati veću zastupljenost lokalnih jedinica pokazateljima vezanima za stambeno tržište.

SLIKA 6. Usporedna stabilnost medijalne cijene stanova u izabranim gradovima



Izvor: izračun autorice.

TABLICA 21. Koeficijenti varijacije i pragovi pouzdanosti medijalne cijene stanova s obzirom na broj transakcija

Jedinica lokalne samouprave	Ukupan broj transakcija	Broj transakcija pri kojem je koeficijent varijacije medijalne cijene stana manji od 5 %	Stabilnost medijalne cijene stanova mjerena koeficijentom varijacije za N transakcija (%)						
			N=10	N=20	N=30	N=40	N=50	N=60	N=70
Bjelovar	127	26	8,4	5,9	4,6	3,9	3,4	3,2	3,0
Čakovec	171	50	11,3	8,1	6,5	5,4	5,0	4,4	4,2
Crikvenica	217	Preko 70	15,4	12,1	10,6	9,5	8,5	8,0	7,5
Dubrovnik	123	63	12,5	8,7	7,3	6,5	5,7	5,1	4,6
Dugo Selo	123	64	12,9	8,6	7,1	6,1	5,4	5,1	4,8
Grad Zagreb	6.778	36	9,5	6,7	5,4	4,8	4,2	3,8	3,5
Karlovac	247	Preko 70	11,7	8,7	7,3	6,5	5,9	5,4	5,2
Kaštela	112	Preko 70	17,3	13,4	11,7	10,1	9,4	8,8	8,1
Koprivnica	120	20	6,5	4,8	4,2	3,8	3,4	3,3	3,1
Makarska	264	Preko 70	14,2	10,4	8,7	7,5	6,9	6,3	5,8
Malinska	104	Preko 70	12,5	10,1	8,8	7,9	7,2	6,5	5,9
Medulin	104	Preko 70	14,1	11,1	9,6	8,7	8,2	7,8	7,1
Novalja	165	66	11,5	8,8	7,1	6,3	5,7	5,2	4,9
Osijek	639	Preko 70	13,9	10,5	8,4	7,6	6,8	6,2	5,9
Poreč	188	50	12,8	8,9	6,6	5,8	5,0	4,4	4,3
Pula	364	46	10,9	7,9	6,3	5,5	4,9	4,5	4,2
Rijeka	1.035	46	10,8	7,6	6,5	5,6	4,9	4,5	4,1
Rovinj	122	Preko 70	13,6	10,0	8,2	7,0	6,5	6,0	5,6
Samobor	108	Preko 70	11,2	8,7	7,5	6,7	6,2	5,9	5,5
Šibenik	154	50	11,3	7,8	6,6	5,4	4,8	4,6	4,3
Sisak	111	Preko 70	17,0	12,9	10,7	9,5	8,4	7,8	7,2
Slavonski Brod	131	36	8,2	6,2	5,2	4,6	4,2	3,8	3,7
Solin	116	54	9,2	6,9	6,1	5,4	5,1	4,9	4,6
Split	704	27	8,9	5,9	4,6	4,0	3,6	3,1	3,0
Umag	151	59	10,6	8,0	6,6	5,6	5,2	4,9	4,5
Varaždin	147	36	8,8	6,5	5,4	4,8	4,4	4,0	3,8
Velika Gorica	391	39	10,0	7,3	5,7	4,9	4,1	3,8	3,7
Vinkovci	133	35	9,9	6,4	5,3	4,7	4,1	3,6	3,5
Vir	168	Preko 70	14,5	10,2	8,4	7,4	6,6	6,2	5,8
Viškovo	120	42	13,4	8,6	6,3	5,2	4,6	4,0	3,5
Vukovar	131	Preko 70	21,1	14,3	10,8	9,1	7,9	7,0	6,5
Zadar	642	45	10,7	7,5	6,1	5,3	4,7	4,2	4,0
Zaprešić	354	42	9,0	6,7	5,8	5,1	4,7	4,3	4,0
Prosjek	-	62	11,9	8,7	7,2	6,2	5,6	5,2	4,8

Izvor: Izračun autorice.

Analiza rezultata prikazanih u Tablici 21. otkriva duboku strukturnu heterogenost hrvatskog tržišta nekretnina i ukazuje na značajne izazove u postizanju statističke pouzdanosti procjena na lokalnoj razini. Podaci jasno pokazuju da zakonski definirani minimum od deset transakcija nije dostatan za postizanje visoke preciznosti izračuna prosječnih tržišnih vrijednosti, što potvrđuju vrijednosti koeficijenta varijacije (KV) za uzorak od deset transakcija koje se u gotovo svim promatranim jedinicama lokalne samouprave nalaze iznad kritičnog praga od 10 %. Primjerice, u gradovima poput Kaštela ili Splita, KV pri deset transakcija iznosi visokih 17,3 %, odnosno 15,2 %, što sugerira da su medijalne cijene izračunate na tako malim uzorcima izrazito nestabilne i podložne snažnom utjecaju ekstremnih vrijednosti. Čak i u Gradu Zagrebu, koji raspolaže najvećim brojem transakcija u bazi, koeficijent varijacije pri uzorku od deset transakcija iznosi 9,5 %, što je na samoj granici prihvatljivosti i potvrđuje nužnost uvođenja strožih kriterija za modeliranje. Prosječni koeficijent varijacije za čitav lokalnih jedinica je čak 11,9 %, što izračun medijalne cijene stana na temelju svega 10 transakcija čini neprihvatljivo nepouzdanima i nepodobnima za bilo kakvu analizu, uključujući i analizu provedenu u ovoj Studiji.

Posebno su indikativni rezultati koji se odnose na minimalni prag transakcija potreban za postizanje koeficijenta varijacije manjeg od 5 %. Iz tablice je razvidno da velik broj gradova, unatoč tome što imaju bazu od preko 100 transakcija, ne uspijeva dosegnuti ovaj prag stabilnosti čak ni s većim uzorkom. Jedinice poput Dubrovnika, Osijeka, Rijeke, Zadra i Pule označene su kategorijom “Preko 60”, što implicira izrazitu internu varijabilnost cijena unutar tih tržišta. Na takvim područjima prodaju se nekretnine vrlo različitih karakteristika i cjenovnih razreda, što onemogućuje medijanu da se “smiri” na manjem broju dostupnih transakcija. S druge strane, gradovi poput Bjelovara, Koprivnice ili Varaždina pokazuju znatno bržu konvergenciju prema stabilnim vrijednostima medijalne cijene stanova, dostižući prag od 5 % varijacije pri uzorcima od 20 do 26 transakcija, što ukazuje na homogenija lokalna tržišta s ujednačenijim cijenama stanova.

Daljnjom analizom kretanja preciznosti u odnosu na volumen podataka, uočava se značajna, ali neujednačena stabilizacija procjena pri povećanju uzorka na 20, 30 i 40 transakcija. Na razini od 20 transakcija, prosječni koeficijent varijacije za sve promatrane jedinice spušta se na 8,4 %, što predstavlja osjetan napredak u odnosu na početnih 11,9 %, ali i dalje ostaje iznad praga od 5 % koji se smatra standardom visoke pouzdanosti. Iako kod homogenijih tržišta, poput onih u Koprivnici ili Varaždinu, ovaj broj transakcija već osigurava zadovoljavajuću preciznost, za većinu obalnih i velikih urbanih središta razina od 20 transakcija i dalje podrazumijeva rizik od značajnijih statističkih odstupanja. Dodatno povećanje uzorka na 30 transakcija rezultira padom prosječnog koeficijenta varijacije na 6,9 %. U ovoj točki, nagib krivulje stabilnosti kod većine jedinica lokalne samouprave počinje se vidljivo ublažavati. Pri uzorku od 30 transakcija, veći broj jedinica primiče se pragu od 5 %, a ekstremni utjecaj pojedinačnih netipičnih kupoprodaja postaje

znatno amortiziran. Ipak, prosječna vrijednost od 6,9 % jasno ukazuje na to da čak ni trostruko veći broj transakcija od zakonskog minimuma nije dovoljan da bi se na razini čitave države osigurala beskompromisno stabilna procjena medijana. Kao što je već navedeno u uvodnom poglavlju, zakonski minimum se primjenjuje kako bi se zaštitilo pravo na privatnost podataka individualnih sudionika na stambenom tržištu, bilo da je riječ o kupcima, prodavateljima, najmoprimcima ili najmodavcima.

Kod uzorka od 40 transakcija, prosječni koeficijent varijacije dodatno konvergira prema vrijednosti od 6,0 %, ali po znatno usporenijoj stopi. Drugim riječima, iako je riječ o razini koja je blizu ciljanog praga, važno je primijetiti da se dinamika poboljšanja preciznosti usporava, što potvrđuje djelovanje zakona opadajućih prinosa u statističkom uzorkovanju. Za gradove poput Zagreba, uzorak od 40 transakcija (gdje koeficijent varijacije iznosi 4,8 %) već osigurava ulazak u zonu visoke robusnosti, no na razini općeg prosjeka, varijabilnost od 6,0 % sugerira da su hrvatska lokalna tržišta nekretnina toliko heterogena da čak i relativno veliki uzorci ostavljaju prostor za određenu nesigurnost.

I konačno, analiza pokazuje da prosječni koeficijent varijacije pada ispod strogo zadanog praga od 5 % tek kada broj transakcija dosegne kritičnu masu od 62 jedinice. Međutim, primjena ovog "idealnog" praga u praksi bi značila drastičan gubitak podataka i geografske pokrivenosti. Naime, prag od 62 transakcije zadovoljava svega 47 lokalnih jedinica, što predstavlja tek 15,51 % od ukupnog broja JLS-ova u Hrvatskoj. Usporedbe radi, spuštanje praga na 30 transakcija povećava broj obuhvaćenih jedinica na 74 (24,4 %), što ujedno predstavlja i uzorak koji je korišten za strojno učenje modela medijalnih i prosječnih cijena stanova implementirano u ovoj Studiji. Istovremeno, metodološki kompromis od 20 transakcija omogućuje analizu 98 lokalnih jedinica, čime se obuhvaća 32,3 % tržišta. Taj prag je pak u ovoj Studiji korišten prilikom procjene medijalnih najamnina.

Ovi nalazi upućuju na to da bi prag od najmanje 20, a poželjno 30 kupoprodajnih transakcija trebalo smatrati razumnim metodološkim minimumom za relativno pouzdan izračun prosječnih tržišnih pokazatelja stambenih nekretnina na razini lokalnih jedinica.

Tek u tom rasponu dolazi do osjetnog smanjenja statističkog šuma, čime se omogućuje smislenija i interpretabilnija analiza. Naime, pri razini od približno 20 do 30 transakcija koeficijent varijacije ulazi u prijelaznu zonu između 5 % i 10 %, pri čemu se njegov nagib počinje ublažavati, što signalizira postupni prelazak prema umjerenoj razini pouzdanosti izračuna medijalne cijene stanova. Iako rizik od slučajne pogreške u toj fazi i dalje ostaje prisutan, on je znatno manji nego pri nižim razinama tržišne aktivnosti.

Istodobno, analiza pokazuje da se statistički šum u potpunosti reducira tek pri razinama od više od 60 transakcija, što u praksi znači da bi inzistiranje na tako visokom pragu znatno ograničilo obuhvat analize. Naime, primjena minimalnog praga od 62 transakcije rezultirala bi izostankom bilo kakvih podataka za gotovo 85 % svih jedinica lokalne

samouprave. Stoga je u analizi lokalnih tržišta nekretnina nužno prihvatiti određenu razinu nepouzdanosti pokazatelja, ali ju istodobno nastojati svesti na najmanju moguću mjeru.

U tom smislu, primjena minimalnog praga od 20 do 30 kupoprodajnih transakcija predstavlja optimalan kompromis između statističke preciznosti i prostorne pokrivenosti podataka, osiguravajući dovoljnu informativnost uz prihvatljiv statistički rizik. Upravo je takav pristup primijenjen u ovoj Studiji, čime se postiže uravnotežen odnos između pouzdanosti procjena i širine obuhvata lokalnih jedinica.

9. Zaključak

Ova studija pruža sustavnu i empirijski utemeljenu analizu lokalnih tržišta stambenih nekretnina u Republici Hrvatskoj s ciljem izrade pouzdanih, prostorno konzistentnih i metodološki usporedivih procjena cijena stanova i najamnina na razini jedinica lokalne samouprave. Polazeći od izraženih ograničenja dostupnosti transakcijskih podataka i velike heterogenosti lokalnih tržišta, razvijen je analitički okvir koji kombinira simulacijsku analizu pouzdanosti prosječnih pokazatelja tržišnih vrijednosti stambenih nekretnina u jedinicama lokalne samouprave, višemodelski pristup strojnog učenja te hijerarhijsko regresijsko modeliranje.

Analiza pouzdanosti prosječnih pokazatelja cijena stanova u ovisnosti o broju ostvarenih kupoprodajnih transakcija jasno je pokazala da pouzdanost izravno izračunatih medijalnih cijena stanova snažno ovisi o broju ostvarenih kupoprodajnih transakcija. Rezultati upućuju na to da se statistički šum značajno smanjuje tek pri razinama od najmanje 20 do 30 transakcija, dok se potpuna stabilizacija postiže tek pri znatno višim razinama tržišne aktivnosti. Budući da bi inzistiranje na visokim pragovima transakcija rezultiralo isključenjem velike većine lokalnih jedinica iz analize, u studiji je prihvaćen svjestan metodološki kompromis koji omogućuje istodobno razumnu statističku preciznost i dostatnu prostornu pokrivenost podataka. Time se osigurava da rezultati ostanu informativni, usporedivi i primjenjivi u kontekstu oblikovanja i provedbe javnih politika.

Usporedno s analizom cijena stanova, u studiji su izračunate i medijalne najamnine stanova, koje predstavljaju ključan pokazatelj za razumijevanje stanja i dinamike tržišta dugoročnog najma. Medijalne najamnine izravno su izračunate za one jedinice lokalne samouprave koje u promatranoj godini bilježe dovoljan broj registriranih ugovora o najmu, dok su za jedinice s nedostatnom tržišnom aktivnošću procijenjene primjenom istog prostorno i ekonomski informiranog metodološkog okvira strojnog učenja kao i kod cijena stanova. Pri tome su u izračun uključeni isključivo stanovi, a ne obiteljske kuće, kako bi se osigurala konzistentnost i usporedivost pokazatelja, s obzirom na ograničenja dostupnih podataka o korisnoj površini obiteljskih kuća. Takav pristup omogućuje dobivanje metodološki ujednačenih procjena najamnina za sve lokalne jedinice, uz jasno razlikovanje između izravno opaženih i modeliranih vrijednosti.

U središnjem dijelu analize primijenjen je višemodelski pristup strojnog učenja s ciljem identifikacije ključnih determinanti cijena stanova, a ne odabira jednog optimalnog algoritma. Rezultati dobiveni za medijalne cijene stanova s PDV-om i prosječne cijene stanova bez PDV-a pokazuju visok stupanj konzistentnosti, kako u pogledu dostupnosti podataka po lokalnim jedinicama, tako i u pogledu prostorne varijabilnosti cijena.

U oba slučaja kao ključne varijable dosljedno se izdvajaju indeks razvijenosti, prostorni pomak cijena stanova, udaljenost od mora te indeks turističke razvijenosti. Ova podudarnost rezultata potvrđuje stabilnost identificiranih ekonomskih i prostornih mehanizama formiranja cijena, neovisno o točnoj definiciji ciljne varijable.

Usporedba performansi različitih algoritama pokazala je da linearni i penalizirani regresijski modeli sustavno nadmašuju nelinearne metode u uvjetima ograničenog uzorka i visoke heterogenosti, uz istodobno veću interpretabilnost i stabilnost procjena. Na temelju kriterija RMSE-a kao ključne mjere prediktivne uspješnosti, višestruka linearna regresija s reduciranim skupom ključnih varijabli identificirana je kao najprikladniji okvir u fazi strojnog učenja, dok je završna faza analize dodatno unaprijeđena primjenom linearnog mješovitog modela sa slučajnim odsječkom na razini razvojnih skupina jedinica lokalne samouprave.

Rezultati mješovitog modela potvrđuju da uvođenje hijerarhijske strukture dovodi do osjetnog poboljšanja prediktivne uspješnosti u odnosu na čistu linearnu regresiju, uz smanjenje pogrešaka predikcije i povećanje objašnjene varijance. Procijenjeni intraklasni koeficijent korelacije pokazuje da dio varijabilnosti cijena proizlazi iz sistematskih razlika među razvojnim skupinama, što jasno upućuje na postojanje regionalne heterogenosti koju je nužno eksplicitno uvažiti u analizi tržišta nekretna.

U konačnici, studija pokazuje da je u uvjetima strukturnog nedostatka podataka nužno prihvatiti određenu razinu statističke nesigurnosti, ali i da se ona može sustavno kontrolirati, kvantificirati i svesti na razumnu mjeru primjenom prostorno informiranih i hijerarhijskih metoda. Dobivene procjene cijena stanova i najamnina za sve jedinice lokalne samouprave pružaju konzistentnu i transparentnu analitičku osnovu za usporedbu lokalnih tržišta, evaluaciju priuštivosti stanovanja i provedbu ciljanih mjera stambene politike na nacionalnoj i lokalnoj razini.

Unatoč primijenjenom metodološkom okviru i korištenju naprednih analitičkih pristupa, rezultate ove studije nužno je interpretirati uz jasno razumijevanje njezinih ograničenja. Temeljno ograničenje proizlazi iz izrazito ograničene dostupnosti izravnih tržišnih podataka na razini jedinica lokalne samouprave. U promatranoj godini izravni podaci o prosječnim cijenama stanova bez PDV-a raspoloživi su za 142 jedinice lokalne samouprave koje bilježe najmanje deset kupoprodajnih transakcija, što i dalje predstavlja tek ograničen dio ukupnog skupa JLS-ova u Republici Hrvatskoj. Situacija je još izraženija na tržištu dugoročnog najma: samo 51 jedinica lokalne samouprave bilježi deset ili više evidentiranih ugovora o najmu stanova, što čini svega 9,2 % svih JLS-ova, dok za preostalih više od 90 % lokalnih jedinica ne postoje administrativno vidljivi podaci koji bi omogućili izravni izračun medijalnih najamnina. Ovakva izrazito ograničena prostorna pokrivenost i mala veličina uzoraka temeljno su strukturno ograničenje analize te ujedno glavni razlog za primjenu modeliranih procjena cijena i najamnina u ovoj Studiji.

Dodatno ograničenje proizlazi iz statističkog šuma prisutnog u izravno izračunatim prosječnim i medijalnim cijenama stanova, osobito u lokalnim jedinicama s manjim brojem transakcija. Iako je u studiji primijenjen minimalni prag broja transakcija kako bi se smanjila nestabilnost pokazatelja, preostala varijabilnost može negativno utjecati na rezultate strojnog učenja, kako u fazi identifikacije ključnih objašnjavajućih varijabli, tako i u fazi procjene cijena i najamnina za jedinice bez izravnih opažanja. Taj se problem dodatno prenosi i na procjene najamnina, gdje je broj opažanja još manji, a heterogenost ugovornih uvjeta izraženija.

Važno ograničenje odnosi se i na prirodu pojedinih ključnih objašnjavajućih varijabli. Indeks razvijenosti i indeks turističke razvijenosti, koji se u svim modelima dosljedno pojavljuju kao najvažniji prediktori cijena, nisu izravno opazive varijable, već su rezultat složenih procjena koje su izradili eksperti na temelju više pokazatelja. Iako su ti indeksi široko prihvaćeni i institucionalno korišteni, oni nužno sadrže određenu razinu mjernog i procjenjivačkog rizika, koji se ne može u potpunosti kvantificirati unutar primijenjenog modelskog okvira.

Analiza je dodatno ograničena prostornom razinom na kojoj je bilo moguće primijeniti strojno učenje. Za katastarske općine Grada Zagreba strojno učenje nije bilo moguće provesti, ne zbog nedostatka tržišnih transakcija, već zbog nepostojanja objašnjavajućih varijabli na istoj prostornoj razini. Naime, iako su cijene stanova i najamnine u Gradu Zagrebu raspoložive na razini katastarskih općina, ključne objašnjavajuće varijable korištene u modelima strojnog učenja – poput indeksa razvijenosti, indeksa turističke razvijenosti, demografskih pokazatelja i razvojnih skupina – definirane su isključivo na razini jedinica lokalne samouprave i ne postoje u konzistentnom, službenom obliku za niže administrativne jedinice. U takvim okolnostima primjena strojnog učenja na razini katastarskih općina zahtijevala bi proizvoljnu disagregaciju ili imputaciju ključnih objašnjavajućih varijabli, što bi značajno narušilo metodološku utemeljenost i interpretabilnost rezultata. Zbog toga je analiza za katastarske općine Grada Zagreba provedena primjenom hijerarhijski strukturiranih i deskriptivno-modeliranih postupaka, ali ne i punog strojnog učenja.

Iako bi s analitičkog stajališta bilo poželjno provesti sličnu prostornu razdiobu i za druge velike gradove, poput Splita ili Rijeke, takva razina administrativne i statističke prostorne podjele u tim gradovima trenutačno ne postoji, što onemogućuje konzistentnu i usporedivu primjenu istog metodološkog pristupa koji je primijenjen u ovoj Studiji.

Među dodatnim ograničenjima valja istaknuti i činjenicu da analiza ne obuhvaća kvalitativne karakteristike stanova, poput starosti zgrade, energetske učinkovitosti, opremljenosti ili mikrolokacijskih obilježja, budući da takvi podaci nisu sustavno dostupni na nacionalnoj razini. Također, korištenje administrativnih granica jedinica

lokalne samouprave kao osnovne prostorne jedinice analize ne mora u potpunosti odgovarati stvarnim granicama lokalnih tržišta nekretnina, unatoč pokušaju da se taj problem ublaži uvođenjem prostornog pomaka ciljne varijable.

Naposljetku, rezultati ove studije predstavljaju presjek tržišnih odnosa u jednoj promatranoj godini te ne zahvaćaju dinamičke promjene u vremenu, što je osobito važno u kontekstu izrazite volatilnosti tržišta nekretnina u recentnom razdoblju. Unatoč tim ograničenjima, primijenjeni metodološki okvir omogućuje transparentnu, konzistentnu i empirijski utemeljenu procjenu cijena stanova i najamnina, uz jasno naznačene granice interpretacije rezultata. Pritom je posebna pažnja posvećena tome da se izbjegne bilo kakva diskrecijska ili ad hoc ljudska intervencija u sam proces procjene, te da se sve procjene temelje isključivo na unaprijed definiranim pravilima, statističkim kriterijima i modelskim postupcima. Time se osigurava da su dobivene vrijednosti što je moguće objektivnije, neovisne o pojedinačnim procjenama ili normativnim pretpostavkama istraživača, te u potpunosti replikabilne u slučaju ponovne primjene istog metodološkog okvira na iste ili nove podatke.

Takav je pristup osobito važan s obzirom na to da procijenjene cijene stanova i najmnine služe kao analitička osnovica za provedbu Programa potpora za stjecanje prve stambene nekretnine te Programa priuštivog najma, odnosno programa koji imaju izravne fiskalne implikacije i uključuju raspodjelu javnih sredstava građanima. U tom kontekstu, pravednost ne proizlazi iz naknadnih korekcija ili diskrecijskih odluka, već iz samog dizajna metodologije: jednake ulazne informacije, jednaki statistički kriteriji i jednaki modelski postupci moraju dovesti do jednakog tretmana svih jedinica lokalne samouprave i svih potencijalnih korisnika programa. Time se smanjuje rizik teritorijalnih nejednakosti koje bi proizlazile iz proizvoljnih ili selektivnih procjena, jača se legitimnost korištenja analitičkih rezultata u provedbi javnih politika te se povećava razina povjerenja u sustav raspodjele javnih sredstava. Upravo ta kombinacija metodološke dosljednosti, transparentnosti i potpune replikabilnosti predstavlja jednu od ključnih prednosti ove studije te, uz jasno uvažavanje njezinih analitičkih ograničenja, čini dobivene rezultate posebno prikladnima za primjenu u području stambenih politika, u kojima je nužno osigurati visoku razinu objektivnosti, usporedivosti i pravednosti među teritorijalnim jedinicama.

Literatura

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). *Fitting Linear Mixed-Effects Models Using lme4*. Journal of Statistical Software, 67(1), 1–48.
- Breiman, L. (2001). *Random forests*. Machine Learning, 45(1), 5–32.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth.
- Brown, L. D., Cai, T. T., & DasGupta, A. (2001). "Interval Estimation for a Binomial Proportion". Statistical Science, 16(2), 101–117.
- Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.
- Cover, T. M., & Hart, P. E. (1967). *Nearest neighbor pattern classification*. IEEE Transactions on Information Theory, 13(1), 21–27.
- Efron, B. (1979). *Bootstrap Methods: Another Look at the Jackknife*. The Annals of Statistics, 7(1), 1–26.
- Ekonomski institut, Zagreb (2024). *Modeliranje i projiciranje ponude i potražnje za stambenim jedinicama u jedinicama lokalne samouprave Republike Hrvatske*. Mimeo Ekonomski institut, Zagreb.
- Friedman, J. H. (2001). *Greedy function approximation: A gradient boosting machine*. Annals of Statistics, 29(5), 1189–1232.
- Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Geltner, D., Miller, N. G., Clayton, J., & Eichholtz, P. (2001). *Commercial Real Estate Analysis and Investments*. South-Western Educational Publishing.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York: Springer.
- Hoerl, A. E., & Kennard, R. W. (1970). *Ridge regression: Biased estimation for nonorthogonal problems*. Technometrics, 12(1), 55–67.
- Institut za turizam (2025). Indeks turističke razvijenosti za 2024. godinu. Dostupno na: <https://itzg.hr/hr/projekti/?tid=247>.
- Ministarstvo prostornoga razvoja, graditeljstva i državne imovine Republike Hrvatske (2025). *Nacionalni plan stambene politike Republike Hrvatske do 2030*. Dostupno na: https://mpgi.gov.hr/UserDocsImages/dokumenti/NPSP/2025_03_26_NPSP.pdf
- Ministarstvo regionalnog razvoja i fondova Europske Unije Republike Hrvatske (2024). *Vrijednost indeksa razvijenosti i pokazatelja za izračun indeksa razvijenost*. Dostupno na: <https://razvoj.gov.hr/UserDocsImages/O%20ministarstvu/Regionalni%20razvoj/>

indeks%20razvijenosti/Vrijednosti%20indeksa%20razvijenosti%20i%20pokazatelja%20za%20
izrac%CC%8Cun%20indeksa%20razvijenosti.pdf.

Narodne novine (2025). Pravilnik o potporama za stjecanje prve nekretnine. Br. 86/2025.

Tibshirani, R. (1996). *Regression shrinkage and selection via the Lasso*. Journal of the Royal Statistical Society: Series B, 58(1), 267–288.

Zou, H., & Hastie, T. (2005). *Regularization and variable selection via the elastic net*. Journal of the Royal Statistical Society: Series B, 67(2), 301–320.

Popis kratica

APN – Agencija za pravni promet i posredovanje nekretninama
CORINE – Coordination of Information on the Environment (baza podataka o pokrovu zemljišta)
DZS – Državni zavod za statistiku
EU – Europska unija
GIS – Geografski informacijski sustav
JLS – jedinica lokalne samouprave
KNN – K-najbližih susjeda (engl. <i>K-Nearest Neighbors</i>)
LASSO – Least Absolute Shrinkage and Selection Operator
MAE – srednja apsolutna pogreška (engl. <i>Mean Absolute Error</i>)
OLS – metoda najmanjih kvadrata (engl. <i>Ordinary Least Squares</i>)
PDV – porez na dodanu vrijednost
RMSE – korijen srednje kvadratne pogreške (engl. <i>Root Mean Squared Error</i>)
R ² – koeficijent determinacije
XGBoost – eXtreme Gradient Boosting

Popis slika

SLIKA 1.	Rang objašnjavajućih varijabli u izabranim modelima strojnog učenja	31
SLIKA 2.	Procijenjene medijalne cijene stanova s uključenim PDV-om	41
SLIKA 3.	Procijenjene mjesečne medijalne najamnine stanova	52
SLIKA 4.	Rang objašnjavajućih varijabli u izabranim modelima strojnog učenja	62
SLIKA 5.	Procijenjene prosječne cijene stanova bez PDV-a	72
SLIKA 6.	Usporedna stabilnost medijalne cijene stanova u izabranim gradovima	77
SLIKA P1.	Usporedna analiza stabilnosti medijalne cijene stanova u ovisnosti na broj kupoprodajnih transakcija u izabranim gradovima	107-111

Popis tablica

TABLICA 1.	Frekvencijska raspodjela broja jedinica lokalne samouprave prema broju transakcija ostvarenih na tržištu stanova u 2025.	10
TABLICA 2.	Usporedba performansi modela s punim skupom od 20 objašnjavajućih varijabli	28
TABLICA 3.	Višemodelska procjena važnosti objašnjavajućih varijabli	30
TABLICA 4.	Rangiranje važnosti varijabli prema pojedinim modelima i konsenzusnom rangui	32
TABLICA 5.	Najbolje specifikacije modela prema različitim kriterijima uspješnosti	35
TABLICA 6.	Usporedba modela za završnu specifikaciju (ir, itr, lag_target, dist_sea)	36
TABLICA 7.	Procjena fiksnih efekata u linearnom mješovitom modelu	39
TABLICA 8.	Varijanca slučajnih i rezidualnih komponenti	39
TABLICA 9.	Usporedba višestruke linearne regresije i linearnog mješovitog modela	41
TABLICA 10.	Ostvarene medijalne cijene stanova i najamnine te broj ugovora o kupoprodaji i najmu stanova u 2025.	48
TABLICA 11.	Omjer medijalne najamnine i medijalne cijene stana za županije i regije	49
TABLICA 12.	Medijalne cijene stanova i najamnina po katastarskim općinama Grada Zagreba	54
TABLICA 13.	Usporedba performansi modela s punim skupom od 20 objašnjavajućih varijabli	60
TABLICA 14.	Višemodelska procjena važnosti objašnjavajućih varijabli	61
TABLICA 15.	Rangiranje važnosti varijabli prema pojedinim modelima i konsenzusnom rangui	63
TABLICA 16.	Najbolje specifikacije modela prema različitim kriterijima uspješnosti	64
TABLICA 17.	Usporedba modela za završnu specifikaciju (ir, itr, lag_target, dist_sea, vacapart)	66
TABLICA 18.	Procjena fiksnih efekata u linearnom mješovitom modelu	69
TABLICA 19.	Varijanca slučajnih i rezidualnih komponenti	70
TABLICA 20.	Usporedba višestruke linearne regresije i linearnog mješovitog modela	71
TABLICA 21.	Koeficijenti varijacije i pragovi pouzdanosti medijalne cijene stanova s obzirom na broj transakcija	78
TABLICA P1.	Performanse alternativnih modela strojnog učenja za različite specifikacije objašnjavajućih varijabli u modelu s medijalnom cijenom stanova s PDV-om kao ciljanom varijablom	91
TABLICA P2.	Performanse alternativnih modela strojnog učenja za različite specifikacije objašnjavajućih varijabli u modelu s prosječnom cijenom stanova bez PDV-a kao ciljanom varijablom	96
TABLICA P3.	Prosječne cijene stanova bez PDV-a i medijalne cijene stanova s PDV-om po katastarskim općinama Grada Zagreba u 2025.	103
TABLICA P4.	Mjesečne medijalne najamnine za stanove po katastarskim općinama Grada Zagreba u 2025.	105

Prilozi Studiji

TABLICA P1. Performanse alternativnih modela strojnog učenja za različite specifikacije objašnjavajućih varijabli u modelu s medijalnom cijenom stanova s PDV-om kao ciljanom varijablom

Broj varijabli	Objašnjavajuće varijable korištene u modelu	Metoda strojnog učenja	MAE	RMSE	R ²	Prilagođeni R2
2	ir, itr	Linear	378,70	502,25	0,65	0,60
2	ir, itr	LASSO	377,97	502,75	0,65	0,60
2	ir, itr	Ridge	375,28	508,17	0,65	0,59
2	ir, itr	Elastic Net	375,81	505,05	0,65	0,59
2	ir, itr	Random Forest	417,24	579,16	0,54	0,46
2	ir, itr	Regression Tree	494,68	617,48	0,48	0,39
2	ir, itr	XGBoost	425,10	608,93	0,49	0,41
2	ir, lag_target	Linear	373,93	546,41	0,59	0,52
2	ir, lag_target	LASSO	371,23	546,15	0,59	0,52
2	ir, lag_target	Ridge	360,59	542,39	0,60	0,53
2	ir, lag_target	Elastic Net	368,11	545,30	0,59	0,52
2	ir, lag_target	Random Forest	387,34	561,97	0,57	0,49
2	ir, lag_target	Regression Tree	433,41	591,94	0,52	0,44
2	ir, lag_target	XGBoost	420,83	592,45	0,52	0,44
2	ir, dist_sea	Linear	379,16	533,89	0,61	0,54
2	ir, dist_sea	LASSO	379,60	534,57	0,61	0,54
2	ir, dist_sea	Ridge	378,57	538,05	0,60	0,54
2	ir, dist_sea	Elastic Net	379,43	534,50	0,61	0,54
2	ir, dist_sea	Random Forest	346,27	518,58	0,63	0,57
2	ir, dist_sea	Regression Tree	430,57	591,70	0,52	0,44
2	ir, dist_sea	XGBoost	445,47	610,29	0,49	0,40
2	ir, vacapart	Linear	428,92	604,02	0,50	0,42
2	ir, vacapart	LASSO	437,85	607,84	0,49	0,41
2	ir, vacapart	Ridge	444,27	611,25	0,49	0,40
2	ir, vacapart	Elastic Net	441,62	609,76	0,49	0,40
2	ir, vacapart	Random Forest	405,25	530,40	0,61	0,55
2	ir, vacapart	Regression Tree	416,62	556,11	0,58	0,50
2	ir, vacapart	XGBoost	367,04	534,29	0,61	0,54
2	itr, lag_target	Linear	381,12	532,03	0,61	0,55
2	itr, lag_target	LASSO	381,72	532,64	0,61	0,55
2	itr, lag_target	Ridge	384,19	534,58	0,61	0,54
2	itr, lag_target	Elastic Net	381,66	532,52	0,61	0,55

2	itr, lag_target	Random Forest	418,65	586,19	0,53	0,45
2	itr, lag_target	Regression Tree	420,63	557,68	0,57	0,50
2	itr, lag_target	XGBoost	425,55	589,05	0,52	0,44
2	itr, dist_sea	Linear	463,14	589,01	0,52	0,44
2	itr, dist_sea	LASSO	464,16	590,02	0,52	0,44
2	itr, dist_sea	Ridge	469,09	594,34	0,51	0,43
2	itr, dist_sea	Elastic Net	464,12	589,93	0,52	0,44
2	itr, dist_sea	Random Forest	480,07	617,65	0,48	0,39
2	itr, dist_sea	Regression Tree	519,34	646,48	0,43	0,33
2	itr, dist_sea	XGBoost	464,89	608,71	0,49	0,41
2	itr, vacapart	Linear	517,89	617,14	0,48	0,39
2	itr, vacapart	LASSO	522,87	621,81	0,47	0,38
2	itr, vacapart	Ridge	534,37	631,68	0,45	0,36
2	itr, vacapart	Elastic Net	534,60	632,47	0,45	0,36
2	itr, vacapart	Random Forest	606,50	721,48	0,29	0,17
2	itr, vacapart	Regression Tree	566,80	666,81	0,39	0,29
2	itr, vacapart	XGBoost	577,54	673,74	0,38	0,27
2	lag_target, dist_sea	Linear	455,94	617,37	0,48	0,39
2	lag_target, dist_sea	LASSO	454,83	617,39	0,48	0,39
2	lag_target, dist_sea	Ridge	448,97	618,12	0,48	0,39
2	lag_target, dist_sea	Elastic Net	453,75	617,40	0,48	0,39
2	lag_target, dist_sea	Random Forest	450,92	608,13	0,49	0,41
2	lag_target, dist_sea	Regression Tree	489,86	622,99	0,47	0,38
2	lag_target, dist_sea	XGBoost	430,43	612,18	0,49	0,40
2	lag_target, vacapart	Linear	463,45	633,41	0,45	0,36
2	lag_target, vacapart	LASSO	463,23	633,21	0,45	0,36
2	lag_target, vacapart	Ridge	461,81	633,85	0,45	0,36
2	lag_target, vacapart	Elastic Net	463,24	633,26	0,45	0,36
2	lag_target, vacapart	Random Forest	449,14	581,67	0,54	0,46
2	lag_target, vacapart	Regression Tree	476,87	597,82	0,51	0,43
2	lag_target, vacapart	XGBoost	434,34	594,71	0,51	0,43
2	dist_sea, vacapart	Linear	539,78	683,81	0,36	0,25
2	dist_sea, vacapart	LASSO	540,47	684,08	0,36	0,25
2	dist_sea, vacapart	Ridge	550,53	687,12	0,35	0,24
2	dist_sea, vacapart	Elastic Net	543,98	685,30	0,36	0,25
2	dist_sea, vacapart	Random Forest	482,48	601,11	0,50	0,42
2	dist_sea, vacapart	Regression Tree	508,83	645,91	0,43	0,33
2	dist_sea, vacapart	XGBoost	498,39	604,37	0,50	0,41
3	ir, itr, lag_target	Linear	303,57	466,76	0,70	0,62
3	ir, itr, lag_target	LASSO	302,60	467,27	0,70	0,62
3	ir, itr, lag_target	Ridge	296,25	466,76	0,70	0,62

3	ir, itr, lag_target	Elastic Net	302,01	467,11	0,70	0,62
3	ir, itr, lag_target	Random Forest	358,31	539,03	0,60	0,49
3	ir, itr, lag_target	Regression Tree	484,75	599,09	0,51	0,37
3	ir, itr, lag_target	XGBoost	393,00	553,09	0,58	0,47
3	ir, itr, dist_sea	Linear	311,13	473,28	0,69	0,61
3	ir, itr, dist_sea	LASSO	311,28	474,35	0,69	0,61
3	ir, itr, dist_sea	Ridge	312,49	479,07	0,68	0,60
3	ir, itr, dist_sea	Elastic Net	311,48	475,62	0,69	0,60
3	ir, itr, dist_sea	Random Forest	376,61	540,05	0,60	0,49
3	ir, itr, dist_sea	Regression Tree	424,22	564,39	0,56	0,44
3	ir, itr, dist_sea	XGBoost	372,58	545,57	0,59	0,48
3	ir, itr, vacapart	Linear	375,21	509,61	0,64	0,55
3	ir, itr, vacapart	LASSO	370,06	513,84	0,64	0,54
3	ir, itr, vacapart	Ridge	372,06	516,44	0,63	0,53
3	ir, itr, vacapart	Elastic Net	374,55	510,18	0,64	0,55
3	ir, itr, vacapart	Random Forest	437,58	567,15	0,56	0,44
3	ir, itr, vacapart	Regression Tree	390,97	538,96	0,60	0,49
3	ir, itr, vacapart	XGBoost	423,02	545,46	0,59	0,48
3	ir, lag_target, dist_sea	Linear	346,25	520,48	0,63	0,53
3	ir, lag_target, dist_sea	LASSO	346,84	520,77	0,63	0,53
3	ir, lag_target, dist_sea	Ridge	352,72	523,29	0,62	0,52
3	ir, lag_target, dist_sea	Elastic Net	351,17	523,08	0,62	0,52
3	ir, lag_target, dist_sea	Random Forest	385,86	548,22	0,59	0,47
3	ir, lag_target, dist_sea	Regression Tree	451,97	619,51	0,47	0,33
3	ir, lag_target, dist_sea	XGBoost	394,32	536,40	0,60	0,50
3	ir, lag_target, vacapart	Linear	369,59	542,14	0,60	0,49
3	ir, lag_target, vacapart	LASSO	368,55	542,13	0,60	0,49
3	ir, lag_target, vacapart	Ridge	361,75	540,46	0,60	0,49
3	ir, lag_target, vacapart	Elastic Net	368,49	542,02	0,60	0,49
3	ir, lag_target, vacapart	Random Forest	383,13	544,09	0,59	0,48
3	ir, lag_target, vacapart	Regression Tree	431,47	594,69	0,51	0,38
3	ir, lag_target, vacapart	XGBoost	376,31	512,52	0,64	0,54
3	ir, dist_sea, vacapart	Linear	382,26	529,76	0,61	0,51
3	ir, dist_sea, vacapart	LASSO	382,91	530,96	0,61	0,51
3	ir, dist_sea, vacapart	Ridge	384,01	541,19	0,60	0,49
3	ir, dist_sea, vacapart	Elastic Net	384,60	535,93	0,61	0,50
3	ir, dist_sea, vacapart	Random Forest	385,78	531,98	0,61	0,51
3	ir, dist_sea, vacapart	Regression Tree	446,45	579,37	0,54	0,41
3	ir, dist_sea, vacapart	XGBoost	340,25	498,80	0,66	0,57
3	itr, lag_target, dist_sea	Linear	381,78	532,74	0,61	0,50
3	itr, lag_target, dist_sea	LASSO	382,71	533,75	0,61	0,50

3	itr, lag_target, dist_sea	Ridge	385,37	538,00	0,60	0,49
3	itr, lag_target, dist_sea	Elastic Net	382,54	533,54	0,61	0,50
3	itr, lag_target, dist_sea	Random Forest	437,69	585,37	0,53	0,40
3	itr, lag_target, dist_sea	Regression Tree	425,51	579,62	0,54	0,41
3	itr, lag_target, dist_sea	XGBoost	430,62	591,09	0,52	0,39
3	itr, lag_target, vacapart	Linear	379,10	530,25	0,61	0,51
3	itr, lag_target, vacapart	LASSO	385,73	537,40	0,60	0,50
3	itr, lag_target, vacapart	Ridge	388,81	538,70	0,60	0,49
3	itr, lag_target, vacapart	Elastic Net	386,07	537,55	0,60	0,49
3	itr, lag_target, vacapart	Random Forest	461,54	584,74	0,53	0,40
3	itr, lag_target, vacapart	Regression Tree	476,87	597,82	0,51	0,38
3	itr, lag_target, vacapart	XGBoost	423,02	580,56	0,54	0,41
3	itr, dist_sea, vacapart	Linear	465,32	592,03	0,52	0,39
3	itr, dist_sea, vacapart	LASSO	476,43	601,19	0,50	0,37
3	itr, dist_sea, vacapart	Ridge	491,79	610,30	0,49	0,35
3	itr, dist_sea, vacapart	Elastic Net	470,97	596,90	0,51	0,38
3	itr, dist_sea, vacapart	Random Forest	508,94	618,62	0,47	0,33
3	itr, dist_sea, vacapart	Regression Tree	527,75	649,00	0,42	0,26
3	itr, dist_sea, vacapart	XGBoost	501,78	619,57	0,47	0,33
3	lag_target, dist_sea, vacapart	Linear	455,56	622,53	0,47	0,32
3	lag_target, dist_sea, vacapart	LASSO	455,40	622,49	0,47	0,32
3	lag_target, dist_sea, vacapart	Ridge	455,28	625,66	0,46	0,32
3	lag_target, dist_sea, vacapart	Elastic Net	454,81	623,58	0,47	0,32
3	lag_target, dist_sea, vacapart	Random Forest	443,29	583,00	0,53	0,41
3	lag_target, dist_sea, vacapart	Regression Tree	458,55	614,54	0,48	0,34
3	lag_target, dist_sea, vacapart	XGBoost	424,69	589,92	0,52	0,39
4	ir, itr, lag_target, dist_sea	Linear	292,84	463,01	0,71	0,59
4	ir, itr, lag_target, dist_sea	LASSO	292,15	464,02	0,70	0,59
4	ir, itr, lag_target, dist_sea	Ridge	281,54	463,48	0,71	0,59
4	ir, itr, lag_target, dist_sea	Elastic Net	286,56	465,96	0,70	0,58
4	ir, itr, lag_target, dist_sea	Random Forest	404,92	566,52	0,56	0,38
4	ir, itr, lag_target, dist_sea	Regression Tree	451,97	619,51	0,47	0,26

4	ir, itr, lag_target, dist_sea	XGBoost	402,83	561,57	0,57	0,39
4	ir, itr, lag_target, vacapart	Linear	305,63	469,66	0,70	0,58
4	ir, itr, lag_target, vacapart	LASSO	304,58	469,62	0,70	0,58
4	ir, itr, lag_target, vacapart	Ridge	300,16	472,65	0,69	0,57
4	ir, itr, lag_target, vacapart	Elastic Net	303,97	469,92	0,70	0,58
4	ir, itr, lag_target, vacapart	Random Forest	397,53	557,04	0,57	0,40
4	ir, itr, lag_target, vacapart	Regression Tree	431,47	594,69	0,51	0,32
4	ir, itr, lag_target, vacapart	XGBoost	374,12	518,26	0,63	0,48
4	ir, itr, dist_sea, vacapart	Linear	314,61	478,07	0,69	0,56
4	ir, itr, dist_sea, vacapart	LASSO	321,56	488,84	0,67	0,54
4	ir, itr, dist_sea, vacapart	Ridge	325,25	490,00	0,67	0,54
4	ir, itr, dist_sea, vacapart	Elastic Net	317,51	481,33	0,68	0,55
4	ir, itr, dist_sea, vacapart	Random Forest	427,42	571,03	0,55	0,37
4	ir, itr, dist_sea, vacapart	Regression Tree	424,22	564,39	0,56	0,39
4	ir, itr, dist_sea, vacapart	XGBoost	387,36	532,41	0,61	0,46
4	ir, lag_target, dist_sea, vacapart	Linear	346,69	522,49	0,63	0,48
4	ir, lag_target, dist_sea, vacapart	LASSO	346,87	523,09	0,62	0,47
4	ir, lag_target, dist_sea, vacapart	Ridge	359,07	530,91	0,61	0,46
4	ir, lag_target, dist_sea, vacapart	Elastic Net	348,28	526,33	0,62	0,47
4	ir, lag_target, dist_sea, vacapart	Random Forest	387,71	554,05	0,58	0,41
4	ir, lag_target, dist_sea, vacapart	Regression Tree	451,97	619,51	0,47	0,26
4	ir, lag_target, dist_sea, vacapart	XGBoost	395,96	530,76	0,61	0,46
4	itr, lag_target, dist_ sea, vacapart	Linear	379,29	530,55	0,61	0,46
4	itr, lag_target, dist_ sea, vacapart	LASSO	383,92	535,15	0,61	0,45

4	itr, lag_target, dist_sea, vacapart	Ridge	388,54	540,77	0,60	0,44
4	itr, lag_target, dist_sea, vacapart	Elastic Net	384,77	536,27	0,61	0,45
4	itr, lag_target, dist_sea, vacapart	Random Forest	464,45	597,21	0,51	0,31
4	itr, lag_target, dist_sea, vacapart	Regression Tree	425,51	579,62	0,54	0,35
4	itr, lag_target, dist_sea, vacapart	XGBoost	445,68	583,07	0,53	0,35
5	ir, itr, lag_target, dist_sea, vacapart	Linear	294,85	465,62	0,70	0,54
5	ir, itr, lag_target, dist_sea, vacapart	LASSO	292,82	466,27	0,70	0,54
5	ir, itr, lag_target, dist_sea, vacapart	Ridge	293,97	478,19	0,69	0,51
5	ir, itr, lag_target, dist_sea, vacapart	Elastic Net	287,47	468,40	0,70	0,53
5	ir, itr, lag_target, dist_sea, vacapart	Random Forest	402,39	554,28	0,58	0,34
5	ir, itr, lag_target, dist_sea, vacapart	RegressionTree	451,97	619,51	0,47	0,18
5	ir, itr, lag_target, dist_sea, vacapart	XGBoost	394,67	523,70	0,62	0,41

Izvor: izračun autorice.

TABLICA P2. Performanse alternativnih modela strojnog učenja za različite specifikacije objašnjavajućih varijabli u modelu s prosječnom cijenom stanova bez PDV-a kao ciljanom varijablom

Broj varijabli	Objašnjavajuće varijable korištene u modelu	Metoda strojnog učenja	MAE	RMSE	R2	Prilagođeni R2
2	ir, itr	Linear	369,89	476,93	0,69	0,63
2	ir, itr	LASSO	370,11	477,12	0,69	0,63
2	ir, itr	Ridge	373,50	482,35	0,68	0,63
2	ir, itr	Elastic Net	370,18	477,22	0,69	0,63
2	ir, itr	Random Forest	381,69	531,71	0,61	0,54
2	ir, itr	Regression Tree	403,28	516,30	0,63	0,57
2	ir, itr	XGBoost	399,43	547,50	0,59	0,52
2	ir, vacapart	Linear	433,35	604,43	0,50	0,41
2	ir, vacapart	LASSO	443,05	607,72	0,49	0,41
2	ir, vacapart	Ridge	455,58	614,28	0,48	0,39
2	ir, vacapart	Elastic Net	438,79	606,13	0,49	0,41
2	ir, vacapart	Random Forest	385,58	486,33	0,67	0,62

2	ir, vacapart	Regression Tree	431,25	546,32	0,59	0,52
2	ir, vacapart	XGBoost	382,02	525,31	0,62	0,56
2	ir, lag_target	Linear	357,07	558,09	0,57	0,50
2	ir, lag_target	LASSO	355,99	558,00	0,57	0,50
2	ir, lag_target	Ridge	347,45	554,21	0,58	0,51
2	ir, lag_target	Elastic Net	355,04	557,62	0,57	0,50
2	ir, lag_target	Random Forest	362,93	549,33	0,58	0,51
2	ir, lag_target	Regression Tree	468,87	586,50	0,53	0,45
2	ir, lag_target	XGBoost	391,24	579,68	0,54	0,46
2	ir, dist_sea	Linear	389,74	538,02	0,60	0,53
2	ir, dist_sea	LASSO	388,15	541,19	0,60	0,53
2	ir, dist_sea	Ridge	386,20	540,19	0,60	0,53
2	ir, dist_sea	Elastic Net	388,50	539,36	0,60	0,53
2	ir, dist_sea	Random Forest	379,83	519,65	0,63	0,57
2	ir, dist_sea	Regression Tree	444,89	574,44	0,54	0,47
2	ir, dist_sea	XGBoost	433,96	586,23	0,53	0,45
2	itr, vacapart	Linear	479,60	558,49	0,57	0,50
2	itr, vacapart	LASSO	492,05	573,92	0,55	0,47
2	itr, vacapart	Ridge	496,73	579,31	0,54	0,46
2	itr, vacapart	Elastic Net	492,13	573,92	0,55	0,47
2	itr, vacapart	Random Forest	582,04	659,29	0,40	0,30
2	itr, vacapart	Regression Tree	575,46	676,16	0,37	0,26
2	itr, vacapart	XGBoost	572,09	650,24	0,42	0,32
2	itr, lag_target	Linear	340,62	516,98	0,63	0,57
2	itr, lag_target	LASSO	341,97	517,85	0,63	0,57
2	itr, lag_target	Ridge	348,07	521,64	0,62	0,56
2	itr, lag_target	Elastic Net	341,89	517,76	0,63	0,57
2	itr, lag_target	Random Forest	399,53	539,16	0,60	0,53
2	itr, lag_target	Regression Tree	484,38	635,16	0,44	0,35
2	itr, lag_target	XGBoost	400,84	537,36	0,60	0,54
2	itr, dist_sea	Linear	431,60	551,13	0,58	0,51
2	itr, dist_sea	LASSO	432,89	552,17	0,58	0,51
2	itr, dist_sea	Ridge	439,88	558,29	0,57	0,50
2	itr, dist_sea	Elastic Net	432,88	552,17	0,58	0,51
2	itr, dist_sea	Random Forest	454,87	576,90	0,54	0,46
2	itr, dist_sea	Regression Tree	476,94	592,66	0,52	0,43
2	itr, dist_sea	XGBoost	419,66	559,85	0,57	0,50
2	vacapart, lag_target	Linear	445,21	641,44	0,43	0,34
2	vacapart, lag_target	LASSO	446,72	640,88	0,43	0,34
2	vacapart, lag_target	Ridge	453,29	641,19	0,43	0,34
2	vacapart, lag_target	Elastic Net	448,03	640,74	0,43	0,34

2	vacapart, lag_target	Random Forest	411,60	547,00	0,59	0,52
2	vacapart, lag_target	Regression Tree	468,26	627,19	0,46	0,37
2	vacapart, lag_target	XGBoost	400,09	527,65	0,62	0,55
2	vacapart, dist_sea	Linear	518,27	676,47	0,37	0,26
2	vacapart, dist_sea	LASSO	518,28	677,03	0,37	0,26
2	vacapart, dist_sea	Ridge	520,20	678,85	0,36	0,26
2	vacapart, dist_sea	Elastic Net	518,56	678,68	0,36	0,26
2	vacapart, dist_sea	Random Forest	477,86	587,80	0,52	0,44
2	vacapart, dist_sea	Regression Tree	524,40	635,03	0,44	0,35
2	vacapart, dist_sea	XGBoost	485,94	576,69	0,54	0,46
2	lag_target, dist_sea	Linear	434,73	623,42	0,46	0,37
2	lag_target, dist_sea	LASSO	433,91	623,83	0,46	0,37
2	lag_target, dist_sea	Ridge	434,15	623,95	0,46	0,37
2	lag_target, dist_sea	Elastic Net	433,99	623,64	0,46	0,37
2	lag_target, dist_sea	Random Forest	424,74	580,92	0,53	0,46
2	lag_target, dist_sea	Regression Tree	472,58	630,34	0,45	0,36
2	lag_target, dist_sea	XGBoost	402,40	571,04	0,55	0,48
3	ir, itr, vacapart	Linear	368,50	478,47	0,68	0,60
3	ir, itr, vacapart	LASSO	372,58	482,20	0,68	0,59
3	ir, itr, vacapart	Ridge	370,06	485,83	0,67	0,59
3	ir, itr, vacapart	Elastic Net	369,52	479,78	0,68	0,60
3	ir, itr, vacapart	Random Forest	409,19	511,39	0,64	0,54
3	ir, itr, vacapart	Regression Tree	431,25	546,32	0,59	0,48
3	ir, itr, vacapart	XGBoost	405,66	498,20	0,66	0,56
3	ir, itr, lag_target	Linear	287,58	449,53	0,72	0,65
3	ir, itr, lag_target	LASSO	285,00	449,86	0,72	0,64
3	ir, itr, lag_target	Ridge	278,59	452,41	0,72	0,64
3	ir, itr, lag_target	Elastic Net	285,17	449,85	0,72	0,64
3	ir, itr, lag_target	Random Forest	350,05	507,28	0,64	0,55
3	ir, itr, lag_target	Regression Tree	481,90	592,89	0,52	0,38
3	ir, itr, lag_target	XGBoost	361,45	540,25	0,60	0,49
3	ir, itr, dist_sea	Linear	329,05	452,48	0,72	0,64
3	ir, itr, dist_sea	LASSO	327,95	452,92	0,72	0,64
3	ir, itr, dist_sea	Ridge	326,38	458,01	0,71	0,63
3	ir, itr, dist_sea	Elastic Net	327,97	452,97	0,72	0,64
3	ir, itr, dist_sea	Random Forest	373,96	511,27	0,64	0,54
3	ir, itr, dist_sea	Regression Tree	455,93	574,24	0,55	0,42
3	ir, itr, dist_sea	XGBoost	410,10	545,82	0,59	0,48
3	ir, vacapart, lag_target	Linear	369,68	551,03	0,58	0,47
3	ir, vacapart, lag_target	LASSO	367,20	551,15	0,58	0,47

3	ir, vacapart, lag_target	Ridge	361,43	549,04	0,58	0,47
3	ir, vacapart, lag_target	Elastic Net	368,54	550,93	0,58	0,47
3	ir, vacapart, lag_target	Random Forest	354,01	497,24	0,66	0,57
3	ir, vacapart, lag_target	Regression Tree	431,25	546,32	0,59	0,48
3	ir, vacapart, lag_target	XGBoost	359,28	465,52	0,70	0,62
3	ir, vacapart, dist_sea	Linear	382,79	530,22	0,61	0,51
3	ir, vacapart, dist_sea	LASSO	383,35	531,91	0,61	0,50
3	ir, vacapart, dist_sea	Ridge	380,84	533,23	0,61	0,50
3	ir, vacapart, dist_sea	Elastic Net	382,89	531,69	0,61	0,50
3	ir, vacapart, dist_sea	Random Forest	382,34	502,69	0,65	0,56
3	ir, vacapart, dist_sea	Regression Tree	483,31	590,34	0,52	0,39
3	ir, vacapart, dist_sea	XGBoost	396,84	473,01	0,69	0,61
3	ir, lag_target, dist_sea	Linear	342,41	524,33	0,62	0,52
3	ir, lag_target, dist_sea	LASSO	342,90	526,01	0,62	0,51
3	ir, lag_target, dist_sea	Ridge	343,77	525,34	0,62	0,52
3	ir, lag_target, dist_sea	Elastic Net	344,35	526,85	0,62	0,51
3	ir, lag_target, dist_sea	Random Forest	363,34	531,96	0,61	0,50
3	ir, lag_target, dist_sea	Regression Tree	444,89	574,44	0,54	0,42
3	ir, lag_target, dist_sea	XGBoost	352,81	510,07	0,64	0,54
3	itr, vacapart, lag_target	Linear	332,80	508,02	0,64	0,55
3	itr, vacapart, lag_target	LASSO	350,15	523,74	0,62	0,52
3	itr, vacapart, lag_target	Ridge	346,81	520,13	0,63	0,52
3	itr, vacapart, lag_target	Elastic Net	352,46	525,30	0,62	0,52
3	itr, vacapart, lag_target	Random Forest	441,16	542,02	0,59	0,48
3	itr, vacapart, lag_target	Regression Tree	468,26	627,19	0,46	0,31
3	itr, vacapart, lag_target	XGBoost	416,82	533,88	0,61	0,50
3	itr, vacapart, dist_sea	Linear	419,39	541,73	0,60	0,48
3	itr, vacapart, dist_sea	LASSO	443,69	561,50	0,57	0,45

3	itr, vacapart, dist_sea	Ridge	435,45	554,89	0,58	0,46
3	itr, vacapart, dist_sea	Elastic Net	445,72	563,51	0,56	0,44
3	itr, vacapart, dist_sea	Random Forest	491,84	583,12	0,53	0,40
3	itr, vacapart, dist_sea	Regression Tree	514,55	630,95	0,45	0,30
3	itr, vacapart, dist_sea	XGBoost	499,82	588,69	0,52	0,39
3	itr, lag_target, dist_sea	Linear	340,07	516,76	0,63	0,53
3	itr, lag_target, dist_sea	LASSO	341,43	517,69	0,63	0,53
3	itr, lag_target, dist_sea	Ridge	365,72	534,62	0,61	0,50
3	itr, lag_target, dist_sea	Elastic Net	347,64	522,62	0,62	0,52
3	itr, lag_target, dist_sea	Random Forest	405,45	551,16	0,58	0,47
3	itr, lag_target, dist_sea	Regression Tree	454,80	625,08	0,46	0,31
3	itr, lag_target, dist_sea	XGBoost	442,46	582,05	0,53	0,41
3	vacapart, lag_target, dist_sea	Linear	435,13	626,42	0,46	0,31
3	vacapart, lag_target, dist_sea	LASSO	436,97	626,51	0,46	0,31
3	vacapart, lag_target, dist_sea	Ridge	445,08	628,74	0,45	0,31
3	vacapart, lag_target, dist_sea	Elastic Net	441,00	627,26	0,46	0,31
3	vacapart, lag_target, dist_sea	Random Forest	403,59	558,30	0,57	0,45
3	vacapart, lag_target, dist_sea	Regression Tree	447,80	616,07	0,48	0,33
3	vacapart, lag_target, dist_sea	XGBoost	401,97	536,20	0,60	0,50
4	ir, itr, vacapart, lag_target	Linear	286,68	447,42	0,72	0,61
4	ir, itr, vacapart, lag_target	LASSO	278,97	451,73	0,72	0,61
4	ir, itr, vacapart, lag_target	Ridge	280,57	454,29	0,72	0,60
4	ir, itr, vacapart, lag_target	Elastic Net	277,05	453,27	0,72	0,60
4	ir, itr, vacapart, lag_target	Random Forest	370,90	504,47	0,65	0,51
4	ir, itr, vacapart, lag_target	Regression Tree	431,25	546,32	0,59	0,42
4	ir, itr, vacapart, lag_target	XGBoost	410,84	524,78	0,62	0,47
4	ir, itr, vacapart, dist_sea	Linear	330,07	450,00	0,72	0,61

4	ir, itr, vacapart, dist_sea	LASSO	324,89	457,98	0,71	0,59
4	ir, itr, vacapart, dist_sea	Ridge	326,33	456,50	0,71	0,60
4	ir, itr, vacapart, dist_sea	Elastic Net	325,38	457,67	0,71	0,60
4	ir, itr, vacapart, dist_sea	Random Forest	408,62	517,70	0,63	0,48
4	ir, itr, vacapart, dist_sea	Regression Tree	483,31	590,34	0,52	0,33
4	ir, itr, vacapart, dist_sea	XGBoost	386,75	503,68	0,65	0,51
4	ir, itr, lag_target, dist_sea	Linear	283,36	442,21	0,73	0,62
4	ir, itr, lag_target, dist_sea	LASSO	281,20	442,65	0,73	0,62
4	ir, itr, lag_target, dist_sea	Ridge	284,43	452,36	0,72	0,60
4	ir, itr, lag_target, dist_sea	Elastic Net	277,50	446,72	0,72	0,61
4	ir, itr, lag_target, dist_sea	Random Forest	391,68	533,08	0,61	0,45
4	ir, itr, lag_target, dist_sea	Regression Tree	455,93	574,24	0,55	0,36
4	ir, itr, lag_target, dist_sea	XGBoost	396,71	547,89	0,59	0,42
4	ir, vacapart, lag_ target, dist_sea	Linear	345,75	523,44	0,62	0,47
4	ir, vacapart, lag_ target, dist_sea	LASSO	346,85	524,99	0,62	0,47
4	ir, vacapart, lag_ target, dist_sea	Ridge	354,21	528,40	0,61	0,46
4	ir, vacapart, lag_ target, dist_sea	Elastic Net	349,90	527,23	0,62	0,46
4	ir, vacapart, lag_ target, dist_sea	Random Forest	400,67	531,92	0,61	0,45
4	ir, vacapart, lag_ target, dist_sea	Regression Tree	483,31	590,34	0,52	0,33
4	ir, vacapart, lag_ target, dist_sea	XGBoost	372,35	476,57	0,69	0,56
4	itr, vacapart, lag_ target, dist_sea	Linear	332,56	505,92	0,65	0,51
4	itr, vacapart, lag_ target, dist_sea	LASSO	356,99	528,66	0,61	0,46
4	itr, vacapart, lag_ target, dist_sea	Ridge	344,38	519,75	0,63	0,48
4	itr, vacapart, lag_ target, dist_sea	Elastic Net	349,51	523,90	0,62	0,47

4	itr, vacapart, lag_ target, dist_sea	Random Forest	445,12	567,11	0,56	0,38
4	itr, vacapart, lag_ target, dist_sea	Regression Tree	454,80	625,08	0,46	0,25
4	itr, vacapart, lag_ target, dist_sea	XGBoost	441,72	559,43	0,57	0,40
5	ir, itr, vacapart, lag_ target, dist_sea	Linear	282,74	438,43	0,73	0,59
5	ir, itr, vacapart, lag_ target, dist_sea	LASSO	276,25	444,82	0,73	0,58
5	ir, itr, vacapart, lag_ target, dist_sea	Ridge	282,39	448,86	0,72	0,57
5	ir, itr, vacapart, lag_ target, dist_sea	Elastic Net	279,43	450,39	0,72	0,56
5	ir, itr, vacapart, lag_ target, dist_sea	Random Forest	392,36	521,11	0,63	0,42
5	ir, itr, vacapart, lag_ target, dist_sea	Regression Tree	483,31	590,34	0,52	0,25
5	ir, itr, vacapart, lag_ target, dist_sea	XGBoost	380,01	507,06	0,65	0,45

Izvor: izračun autorice.

TABLICA P3. Prosječne cijene stanova bez PDV-a i medijalne cijene stanova s PDV-om po katastarskim općinama Grada Zagreba u 2025.

Katastarska općina	Broj kupoprodaja	Prosječna cijena stana bez PDV-a (EUR/m ²)	Medijalna cijena stana s PDV-om (EUR/m ²)	Indeks prosječne cijene bez PDV-a (Grad Zagreb = 100)	Indeks medijalne cijene s PDV-om (Grad Zagreb = 100)
Adamovec	-	1.564	1.799	58,0	61,9
Blaguša	-	1.564	1.799	58,0	61,9
Blato	711	2.401	2.769	89,1	95,3
Brezovica	14	2.300	2.451	85,4	84,4
Čehi	-	2.336	2.496	86,7	85,9
Centar	332	3.413	3.379	126,7	116,3
Centar Novi	417	3.307	3.246	122,7	111,7
Črnomerec	288	3.006	3.125	111,6	107,6
Čučerje Novo	-	2.224	2.329	82,5	80,2
Demerje	-	2.300	2.451	85,4	84,4
Donja Dubrava	33	2.407	2.926	89,3	100,7
Donje Vrapče	321	3.065	3.158	113,8	108,7
Dragonožec	-	2.056	2.367	76,3	81,5
Dubrava	172	2.460	2.694	91,3	92,7
Đurđekovec	-	1.564	1.799	58,0	61,9
Glavica	-	1.564	1.799	58,0	61,9
Goranec	-	1.564	1.799	58,0	61,9
Gornje Vrapče	19	2.849	3.221	105,7	110,9
Gornji Stenjevec	30	2.668	2.874	99,0	98,9
Gračani	48	2.224	2.329	82,5	80,2
Granešina	243	2.405	2.893	89,3	99,6
Granešina Nova	25	2.170	2.223	80,5	76,5
Horvati	-	2.300	2.451	85,4	84,4
Jakuševac	-	2.577	2.814	95,6	96,9
Jakuševac Novi	76	2.577	2.814	95,6	96,9
Kašina	-	1.564	1.799	58,0	61,9
Klara	452	2.336	2.496	86,7	85,9
Klara Nova	87	2.473	2.581	91,8	88,8
Kupinečki Kraljevec	-	2.300	2.451	85,4	84,4
Lučko	17	1.909	2.181	70,9	75,1
Lužan	-	1.564	1.799	58,0	61,9
Maksimir	46	3.170	3.288	117,7	113,2
Markuševac	-	2.224	2.329	82,5	80,2
Mikulići	38	2.557	2.539	94,9	87,4
Odra	106	2.056	2.367	76,3	81,5

Odranski Obrež	-	2.056	2.367	76,3	81,5
Odranski Obrež Novi	-	2.056	2.367	76,3	81,5
Peščenica	642	2.642	3.090	98,1	106,4
Planina	-	1.564	1.799	58,0	61,9
Podsused	67	2.124	2.328	78,8	80,1
Remete	37	2.578	2.769	95,7	95,3
Resnik	96	2.353	2.576	87,3	88,7
Rudeš	141	2.954	3.049	109,6	105,0
Šašinovec	39	1.564	1.799	58,0	61,9
Šestine	58	2.808	2.829	104,2	97,4
Sesvete	141	1.828	2.087	67,8	71,8
Sesvete Novo	107	1.988	2.141	73,8	73,7
Sesvetski Kraljevec	149	1.709	1.964	63,4	67,6
Starjak	-	2.056	2.367	76,3	81,5
Stenjevec	265	2.710	2.950	100,6	101,5
Stenjevec Jug	258	2.672	3.202	99,2	110,2
Stenjevec Sjever	-	2.710	2.950	100,6	101,5
Trešnjevka	166	3.052	3.067	113,3	105,6
Trešnjevka Nova	185	2.955	3.029	109,7	104,3
Trnje	321	3.184	3.263	118,2	112,3
Trpuci	-	2.056	2.367	76,3	81,5
Vrapče	137	2.686	2.980	99,7	102,6
Vrapče Novo	-	2.686	2.980	99,7	102,6
Vugrovec	-	1.564	1.799	58,0	61,9
Vurnovec	-	1.564	1.799	58,0	61,9
Zaprudski Otok	419	3.022	2.928	112,2	100,8
Žitnjak	38	2.743	2.820	101,8	97,1
Grad Zagreb	6.778	2.694	2.905	100,0	100,0

Napomena: Za katastarske općine u kojima nije zabilježen dovoljan broj kupoprodajnih transakcija za izravan izračun cijena, prosječne i medijalne cijene procijenjene su imputacijom na temelju vrijednosti najbliže katastarske općine s dovoljnim brojem transakcija, pri čemu je prostorna bliskost definirana geografskom udaljenošću.

Izvor: izračun autorice.

TABLICA P4. Mjesečne medijalne najamnine za stanove po katastarskim općinama Grada Zagreba u 2025.

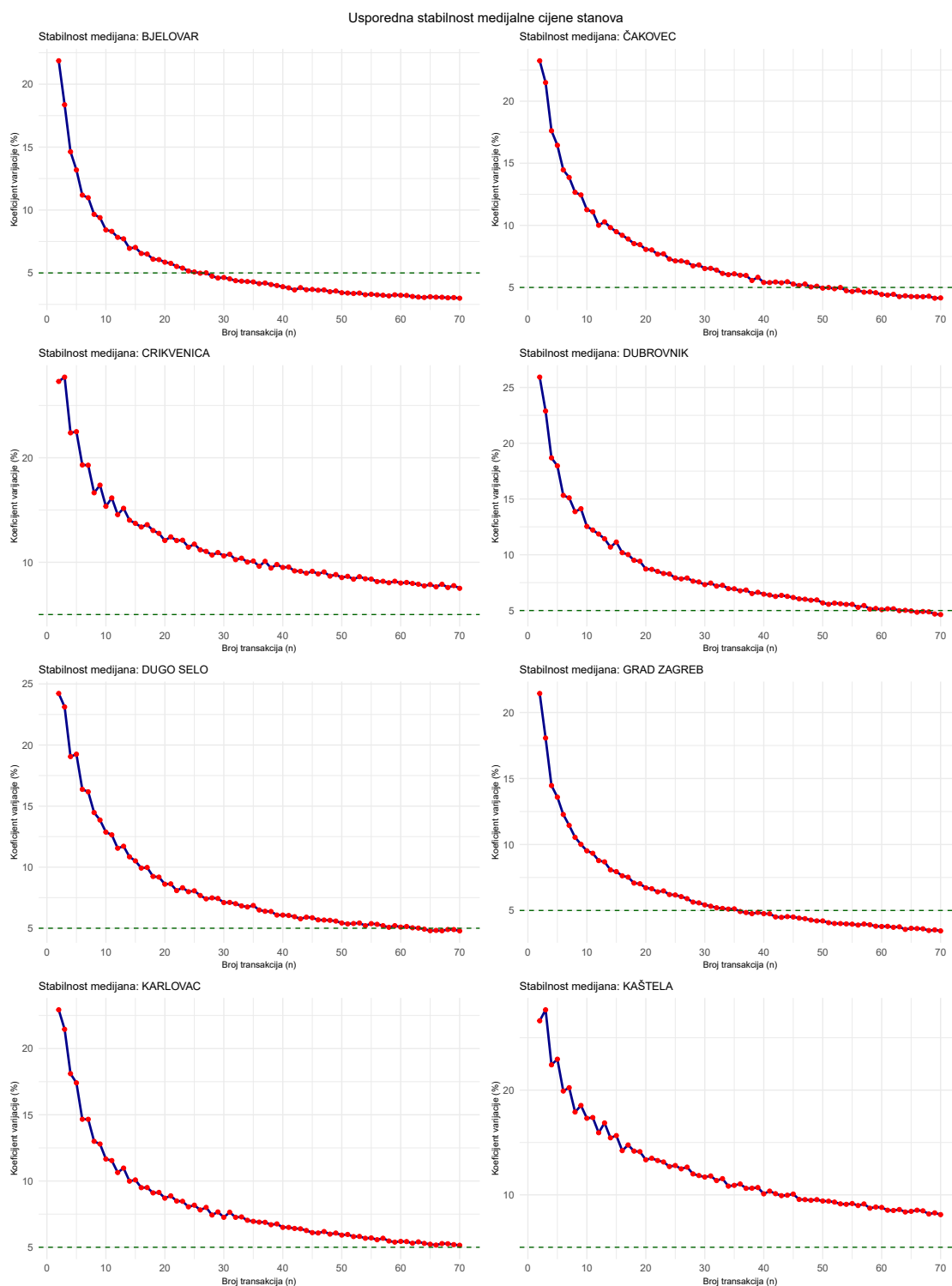
Katastarska općina	Procijenjena mjesečna medijalna najamnina (EUR/m ²)	Indeks mjesečne medijalne najamnine (Grad Zagreb = 100)
Adamovec	7,9	62,2
Blaguša	7,9	62,2
Blato	14,2	111,8
Brezovica	10,7	84,3
Čehi	10,7	84,3
Centar	14,2	111,8
Centar Novi	14,55	114,6
Črnomerec	12,65	99,6
Čučerje Novo	9,8	77,2
Demerje	11	86,6
Donja Dubrava	12,9	101,5
Donje Vrapče	12,75	100,4
Dragonožec	10,7	84,3
Dubrava	12	94,5
Đurđekovec	9,2	72,4
Glavica	7,9	62,2
Goranec	9,8	77,2
Gornje Vrapče	10,7	84,3
Gornji Stenjevec	12,7	99,7
Gračani	10,3	80,8
Granešina	12,7	100,3
Granešina Nova	9,8	77,1
Horvati	10,7	84,3
Jakuševac	13,8	108,7
Jakuševac Novi	12,4	97,6
Kašina	7,9	62,2
Klara	12,5	98,4
Klara Nova	13,1	103,1
Kupinečki Kraljevec	10,7	84,3
Lučko	11	86,6
Lužan	7,9	62,2
Maksimir	14,5	114,0
Markuševac	10,3	81,1
Mikulići	11,2	88,0
Odra	10,7	84,3
Odranski Obrež	10,7	84,3

Odranski Obrež Novi	10,7	84,3
Peščenica	12,7	100,0
Planina	9,8	77,2
Podsused	10,3	80,7
Remete	12,2	96,0
Resnik	11,3	89,3
Rudeš	13,3	104,7
Šašinovec	7,9	62,4
Šestine	12,5	98,1
Sesvete	9,2	72,4
Sesvete Novo	9,4	74,2
Sesvetski Kraljevec	9,9	78,0
Starjak	10,7	84,3
Stenjevec	10,8	85,0
Stenjevec Jug	13,2	103,9
Stenjevec Sjever	10,8	85,0
Trešnjevka	13,9	109,4
Trešnjevka Nova	12,4	97,6
Trnje	12,9	101,6
Trpuci	10,7	84,3
Vrapče	10,7	84,3
Vrapče Novo	10,7	84,3
Vugrovec	9,8	77,2
Vurnovec	7,9	62,2
Zaprudski Otok	12,25	96,5
Žitnjak	13,55	106,7
Grad Zagreb	12,7	100,0

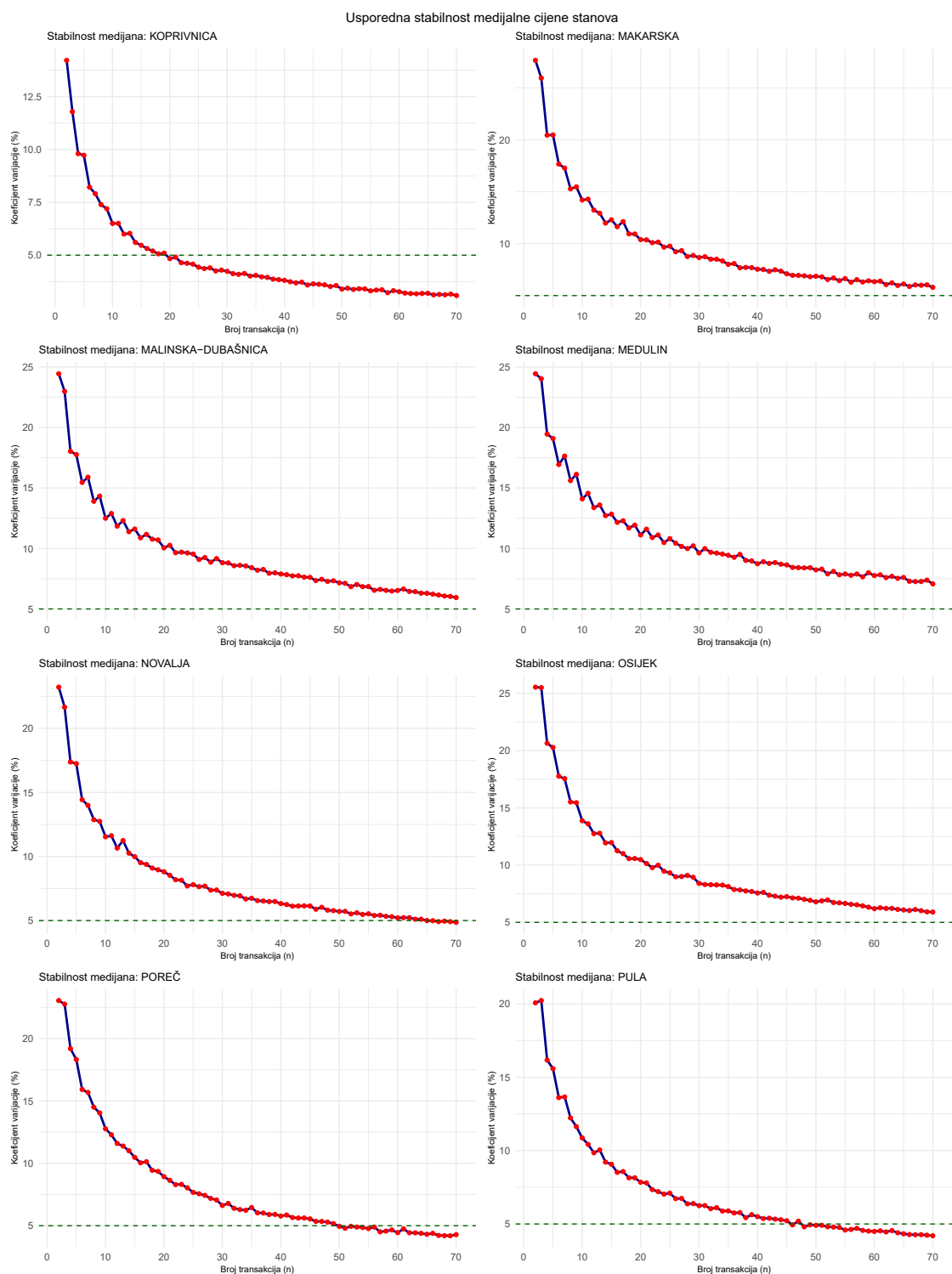
Napomena: Za katastarske općine s manje od 20 zabilježenih ugovora o najmu stanova, ali s 20 ili više kupoprodajnih transakcija u 2025. godini, medijalna najamnina procijenjena je kao umnožak stvarno ostvarene medijalne cijene stana u toj katastarskoj općini i udjela medijalne najamnine u medijalnoj cijeni stana za Grad Zagreb. Za katastarske općine koje nemaju dovoljan broj ni kupoprodajnih transakcija ni ugovora o najmu, medijalna najamnina procijenjena je imputacijom na temelju vrijednosti najbliže katastarske općine za koju postoji stvarno zabilježena ili procijenjena medijalna najamnina, pri čemu je prostorna bliskost definirana geografskom udaljenošću.

Izvor: izračun autorice.

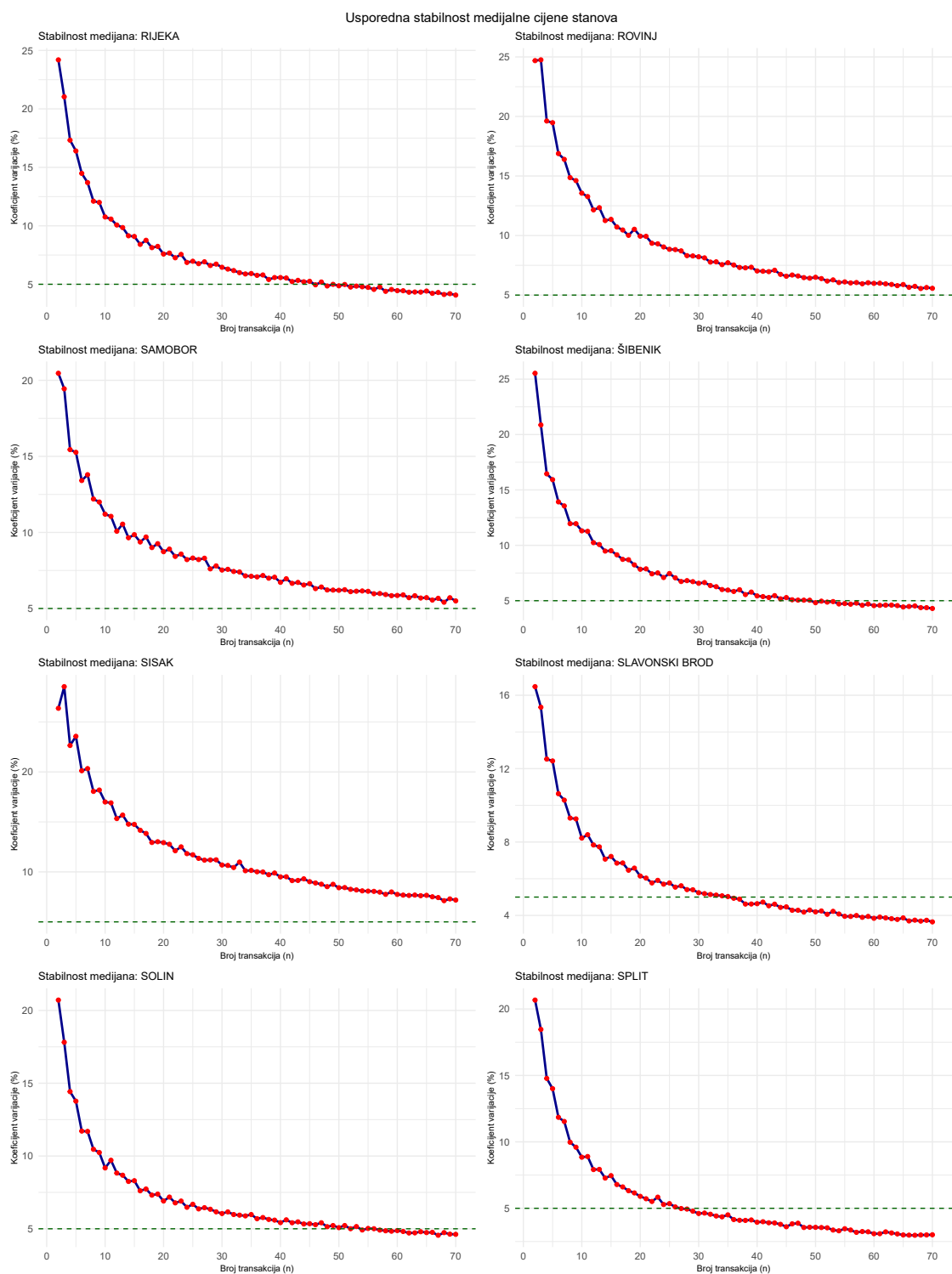
SLIKA P1. Usporedna analiza stabilnosti medijalne cijene stanova u ovisnosti na broj kupoprodajnih transakcija u izabranim gradovima



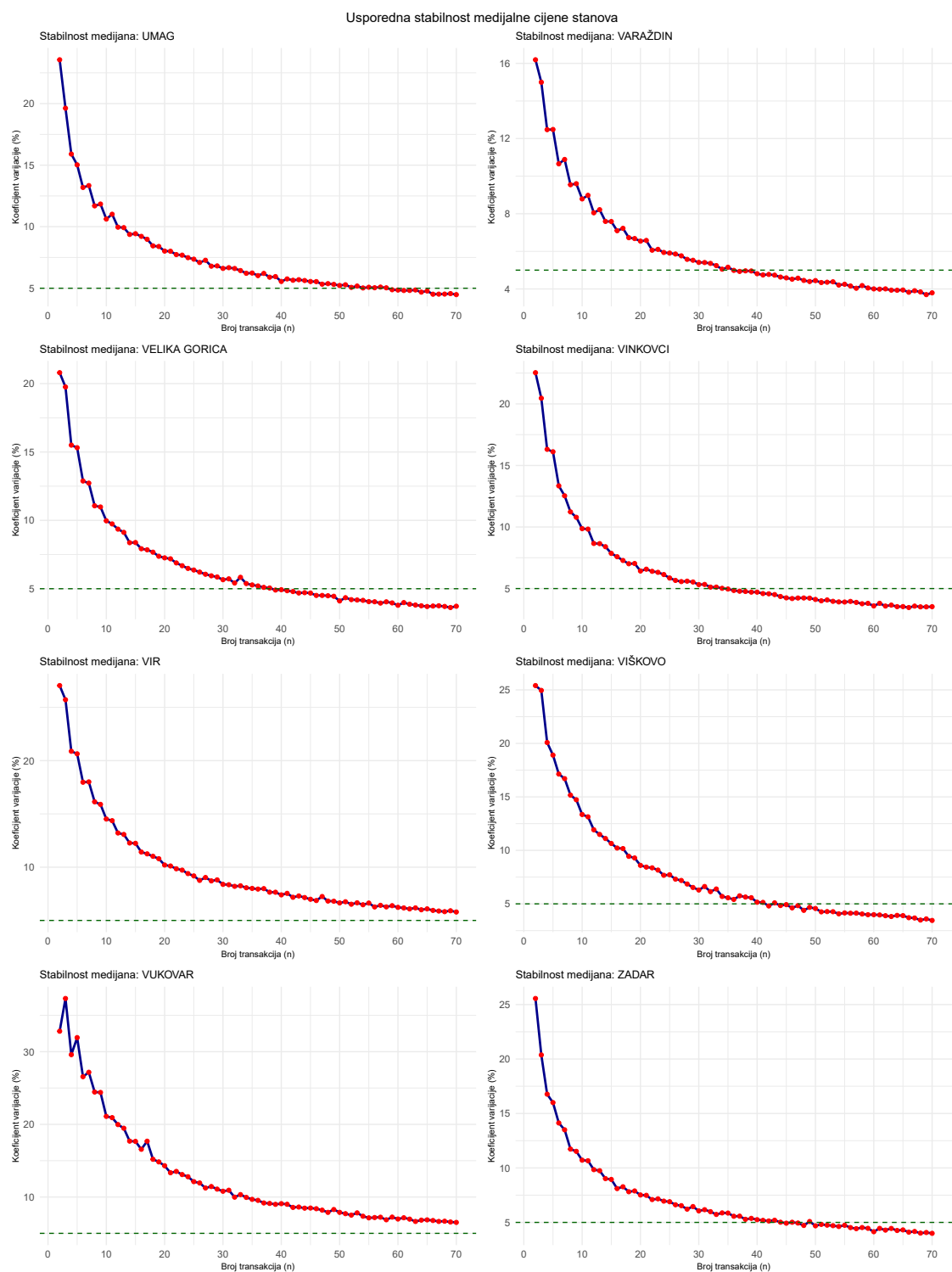
SLIKA P1. Usporedna analiza stabilnosti medijalne cijene stana u ovisnosti na broj kupoprodajnih transakcija u jedinicama lokalne samouprave



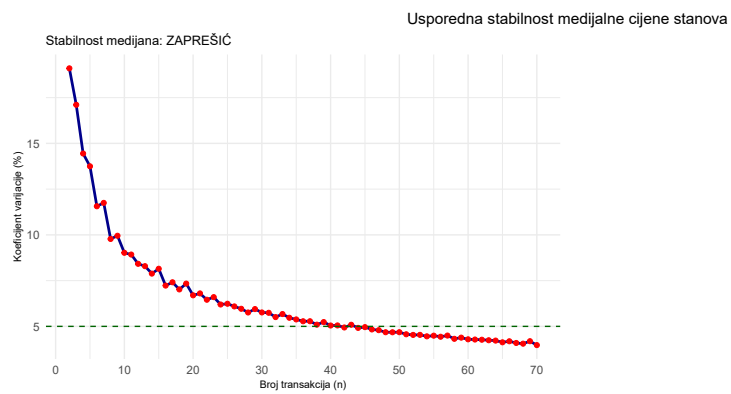
SLIKA P1. Usporedna analiza stabilnosti medijalne cijene stana u ovisnosti na broj kupoprodajnih transakcija u jedinicama lokalne samouprave



SLIKA P1. Usporedna analiza stabilnosti medijalne cijene stana u ovisnosti na broj kupoprodajnih transakcija u jedinicama lokalne samouprave



SLIKA P1. Usporedna analiza stabilnosti medijalne cijene stana u ovisnosti na broj kupoprodajnih transakcija u jedinicama lokalne samouprave



Izvor: izračun autorice.

