

**ПРЕДВИЂАЊЕ ЦЕНА НЕКРЕТНИНА У МОСКВИ ПРИМЕНОМ ХЕДОНИСТИЧКИХ
МОДЕЛА И ПРОТИВРЕЧНИХ СЛУЧАЈЕВА
PREDICTION OF REAL ESTATE PRICES IN MOSCOW USING HEDONISTIC
MODELS AND COUNTERFACTUAL EXAMPLES**

Ђорђе Радивојевић, Сандро Радовановић, др Борис Делибашић

РЕЗИМЕ: Приликом предвиђања цене некретнине потребно је водити рачуна о неколико ствари. Наиме, поред тачности предвиђања потребно је водити рачуна да се предвиђање објасни и да се обезбеди механизам који омогућава шта-ако анализу. Међутим, објашњење одлуке и шта-ако анализа захтевају узрочно-последичну анализу која је по природи веома комплексна. У овом раду поступак предвиђања, објашњења одлуке и узрочно-последичну анализу обезбеђујемо комбинацијом хедонистичких модела и противречних примера. Наиме, хедонистички модели служе да дају предвиђање и опишу колико купци вреднују особине некретнине, а колико особине окружења некретнине, док противречни случајеви дају потенцијална објашњења зашто је купац платио или није платио хедонизам. Експерименти су спроведени над подацима купопродајних трансакција у Московској регији који описују сваку трансакцију са 20 особина од којих је девет особина некретнине и 11 особина окружења некретнине. Експерименти показују да је могуће предвидети цене некретнине прилично успешно применом хедонистичких модела где је корен средње квадратне грешке логаритма цене некретнине једнак 0,2364. Потом, применом противречних случајева установљено је да постоје три појавна облика хедонизма на тржишту. То су приступ транспорту, здравственим и образовним институцијама, те близина додатним образовним и културним установама и на крају хедонизам квалитета животне средине.

КЉУЧНЕ РЕЧИ: Предвиђање цене некретнине, хедонистички модели, противречни примери

ABSTRACT: It is necessary to take into account several factors when predicting the price of real estate. More specifically, in addition to the accuracy of the prediction, it is necessary to explain the prediction and to provide a mechanism that enables what-if analysis. However, the explanation of the decision and what-if analysis requires a cause-and-effect analysis, which is very complex by nature. In this paper, the procedure of prediction, explanation of the decision, and cause-and-effect analysis are provided by a combination of hedonistic models and counterfactual examples. Hedonistic models serve to predict and describe how much the buyers value the properties of the real estate, and how much the properties of the real estate environment, while counterfactual examples give potential explanations why the buyer paid or did not pay for hedonism. Experiments are conducted on the purchase and sale transactions in the Moscow region, which describe each transaction with 20 properties, of which nine are properties of the property and 11 are properties of the real estate environment. Experiments show that it is possible to predict real estate prices quite successfully by applying hedonistic models with the root mean logarithm square error of the real estate price is equal to 0.2364. Then, by applying counterfactual examples, it is established that there are three manifestations of hedonism on the market. These are access to transport, health and educational institutions, and proximity to additional educational and cultural institutions, and finally the hedonism of environmental quality.

KEY WORDS: House price prediction, Hedonic price models, Counterfactual examples

1. УВОД

Предвиђање цене некретнине је задатак који се већ деценијама решава применом метода које долазе из домена статистичке анализе, попут анализе временских серија и регресионом анализом. На овај начин се уграђује динамика која је видљива у продајним ценама некретнина као што су тренд и сезоналност. Временом су класичне статистичке моделе заменили модели машинског учења који имају већу тачност у предвиђањима, али је корисност оваквих модела временом опала. Наиме, доносиоци одлука нису могли да искористе предвиђања због мањка разумевања самих модела и резултата [1].

У циљу додатне анализе и појачавања процеса доношења одлука потребно је да предиктивни модели имају додатне особине попут објашњења одлуке и разумевања узрочно-последичних веза. Додавањем ових особина ће све заинтересоване стране добити корист у виду разумевања догађања на тржишту некретнина. Доносиоци одлука могу да разумеју шта је то што је купцима битно у процесу куповине некретнине. Односно, које особине купаца вреднују приликом куповине некретнине и колика је значајност сваке особине. На пример, у [2] је приказано да је за изнајмљивање соба или станова преко *Airbnb*-а најбит-

није добро представити ткз. интринстичке особине собе или стана као што су квадратура просторија, број места за спавање и слично. Односно, ослањањем на коментаре корисника долази до већег незадовољства корисника, зато што су перципирани однос између домаћина и корисника динамичка категорија која се тешко материјализује пре чина резервисања или плаћања собе или стана. Међутим, познавање таквих информације се може искористити за просторно планирање целокупног насеља или општине где би доносиоци одлуке закључили на основу куповина шта је могуће и пожељно урадити како би грађани били задовољнији. Произвођач некретнина би могао, увидом у преференције купаца, да планира производњу објекта тако да купац буде задовољнији и самим тим издвоји више средстава за некретнину. На крају, купац некретнине је задовољнији уколико су његове потребе задовољене без да их је експлицитно исказао [7, 23]. Пример развоја јавних политика се може наћи у раду [3] у којем се оцењивао утицај приступа јавном превозу на цене некретнина за градове Рим и Сантандер. Након откривања битних особина јавног превоза дефинисан је план проширења јавног превоза и финансирања инфраструктуре јавног превоза са циљем обогаћивања вредности некретнина у одређеним деловима града и последично већег задовољства грађана.

Проблем у спровођењу оваквих анализа је био у проналажењу узрочно-последичних веза. Наиме, ни статистички модели, ни модели машинског учења нису у стању да тврде да је узрок цене некретнине нека особина, а ни да тврде која је јачина тог узрока. Каузални модели су превише компликовани и захтевају велики број тестирања како би се дошло до праве узрочно-последичне везе неке појаве. Чак и када се помисли да је модел направљен, може постојати нека особина која није опсервирана и да поремети тумачење [16]. Други проблем који постоји се налази у објашњавању одлуке и препоручивању одлуке. Наиме, доносилац одлуке не може да утиче на неке особине некретнине. Особине које описују саму некретнину су фиксне по природи и не могу да се мењају (нпр. спрат на којем се некретнина налази или површина некретнине). У том случају објашњење мора да узме у обзир све могуће особине, док препоручивање акција само особине које су променљиве и то тако да препоручивање акција измери ефекат како ће промена утицати на задовољство корисника при том имајући у виду узрочно-последичне везе са непроменљивим особинама [1].

У овом истраживању проблем предвиђања цене некретнине је решен комбинацијом хедонистичких модела и противречних примера. Хедонистички модели служе да опишу и предвиде цену некретнине дељењем особина некретнине на две групе. Прва група се односи на особине некретнине које не могу да се мењају и описују квалитет живљења, тзк. нехедонистичке особине. Друга група описује окружење некретнине које повећава вредност некретнине и које је корисник спреман да плати, тзк. хедонистичке особине [2]. Нехедонистичке особине су непроменљиве и оне описују некретнину преко објективних параметара као што је површина некретнине, спрат на којем се стан налази итд. Са друге стране хедонистичке особине попут удаљености од аутобуске станице могу да се промене и тиме повећа вредност некретнине. Да бисмо у потпуности искористили потенцијал хедонистичких модела потребно је да додатно узрочно-последичне везе. Противречни примери [22] представљају методологију проналаска потенцијалних узрочно-последичних веза које се налазе у моделима предвиђања попут хедонистичког модела цена некретнине. Противречни примери су вештачки примери који су најсличнији референтном примеру који имају другачији исход. Односно, како би изгледала некретнина да је корисник платио (или није платио) хедонизам. Ови примери групишу хедонистичке особине и дају потенцијалне разлоге хедонизма, односно објашњења шта је потребно променити како би корисник био задовољан. Експеримент је спроведен на подацима који описују купопродајне трансакције некретнина у Московској регији [19].

Остатак рада је описан кроз следећа поглавља. У другом поглављу се налази преглед литературе. У трећем поглављу се налази методологија истраживања. Четврто поглавље даје резултате експеримената и њихову дискусију. На крају, у петом поглављу се налази закључак и правци будућих истраживања.

2. ПРЕГЛЕД ЛИТЕРАТУРЕ

Поступак предвиђања цене некретнине је веома изазован задатак због великог броја фактора који утичу на цену некретнине. Ови фактори могу бити макро-економски, микро-економски, особине окружења и самог стамбеног простора. Неки од фактора се могу опсервирати и јасно утврдити, међутим неки од њих су нејасни, невидљиви и није их могуће прецизно измерити. Најчешће се цена некретнине предвиђа применом анализе временских серија, односно нелинеарне законитости у ценама некретнине се моделују кроз тренд, сезоналност и цикличност које су компоненте временских серија. Тачније, користе се методе временских серија као што су ауто-регресивни модели са померајућом средином (енг. *Auto-Regressive Moving Average*) и адаптацијама наведеног модела као што је ауто-регресивна интегрисана померајућа средина (енг. *Auto-Regressive Integrated Moving Average*) [11]. Међутим, ови модели имају проблема у предвиђању цена некретнине у присуству нелинеарних законитости или интеракција два или више фактора на цену некретнине. Због тога за предвиђања цене користе и алгоритми откривања законитости у подацима, попут Гаусових процеса (енг. *Gaussian process*), машине са бекторима подршке (енг. *Support Vector Regression*), а у последње време дубоким неуронским мрежама (енг. *Deep Neural Networks*) [18].

Према литератури постоје два главна приступа за моделовање цене некретнине. Први приступ се заснива искључиво на квантитативно мерљивим карактеристикама некретнине, нпр. површина објекта, број соба, спрат на којем се некретнина налази итд. Други и популарнији приступ је заснован на коришћењу динамике система економских фактора. Како се рад бави предвиђањем цене некретнине на тржишту Москве (Русија), преглед литературе ће се фокусирати на наведена два приступа и у региону руског тржишта.

Радови који се баве анализом цена некретнина на руском тржишту се углавном баве физичким карактеристикама некретнине заједно са економским факторима. Један битан економски фактор јесте цена нафте. Показало се да цена нафте на глобалном тржишту у великој мери утиче на куповну моћ, а тиме и цену некретнине у Русији [11, 17]. Тачније, јачина локалне валуте и њен однос са валутом Европске уније јасно описује цену квадратног метра некретнине на територији већих градова у Русији. Показало се да постоји веома јака корелација између цене нафте на глобалном тржишту и јачине локалне валуте, те се статистичким тестовима (анализа варијансе и Ман-Витни тест) независно за оба фактора показало да постоји статистички значајан утицај фактора на цену некретнине.

Међутим, на овај начин се не долази до одговора на питање колика је цена некретнине. Како би се дошло до тог одговора примењује се анализа временских серија и регресиона анализа. Модели временских серија покушавају да моделују зависности у ценама кроз време и разложе цену на тренд и сезоналност. Показало се да се оваквим моделима може предвидети цена на релативно задовољавајућем нивоу [6]. У раду [18] подељени су фактори који утичу на цену некретнине на макро-економске, микро-економске,

факторе тржишта некретнина и физичке особине некретнине. Неколико интересантних закључака је изведено. Наиме, веома висок утицај на цену некретнине има цена нафте на глобалном тржишту. Наиме, што је већа цена нафте већа је и цена некретнине и овај утицај је значајнији од физичких особина некретнине. Међутим, проблем који постоји код овог приступа јесте мултиколинearност. Односно, уколико су фактори које испитујемо корелисани, онда се не може веровати регресионим коефицијентима приликом тумачења резултата. Један приступ који може да се спроведе јесте анализа главних компоненти [19]. Међутим, тада се губи на јасноћи фактора.

Недавно, као модерни приступ у предвиђању цена (не само некретнина) издвојили су се хедонистички модели. Идеја хедонистичких модела јесте да се фактори поделе на две групе. Једну групу коју чине особине које су неопходне за живот (нпр. површина некретнине, број соба и површина терасе), док другу групу чине особине које представљају хедонистичко понашање купца, односно особине које нису везане за живљење (нпр. удаљеност од центра града, паркинг место или популарност насеља) [2]. На пример, у Токију [5] као негативни хедонистички фактори приликом куповине некретнине су се показали висина зграде и број станова у згради.

Додатни ниво комплексности који овај рад уноси јесте објашњење шта је потребно поправити како би купац платио „хедонизам“. Односно, наша идеја је да предложимо некретнину која је најсличнија постојећој са мањим изменама како би купац издвојио додатни новац за плаћање особина које нису нужно везане за живот. Метод који користимо је нов у области машинског учења и назива се прављење противречних случајева (енг. *counterfactual cases*) [22], те према нашим истраживањима још није употребљен у предвиђању цена некретнине. Термин противречних случајева не треба мешати са термином противречног размишљања. Противречно размишљање је процес описивања жаљења за жељеним исходом који се може статистички моделовати (у виду структурних једначина) у циљу објашњења потенцијалних разлога жаљења. Противречни случајеви покушавају да пронађу најмању могућу промену у скупу улазних атрибута тако да дође до жељене промене у циљном атрибуту. Дакле, уместо описивања исхода, противречни случајеви дају одговор шта је потребно променити како би се исход променио. Овај приступ може да служи пре свега доносиоцима одлука за генерисање политика и просторно планирање (нпр. где је потребно направити објекат културе, школско-васпитни објекат или медицински објекат), инвеститору да повећа вредност некретнине (кроз просторно планирање), али и самом кориснику некретнине како би повећао вредност некретнине.

3. МЕТОДОЛОГИЈА

Хедонистички модели се заснивају на посматрању зависности излазне варијабле од две групе карактеристика. Једна су оне директно везане за саму излазну варијаблу, а друга су оне које индиректно утичу на њу [2]. Цена некретнине је одређена карактеристикама саме некретнине као и спремношћу самог потрошача да плати карактеристике ве-

зане за окружење некретнине. У карактеристике саме некретнине би се убрајале особине попут површине некретнине, као и број соба, присуство балкона, година изградње и остале особине директно везане за саму некретнину. У карактеристике окружења би се убрајале удаљеност од важних објеката, попут школа, болница, аутобуске станице, стопа криминала у том крају, ниво загађења ваздуха и остале особине индиректно везане за саму некретнину [9].

Да бисмо добили комплетан одговор на питање да ли нека особина утиче на цену некретнине или не потребно је да се бавимо каузалном (узрочно-последичном) анализом. За ту потребу користимо противречне случајеве, тј. случајеве са измењеним вредностима атрибута у односу на неки референтни пример. Другим речима, шта би се догодило са ценом некретнине уколико се измене неки од датих атрибута. На тај начин добијамо додатни увид у каузалности атрибута који се могу користити при тумачењу самог модела као и прецизнијем доношењу одлука. Начин на који се постиже узрочно-последична веза код противречних случајева јесте учењем каузалног модела који подразумева да постоји скривени атрибут који утиче на све улазне атрибуте. Учењем новог модела који описује релацију скривеног атрибута на све улазне атрибуте учи се међусобни утицај између улазних атрибута и њихов заједнички утицај на циљни атрибут. Када је добијен такав модел може се искористити како би се испитало шта може да се промени како би се остварио жељени ефекат (нпр. промена циљног атрибута). То се назива противречно објашњење. Противречна објашњења (енг. *counterfactual explanations*) су заправо нова комбинација атрибута која узрокује одређено жељено предвиђање излазне варијабле. Таква инстанца, са одређеном комбинацијом вредности атрибута не мора нужно постојати, већ може бити замишљена инстанца [16]. Поступак учења узрочно-последичних односа модела машинског учења који покушава да научи функцију којом се предвиђа вредност жељеног исхода добијеним хедонистичким моделом са најмањим могућим променама у вредностима улазних атрибута (под променама се подразумева и да се што мање особина промени и да се вредности што мање промене). Реализација каузалних модела се постиже најчешће дубоким неуронским мрежама и заинтересовани читаоци се упућују на [4, 21, 22].

Циљ овог рада јесте да да одговор на питање зависности цене некретнине од карактеристика саме некретнине и доприноса особина околине у формирању цене, као и дефинисању узрочно-последичне везе промена вредности особина некретнине и промена вредности цене. Познавањем тих одговора све заинтересоване стране могу имати добробити. Наиме, доносиоци одлука на нивоу државе могу да управљају просторним развојем на нивоу насеља или општине планирајући развој здравствених, образовних, културних установа или објеката за разоноду. Произвођачи добијају могућност испитивања фактора који ће утицати на цену објекта, те могу планирати објекте који ће највише допринети „плаћању хедонизма“. Испуњавањем ових особина, крајњи корисник ће осећати веће задовољство јер се у околини некретнине налазе управо они објекти који су кориснику потребни.

4. ПОСТАВКА ЕКСПЕРИМЕНТА

4.1. Опис података

Подаци над којима је спроведен експеримент долазе из подручја Московског региона у периоду од 2011. године до 2015. године и податке је обезбедила Сбербанка [19]. Скуп података се састоји укупно 30.471 трансакција описаних преко 20 особина које описују физичке карактеристике некретнине, али и особине које долазе из окружења некретнине. Ове особине чине улазне променљиве које служе за предвиђање циљног атрибута који представља цену некретнине. Подела карактеристика на нехедонистичке и хедонистичке представљена је у табели 1.

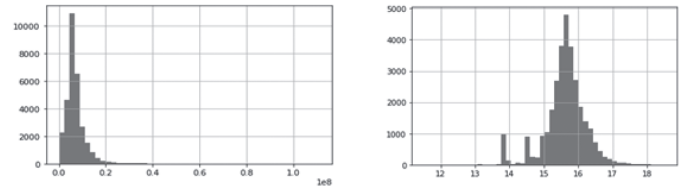
Табела 1. Нехедонистичке и хедонистичке карактеристике некретнине

Нехедонистичке карактеристике	Хедонистичке карактеристике
<i>full_sq</i>	<i>railroad_km</i>
<i>life_sq</i>	<i>ttk_km</i>
<i>apartment_size</i>	<i>average_kshs_distance</i>
<i>num_room</i>	<i>additional_education_km</i>
<i>build_year</i>	<i>theater_km</i>
<i>kitch_sq</i>	<i>radiation_km</i>
<i>floor</i>	<i>cpi</i>
<i>max_floor</i>	<i>cafe_count</i>
<i>state</i>	<i>basketball_km</i>
	<i>0_13_all</i>
	<i>0_13_female</i>

Нехедонистичке особине представљају укупни површину некретнине (*full_sq*), смабену површину некретнине (*life_sq*), величина дневног боравка (*apartment_size*), број просторија (*num_room*), годину изградње (*build_year*), површину кухиње (*kitch_sq*), спрат на којем се налази некретнина (*floor*), највећи спрат у којем се некретнина налази (*max_floor*) и категорија која представља стање некретнине (*state*). Особине које представљају хедонизам се грубо могу поделити у неколико категорија. Прву категорију чине транспортне особине некретнине. То су удаљеност до железничке станице (*railroad_km*) и удаљеност од трећег транспортног прстена (*ttk_km*). Затим, хедонизам је описан и кроз удаљеност од здравствених и образовних установа (*average_kshs_distance*), као и кроз удаљеност од додатних извора образовања као што су школе за децу са посебним потребана, за образовање одраслих, за надарену децу и слично (*additional_education_km*) и близина културних установа (*theater_km*). Квалитет живота је описан кроз удаљеност од радиоактивног отпада (*radiation_km*) и раст цена индекса потрошачких цена (*cpi*). На крају, насеље у којем се налази некретнина је описана кроз број кафића у окружењу од 1500 метара (*cafe_count*), те удаљеност од терена за кошарку (*basketball_km*), број деце у насељу (*0_13_all*) и број женске деце у насељу (*0_13_female*). Наведене особине су изабране након иницијалне анализе где се показало да наведене особине појединачно имају јачи

утицај на циљни атрибут.

Расподела цена некретнина посматрањем скупа података не показује нормалну расподелу (десни део је издужен). Стога је цена непокретности логаритмована, чиме се постиже нормална расподела цена и омогућава успешније предвиђање цена на тестном скупу података.



Слика 1. Расподела вредности цена у а) оригиналном и б) логаритамском облику

Што се припреме података тиче недостајуће вредности су попуњене средњим вредностима за ту особину.

4.23. Поставка истраживања

Као што је већ споменуто, поступак предвиђања цене некретнине је спроведен коришћењем хедонистичких модела. Ови модели, поред самог предвиђања, дају утицај особина које описују живљење (нехедонистичке особине), као и утицај особина које нису битне за живљење али је купац спреман да их плати (хедонистичке особине). Потом, у циљу одређивања узрочно-последичних веза спроведена је методологија противречних случајева која даје додатно сазнање о томе шта је могуће променити како би се некретнина добила на вредности.

За учење хедонистичког модела користи се најчешће алгоритам линеарне регресије [8] што је случај и у овом истраживању. Овај алгоритам представља линеарни модел у којем свака особина има придружен коефицијент који описује утицај особине на цену некретнине. Уколико је коефицијент позитиван, онда та особина позитивно утиче на цену некретнине. Односно са повећањем вредности особине расте и цена некретнине (нпр. са већом површином некретнине већа је и цена). Супротно томе, уколико је коефицијент особине негативан, онда са порастом те особине опада цена некретнине.

Противречни случајеви су имплементирани кроз учење дубоке неуронске мреже коришћењем *DiCE* библиотека [15] у програмском језику *Python*. Противречни случајеви омогућавају додатну анализу и у наставку рада су приказани примери противречних случајева који дају препоруке шта је потребно променити у окружењу некретнине како би купац платио хедонизам. Односно, које хедонистичке особине би требало променити како бу купац осетио задовољство и платио хедонизам. Приликом прављења противречних примера битно је дефинисати значајност да у објашњењима вредности атрибута буду што сличнија референтној инстанци, као и значајност разноликости противречних случајева (да се добијени противречни случајеви међусобно што више разликују).

Приликом учења хедонистичког модела користи се процес унакрсне валидације. Односно, скуп података се дели на десет подједнаких делова од којих се девет користи за учење модела, а преостали за проверу тачности модела. Поступак се понавља десет пута тако да сваки подкуп података тачно једном буде употребљен за проверу тачности. Унутар унакрсне валидације спроводи се додатна валидација за учење противречних случајева тако што се 80% користи за учење модела, а преосталих 20% за проверу. Као мера тачности предвиђања коришћен је корен средње квадратне грешке (*RMSE*) и средња апсолутна грешка (*MAE*) [20].

У циљу детаљније провере и анализе хедонистичких модела и противречних случајева спроведен је експеримент кроз четири поставке. Ове поставке су постављене како би се проучио допринос особина у формирању цене. Поставке су:

1. Све изабране особине се користе као улазни атрибут, а цену као излазни атрибут;
2. Нехедонистичке особине се користе као улазни атрибут, а цена као излазни атрибут;
3. Хедонистичке особине се користе као улазни атрибут, а цена као излазни атрибут; и
4. Након учења модела представљеним у поставци 2, користе се хедонистичке особине као улазни атрибут како би се научили противречни примери.

Први експеримент представља класичан пример учења хедонистичких модела. Наиме, користе се све могуће особине и проверава се утицај појединачних особина на цену некретнине. Зарад поређења коефицијената линеарне регресије, спроведени су и експерименти предвиђања цене некретнине у којем се користе искључиво нехедонистичке особине и искључиво хедонистичке особине. На крају, спроведен је и поступак генерисања противречних случајева. За генерисање противречних случајева потребно је имати класификациони проблем (два могућа исхода). У овом истраживању, дефинисан је исход хедонизам (или одсуство хедонизма) на основу грешке предвиђања када се користе само особине које описују живљење (други експеримент). Наиме, ако је разлика стварне цене некретнине и предвиђања позитивна (модел линеарне регресије је потценио стварну цену некретнине) онда је корисник платио хедонизам. Интуиција иза ове дефиниције јесте да модел није имао на располагању особине хедонизма и да те особине описују разлику стварне и предвиђене вредности. У супротном случају, ако је разлика стварне цене некретнине и предвиђања негативна или једнака нули онда корисник није платио хедонизам.

4. РЕЗУЛТАТИ И ДИСКУСИЈА

Резултати истраживања и дискусија модела је описана у два засебна дела. Наиме, прво су описани резултати добијени приликом учења хедонистичких модела. Овај поступак описује прве три поставке експеримената описане у претходном поглављу. Након приказа и дискусије хедо-

нистичких модела, приказани су противречни случајеви и спроведена је њихова дискусија.

4.1. Хедонистички модел

Део истраживања примене хедонистичког модела описује прве три поставке експеримената описане у претходном поглављу. Коефицијенти линеарне регресије су приказани у табели 2.

Приликом тумачења коефицијената линеарне регресије треба имати у виду да је циљни атрибут (цена некретнине) изражена на логаритамској скали, те се коефицијенти тумаче као промена логаритма цене некретнине за јединични раст особине.

У првој поставци експеримента (сви атрибути истовремено), највећи утицај на цену имају стање некретнине (коефицијент линеарне регресије износи 3,48), удаљеност од радиоактивног отпада (2,85), број соба (2,38) и укупна квадратура (1,44), али и особине удаљеност од здравствених и образовних установа (-2,98), као и удаљеност од трећег транспортног прстена (-2,67). Тумачење ових коефицијената описује да се цена некретнине повећава уколико је стање некретнине боље, ако је некретнина већа, са више засебних просторија, удаљенија од радиоактивног отпада, али ближа здравственим и образовним институцијама и ближа извору транспорта. Односно, можемо приметити да је део битних особина из категорије хедонизма, попут удаљености од радиоактивног отпада, удаљености од здравствених и образовних установа и удаљености од трећег транспортног прстена. Ово је сигнал да је становницима Московског региона битно да живе у здравом окружењу, што је слично као и становницима Токија [5], Белфаста [14] и Нанта [12]. Такође, грађани желе да имају брз приступ транспортном систему. Односно, желе да имају могућност да се брзо и једноставно пребаце са места становања до ресурса који им је потребан. Ова особина се показала као веома битна, посебно у великим градовима као што је Токио [5], и Сиднеј [13], али и у Великој Британији [10].

Уколико се посматра само особине које се везују директно за некретнину, онда су резултати такви да се поред директне пропорционалности која је већ показан као значајна за укупну површину (1,47), посебно издваја стање некретнине (5,50) и број соба (5,97) као далеко најзначајније при формирању цене. То говори да је цена за такве некретнине врло стабилна на промене осталих карактеристика уколико је већи број соба и доброг, тј високо оцењеног стања некретнине. Међутим, значај свакако има укупан број спратова зграде у којој се некретнина налази (1,08), где се повећањем броја спратова повећава и цена некретнине.

Уколико се посматра цена некретнине само кроз призму зависности од особина околине у којој се та некретнина налази (хедонистичке особине), онда се може видети да је најзначајније да некретнина буде што удаљенија од радиоактивног отпада (3,31) али и од железнице (2,13). Али је важно да некретнина буде близу здравствених и образовних установа (-2,78) и трећем транспортном прстену (-2,54).

Табела 2. Коэффициенти линеарне регресије цене и атрибута по поставкама

Карактеристика	Поставка 1	Поставка 2	Поставка 3
<i>full_sq</i>	1,44	1,47	/
<i>life_sq</i>	-0,08	-0,27	/
<i>apartment_size</i>	-0,73	0,21	/
<i>num_room</i>	2,38	5,97	/
<i>build_year</i>	-0,11	-0,58	/
<i>kitch_sq</i>	0,24	-0,11	/
<i>floor</i>	0,47	0,61	/
<i>max_floor</i>	0,28	1,08	/
<i>state</i>	3,48	5,50	/
<i>railroad_km</i>	0,71	/	2,13
<i>tkm_km</i>	-2,67	/	-2,78
<i>average_kshs_distance</i>	-2,98	/	-2,54
<i>additional_education_km</i>	0,61	/	-0,05
<i>theater_km</i>	0,30	/	0,37
<i>radiation_km</i>	2,85	/	3,31
<i>cpi</i>	0,21	/	0,28
<i>cafe_count</i>	0,01	/	0,01
<i>basketball_km</i>	-0,55	/	-1,0
<i>0_13_all</i>	0,01	/	0,01
<i>0_13_female</i>	0,01	/	0,01

Имајући у виду добијене резултате, можемо закључити да је цена некретности у Московској регији зависи и од особина живљења и од хедонистичких особина. Стање у којем се некретнина налази, величина некретности и број соба у великој мери утичу на цену, али утичу и особине које омогућавају здрав живот и брз приступ различитим друштвеним добрима.

Потребно је још одговорити на питање колико добро наведени модели предвиђају. У табели 3 су приказане вредности корена средње квадратне грешке за прве три поставке модела, одакле се може видети тачност предвиђања.

Табела 3. Тачност предвиђања хедонистичких модела

Модел линеарне регресије	Поставка 1	Поставка 2	Поставка 3
<i>RMSE</i>	0,2364	0,2597	0,3291
<i>MAE</i>	0,3126	0,3367	0,3389

Може се уочити да је најбољи модел линеарне регресије онај у којем се користе све особине, и хедонистичке и нехедонистичке. У том случају грешка измерена преко корена средње квадратне грешке износи 0,2364, односно средња апсолутна грешка износи 0,3126. Ова вредност корена средње квадратне грешке сугерише да модел линеарне регресије прилично успешно предвиђа стварне вредности цена некретности. Слични резултати се постижу и за модел који користи особине некретности (нехедонистичке особине), док се слабије предиктивне перформансе (али задовољавајуће добре) постижу када се користе хедонистичке особине. Другим речима, успешност предвиђања више зависи од особина некретности, него од хедонистичких особина.

4.2. Противречни случајеви

Спровођење узрочно-последичне анализе је спроведено коришћењем противречних случајева. Противречни случајеви узимају у обзир међусобну интеракцију коју улазни атрибути имају и прави онакве примере тако да се постигне жељени исход, са најмањом могућом разликом у односу на посматрани случај (у овом истраживању продају некретности). У овом истраживању, дефинисан је исход хедонизам (или одсуство хедонизма) на основу грешке предвиђања када се користе само особине које описују живљење. Потом је испитано како се може променити окружење некретности тако да се добије другачији исход (нпр. да корисник издвоји додатна средства како би платио хедонизам). Тумачењем противречних случајева добија се узрочно-последично објашњење. Како објашњења зашто је корисник платио (или није платио) хедонизам има много, генерише се више случајева и сваки се појединачно тумачи.

У наставку рада приказана су два примера противречних случајева. Један у којем корисник није платио хедонизам и други у којем је платио хедонизам. Најпре је показан пример у којем корисник није платио хедонизам и где су генерисани противречни случајеви у којем би такав корисник платио хедонизам уколико би се промениле вредности особина хедонизма (Табела 4). Како се у свакој куповини може приметити значајнија промена одређених особина околине, такве промене су означене подебљаним словима.

Табела 4. Вредности атрибута некретности за упитну инстанцу и генерисана противречна објашњења (плаћање хедонизма)

Карактеристика	Упитна некретнина	1	2	3	4	5
<i>railroad_km</i>	12,7	6,283	0,604	12,834	0,002	15,315
<i>tkm_km</i>	14,2	1,874	24,701	14,134	0,002	14,235
<i>average_kshs_distance</i>	17	0,203	17,057	0,293	15,837	8,808
<i>additional_education_km</i>	14,2	24,268	10,736	24,224	15,359	14,205
<i>theater_km</i>	18,5	18,413	10,588	78,028	18,587	18,413
<i>radiation_km</i>	0,5	20,033	0,238	0,222	0,252	53,890
<i>cpi</i>	437	444,2	437	437	424,4	441,8
<i>cafe_count</i>	25	381	1011	496	24	24
<i>basketball_km</i>	7	12,025	0,435	6,608	2,104	14,403
<i>0_13_all</i>	10000	6128	9965	6048	9965	15433
<i>0_13_female</i>	5000	5017	7433	4071	4983	6192
Хедонизам	није платио	платио	платио	платио	платио	платио

Један од разлога (први противречни пример) зашто наведени корисник није платио хедонизам је транспорт, животна средина и забава. Наиме, ове три особине заједно су допринеле да корисник не плати хедонизам. Најсличнији пример у којем би корисник платио хедонизам захтева да присуство транспортној мрежи буде знатно мање (особине *railroad_km* и *tkm_km*). Такође, потребно је да се потенцијално размотри о изградњи здравствених и образовних установа јер удаљеност мора да се смањи са 17 км на 200 метара, као и да се извор радиоактивног отпада удаљи са 500 метара на 20 км. На крају, потребно је постоји знатно већи број кафића у окружењу. Други потенцијални разлог (други и четврти противречни случај) може бити спорт-

ске природе. Односно, околина некретнине би могла да се претвори у спортски квартал. Наиме, кошаркашка хала би требало да буде много ближа (са 7 км на 400 м) и да је железничка станица ближа (са скоро 13 км на 600 м), као и да постоји значајно већи број кафића у околини (уместо 25 нешто преко 1000). Потенцијални трећи разлог (трећи и пети противречни пример) би могао бити присуство здравствених и образовних институција. Другим речима, насеље би могло да се претвори у нпр. универзитетски кампус. У том случају потребно је да једна таква институција буде на 300 метара и да се повећа број кафића у окружењу или да се извор радиоактивног отпада склони на знатно удаљеније место (са 500 метара на 54 км). Могу се издвојити и закључити из противречних примера да у Московској регији хедонизам представља близина приступа транспорту (градском превозу или железници), образовне институције, индустрија забаве (спортске дворане и кафићи) и животна средина (удаљеност од радиоактивног отпада). Ове особине се јављају самостално или у комбинацијама.

Затим су приказани примери где је корисник платио хедонизам и где су направљени случајеви где се испитује зашто је корисник платио хедонизам. Добијени противречни случајеви су представљена у табели 5. Како се у свакој куповини може приметити значајнија промена одређених атрибута, такве промене су подебљане.

Табела 5. Вредности атрибута некретнине за упитну инстанцу и генерисана противречна објашњења (неплаћање хедонизма)

Карактеристика	Упитна инстанца	1	2	3	4	5
railroad_km	2,7	2,391	10,338	7,037	2,182	17,071
ttk_km	4,2	64,135	2,509	0,045	1,055	44,928
average_kshs_distance	13	55,52	13,110	0,203	0,360	38,337
additional_education_km	14,2	14,176	22,600	1,062	14,607	14,215
theater_km	8,5	6,193	47,331	7,506	0,083	8,413
radiation_km	24,5	22,315	24,797	0,009	0,634	23,223
spi	437	437	443,8	418,2	437	437
cafe_count	25	1623	2	8	865	20
basketball_km	1	0,052	0,943	0,052	0,652	18,231
0_13_all	10000	10736	16883	9965	322	10033
0_13_female	5000	5017	4983	657	4997	5017
Хедонизам	платио	није платио	није платио	није платио	није платио	није платио

Узрочно-последични разлози зашто је корисник платио хедонизам могу бити различити. Један од разлога је близина трећем транспортном прстену, као и близина здравственим или образовним установама (први противречни случај). Потенцијално објашњење додаје и близину кошаркашкој дворани и приступ железници (пети противречни случај). У оба случаја, битно објашњење је близина приступа транспорту и близина здравственим и образовним институцијама. Друго потенцијално објашњење је близина образовним и културним установама (друго противречно објашњење) у комбинацији са приступом железничкој ста-

ници и разонодом (близина позоришта и број кафића). На крају, разлог зашто је корисник платио хедонизам може да лежи у животној средини. Наиме, некретнина је доста удаљена од извора радиоактивног отпада (трећи и четврти противречни случај). Дакле, посматрањем проблем узрока плаћања хедонизма са друге стране (зашто је неко платио хедонизам, уместо посматрања шта је потребно да би платио хедонизам) добијају се слични али јаснији закључци. Хедонизам се може описати у три облика. Први облик јесте приступ транспорту, здравственим и образовним институцијама. Други облик хедонизма представља близина додатним и допунским образовним и културним установама. На крају, хедонизам представља и животна средина (односно, удаљеност од извора радиоактивног отпада).

Примена комбинације хедонистичких модела и противречних случајева може донети корист свим заинтересованим странама у процесу предвиђања цене некретнине. Хедонизам који је неопходно платити јасно се може објаснити купцу, али истовремено предочити зашто је цена некретнине већа уколико се постави такво питање при куповини. Са друге стране и сами инвеститори могу на основу ових података очекивати веће продајне цене некретнине у које улажу уколико је инфраструктура краја у коме се некретнина налази боља, односно иде у складу са уоченим предностима удаљености од значајних објеката. Самим тим и њихова зарада ће бити већа. На крају, доносиоци одлуке у области просторног планирања би могли да добију информацију како да развијају град тако да сви грађани добију ресурс који им је потребан и који желе да плате.

ЗАКЉУЧАК

У овом раду спроведена је прављење модела за предвиђање цена некретнине који има додатне пожељне особине објашњења предикције, као и могућност спровођења шта-ако анализе са фокусом на узрочно-последичне везе. Тачније, научени су хедонистички модели цена некретнине у комбинацији са противречним примерима и противречним објашњењем.

Експеримент је спроведен над подацима Сбербанке који описују трансакције некретнине у Московској регији преко 20 особина од којих девет особина описује некретнину, а 11 особина окружење некретнине, односно описује хедонизам купца. Хедонистички модели је показао да е могуће прилично успешно предвидети цену некретнине где је корен средње квадратне грешке логаритма цене некретнине једнак 0,2364. Као најбитније нехедонистичке особине показале су се стање некретнине, удаљеност од радиоактивног отпада, број соба, укупна површина некретнине. Док су битне хедонистичке особине удаљеност од здравствених и образовних установа и удаљеност од трећег транспортног прстена. Противречни случајеви дају потенцијална објашњења зашто је корисник платио или није платио хедонизам који може да се користи за процес доношења одлука на нивоу државе или на нивоу произвођача некретнине. Примећено је да се хедонизам може ја-

вити у три облика. Први облик хедонизма описује приступ транспорту, здравственим и образовним институцијама. Други облик хедонизма представља близину додатним и допунским образовним и културним установама. На крају, хедонизам представља и квалитет животне средине коју купци вреднују.

Као правци даљег истраживања планирамо да унапредимо поступак учења хедонистичких модела кроз временску анализу регресионих коефицијената. Наиме, уместо посматрања целог скупа података жеља је да се подаци поделе по годинама и испита да ли су законитости које смо идентификовали важеће ако се скуп података подели на мање јединице, као и да ли ће законитости важити у будућности. Поред наведеног, планирамо да поступак поновимо на различитим географским јединицама, између осталих и на територији Београда.

ЗАХВАЛНИЦА

Овај рад је делимично финансиран са пројекта Канцеларије за поморска истраживања, САД (*ONR/ONR Global* број пројекта N62909-19-1-2008).

РЕФЕРЕНЦЕ

- [1] Biran, O., & Cotton, C. (2017, August). Explanation and justification in machine learning: A survey. In *IJCAI-17 Workshop on Explainable AI (XAI)* (Vol. 8, No. 1, pp. 8-13).
- [2] Chen, Y., & Xie, K. (2017). Consumer valuation of Airbnb listings: a hedonic pricing approach. *International Journal of Contemporary Hospitality Management*, 29(5), 2405-2424.
- [3] Cordera, R., Coppola, P., dell'Olio, L., & Ibeas, A. (2019). The impact of accessibility by public transport on real estate values: A comparison between the cities of Rome and Santander. *Transportation Research Part A: Policy and Practice*, 125, 308-319.
- [4] Dandl, S., Molnar, C., Binder, M., & Bischl, B. (2020). Multi-Objective Counterfactual Explanations. *arXiv preprint arXiv:2004.11165*.
- [5] Diewert, W. E., & Shimizu, C. (2016). Hedonic regression models for Tokyo condominium sales. *Regional Science and Urban Economics*, 60, 300-315.
- [6] Gluschenko, K. (2017). The Moscow Market in Country's Economic Space. *MPRA Paper*, (80901).
- [7] Gosal, A. S., & Ziv, G. (2020). Landscape aesthetics: Spatial modelling and mapping using social media images and machine learning. *Ecological Indicators*, 117, 106638.
- [8] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics. *Springer New York*.
- [9] Heyman, A. V., Law, S., & Berghauer Pont, M. (2019). How is location measured in housing valuation? A systematic review of accessibility specifications in hedonic price models. *Urban Science*, 3(1), 3, 14 pages.
- [10] Higgins, D. M., Rezaei, A., & Wood, P. (2019). The value of a tram station on local house prices: an hedonic modelling approach. *Pacific Rim Property Research Journal*, 25(3), 217-227.
- [11] Kisel, T. N., & Palastrova, Y. I. (2016). Analysis of the main trends in prices in the real estate market of Moscow in the period of economic instability. *Ekonomika Predprijatija*, (2), 847-851.
- [12] Le Boennec, R., & Salladarré, F. (2017). The impact of air pollution and noise on the real estate market. The case of the 2013 European Green Capital: Nantes, France. *Ecological Economics*, 138, 82-89.
- [13] Lieske, S. N., van den Nouwelant, R., Han, J. H., & Pettit, C. (2019). A novel hedonic price modelling approach for estimating the impact of transportation infrastructure on property prices. *Urban Studies*, 0042098019879382.
- [14] McCord, M. J., MacIntyre, S., Bidanset, P., Lo, D., & Davis, P. (2018). Examining the spatial relationship between environmental health factors and house prices. *Journal of European Real Estate Research*, 11(3), 353-398.
- [15] Mothilal, R. K., Sharma, A., & Tan, C. (2020, January). Explaining machine learning classifiers through diverse counterfactual explanations. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 607-617).
- [16] Pearl, J. (2019). The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, 62(3), 54-60.
- [17] Safullin, M. R., El'shin, L. A., & Shakirova, A. I. (2012). Evaluation of business and economic activity as a short-term forecasting tool. *Herald of the Russian Academy of Sciences*, 82(4), 290-294.
- [18] Salmikov, V. A., & Mikheeva, O. M. (2018). Models for predicting prices in the Moscow residential real estate market. *Studies on Russian Economic Development*, 29(1), 94-101.
- [19] Smirnov, S., & Tlostanov, V. (2019). Moscow Real Estate: Pricing Analysis Through the Prism of Statistics and Machine Learning. Available at SSRN 3374436.
- [20] Suknović, M., & Delibašić, B. (2010). Poslovna inteligencija i sistemi za podršku odlučivanju. *Fakultet organizacionih nauka, Beograd*.
- [21] Verma, S., Dickerson, J., & Hines, K. (2020). Counterfactual Explanations for Machine Learning: A Review. *arXiv preprint arXiv:2010.10596*.
- [22] Wachter, S., Mittelstadt, B., & Russell, C. (2017). Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harvard Journal of Law & Technology*, 31, 841.
- [23] Zhao, X., Ma, X., Tang, W., & Liu, D. (2019). An adaptive agent-based optimization model for spatial planning: A case study of Anyue County, China. *Sustainable Cities and Society*, 51, 101733.



Ђорђе Радивојевић, студент,

Универзитет у Београду – Факултет организационих наука

Контакт: djordje.r.it@gmail.com

Област интересовања: Машинско учење, Базе података, Веб апликације, Клауд технологије



Сандро Радовановић, асистент,

Универзитет у Београду – Факултет организационих наука

Контакт: sandro.radovanovic@fon.bg.ac.rs

Област интересовања: Машинско учење, Откривање законитости у подацима, Системи за подршку одлучивању, Теорија одлучивања



др Борис Делибасић, редовни професор,

Универзитет у Београду – Факултет организационих наука

Контакт: boris.delibasic@fon.bg.ac.rs

Област интересовања: Системи за подршку одлучивању, Вишекритеријумско одлучивање, Пословна интелигенција, Машинско учење