

ПРИЛОЖЕНИЕ НА ИЗКУСТВЕНИЯ ИНТЕЛЕКТ ЗА СЪБИРАНЕ И ОБРАБОТКА НА ДАННИ ЗА ПРИЛАГАНЕ В МНОЖЕСТВЕНИ РЕГРЕСИОННИ МОДЕЛИ ПРИ ОЦЕНКА НА НЕДВИЖИМИ ИМОТИ

Яна Стоенчева, Борислав Борисов

APPLICATION OF AI IN DATA COLLECTION AND PROCESSING FOR USE IN MULTIPLE REGRESSION MODELS FOR PROPERTY VALUATION

Yana Stoencheva, Borislav Borisov

Abstract: This article examines the potential of AI as instrument for data collection and processing, and its use in property valuation. Some Bulgarian and international publications on the topic are reviewed. Data from over 1,000 housing sale offers in 11 metropolitan districts have been collected and processed using AI, specifically the ChatGPT platform. A hedonic regression analysis was conducted on a cleaned sample of offers. The study draws conclusions on the role of AI in collecting and processing large-scale databases, as well as on the application of the hedonic pricing model in real estate valuation using the comparative approach. Recommendations for further elaboration of the model and advancing the research have been proposed.

Key words: artificial intelligence; hedonic price; multiple regression; property valuation; sales comparison approach; adjustment factor.

Увод

Най-разпространеният и приложим подход, използван за оценка на недвижими имоти е сравнителният. Неговата философия изхожда от хипотезата, че стойността на всеки актив следва да е близка до стойностите на други обекти, които имат сходни качества по отношение на ключови характеристики, участващи в образуването на пазарната им цена. За сравнение се използва информация за реално сключени сделки, или когато няма такива - оферти за продажба на обекти, подобни на оценявания.

Недвижимите имоти се характеризират със своята хетерогенност. Не съществуват два имота, които да са напълно еднакви. Дори апартаменти, които се намират на един и същи етаж в една и съща сграда, с еднаква площ и довършителни работи, имат различно изложение и разкриват различна гледка. Поради тази причина, не е възможно, стойностите на аналозите ползвани за оценка на даден имот по сравнителния подход, да бъдат приложени директно към обекта на оценка. Необходимо е, чрез извършване на корекции по избрани характеристики, качествата на аналозите да бъдат приравнявани към обекта на оценка. Това се осъществява чрез прилагането на определени от оценителя, на база неговата експертиза и практически опит, корекционни коефициенти. Множество научни разработки търсят решение на проблема със

силното влияние на субективния фактор при оценяването. Такава възможност за дефиниране на емпирично установени зависимости, които да се използват при определянето на корекционни коефициенти, дава използването на Хедоничната регресия.

Моделът на Хедоничната стойност е вид множествен регресионен анализ, чрез който цената на оценявания обект се определя, като функция от цените на различните характеристики на обекта. Това е сложен статистически метод, за чието прилагане е необходимо да се използва голяма по обем, представителна извадка от пазарни свидетелства съотносими към обекта на оценяване. Развитието на технологиите и все по-бързо усъвършенстващите се платформи за обработка на големи обеми от данни, станали популярни с названието „изкуствен интелект“ (*Artificial Intelligence, AI*), дава възможност да бъде чувствително намалено времето необходимо за анализ на големи информационни масиви, което прави възможно изготвянето на бази данни с достатъчна по обем информация, подходящи за прилагане в иконометрични модели.

Темата е актуална поради голямото значение на оценяването за пазара на недвижими имоти и все по-широко дискутираната тема за навлизането на *AI* във всички сфери на живота. На база изготвени оценки се вземат важни инвестиционни решения, определят се размера на отпусканите ипотечни кредити (като допустим про-

цент от пазарната оценка), максималният размер на непаричните (апортни) вноски в капитала на търговски дружества, цените на сделки с имоти общинска и държавна собственост, стойността на дълготрайните материални активи (земя и сгради) за нуждите на счетоводната отчетност и данъчното облагане и т.н. Всичко това налага да бъде изисквана максимална обективност при изготвяне на оценките. Разработването на надеждна методология за определяне на корекционните коефициенти, които се използват при прилагане на сравнителния подход за оценка, би могло да намали в голяма степен субективния фактор при изготвяне на оценките и да увеличи тяхната надеждност.

Целта на изследването е да се покаже как, чрез използването на технологията на *AI* е възможно да бъдат извлечени и обработени голям обем данни, които след това да бъдат приложени в модел от типа Хедонична регресия, за обясняване на зависимостта на стойността на обекта от стойностите на неговите характеристики.

Така получените зависимости могат да бъдат основа за определяне на коефициенти за корекция на стойността на пазарните свидетелства.

Поставените изследователски задачи са:

- Проучване на съществуващия опит;
- Събиране на информация;
- Обработка на данните с инструментите на *AI* и конструиране на модел за статистически анализ на събраните данни;
- Прилагане на модела с помощта на специализиран статистически софтуер и анализ на получените резултати.

Обект на изследване в разработката е пазарът на недвижими имоти в град София.

Предмет на изследване са цените на жилищни имоти (апартаменти) в многофамилни сгради.

Основната изследователска теза е, че специализираните форми на изкуствен интелект могат да бъдат незаменим помощник за определяне на реалистични пазарни цени на недвижими имоти чрез извършване на надежден сравнителен анализ по определени предварително ключови характеристики.

Ограничителни условия:

Обхванати са единадесет квартала в град София, избрани така че да имат ясно разграничими локационни характеристики по отношение на инфраструктурата, отдалечеността от центъра, транспортната достъпност и не на последно

място категорията „престиж“.

Данните са статични и са събрани в определен момент.

Има ограничение и по отношение на видовете апартаменти, като са разгледани само оферти за продажба на апартаменти с една, две и три стаи.

Преглед на литературата

Оценяването на недвижимите имоти е творчески процес, който се базира на задълбочените познания и опита на оценителите. Най-често прилагания подход при изготвянето на пазарни оценки е сравнителният. Неговата методология се състои в определяне на пазарната стойност на основата на сравнение с цените на съпоставими с обекта на оценката недвижими имоти (аналози), с отчитане на влиянието на подбрани фактори, участващи в образуването на цената. Спрямо констатираните различия между характеристиките на аналозите и обекта на оценка се правят корекции на цените на пазарните свидетелства от гледна точка на оценявания обект, като ако неговите качества са по-високи от тези на аналога корекцията е в положителна посока и обратно (Пелов, 2011). Счита се за приемливо, коригираната стойност на обекта на оценка да е в границите от 0,8 до 1,5 пъти от средната цена на пазарните аналози за района, в който се намира обекта (Калинков, 2007). Както беше споменато по-горе определянето на тези корекционни коефициенти е чисто субективно и зависи от нивото на експертиза и професионалния опит на оценителите.

При избора на характеристики, по които се извършва сравнителният анализ, следва да бъдат включени тези, които са най-съществени за формиране на стойността на обектите и за които е достъпна надеждна информация в натурални измерителни. Местоположението е един от най-значимите фактори, както за избор на населено място за обитаване, така и за избор на конкретна локация в самото населено място (Стоенчева, 2020). Стойността на жилище в „добър“ квартал не идва от местоположението само по себе си, а от добавените, благодарение на тази локация характеристики: изложение/гледка, удобство на достъпа до имота, качество на сградата, качество на заобикалящата го среда, демографски и социални характеристики на района, достъп до комунални услуги и много други характеристики, дефинирани от местоположението на жилището. Така всички тези фактори се явяват цено-

образуващи по отношение на имотите, като всеки имот представлява уникална съвкупност от атрибути. (Stoenchev, 2018)

Бумът в развитието на промишленото производство и новите технологии след първата световна война, води до това че в кратки периоди от време определени продукти напълно отпадат от производство, а новите им заместители имат коренно различни качествени характеристики. Използваните тогава методи за изчисляване на ценови индекси, като метода на припокриващите се времеви редове и средните стойности по обобщени категории, стават все по трудно приложими. Необходимо е да бъдат използвани нови методи за изчисляване на ценови индекси. Така възниква техниката за установяване на цената на продукта, като производна от стойностите на специфични негови характеристики на базата на множествена линейна регресия, известна и като модел на Хедоничната стойност (Court, 1939). Хедоничната стойност е дефинирана, като съставна от косвените цени на характеристиките и може да бъде определена, чрез наблюдения върху цените на диференцирани продукти и определен брой характеристики свързани с тях. Иконометрията оценява тези косвени цени, чрез регресионен анализ от първа стъпка (цената на продукта, представена като функция от неговите характеристики) при конструирането на хедонични ценови модели (Rosen, 1974). Философията на този модел е много близка до принципа на корекционните коефициенти, който се прилага при използване на сравнителния подход за оценка на недвижими имоти. Чрез анализ на емпирични резултати от изследване на достатъчен обем данни, съотносими към конкретните обекти на оценяване с методите на Хедоничната регресия, биха могли да бъдат формулирани правила за определяне на съответните корекционни коефициенти.

Коефициентите на модела на множествената линейна регресия, с които се умножават независимите променливи, по същество са частни коефициенти на регресия. Всеки такъв коефициент измерва промяната на зависимата променлива (в случая цена на жилище), която промяна отговаря на увеличение на стойността на съответния фактор с единица при равни други условия (Гоев, 2019). При изграждане на такъв модел, следва да бъдат включени като независими променливи факторите, които обичайно се изследват в процеса на оценяване, например: брой

стаи, площ, състояние на обекта, състояние, вид и възраст на сградата, височина и т.н. При наличието на надеждни резултати от статистическия анализ на данните, коефициентите от модела могат да бъдат трансформирани във вид, който би позволил използването им в практиката на оценителите.

В страните с развити икономики е налична и достъпна огромна по обем информация за пазара на недвижими имоти. В различни интернет платформи се съдържат стотици хиляди оферти за продажби и наеми. Тази информация, макар да е публично достъпна е трудна за анализ и обработка поради огромния си обем. В последните години бяха разработени и придобиха изключителна популярност голям брой софтуерни решения на базата на *AI*, които правят възможна обработката на огромни масиви с данни.

Оценяването на имоти е една трудна и времеемка дейност, изискваща специални знания, често субективна и използваща малък набор от данни (Калчев, 2023). Използването на възможностите предоставяни от *AI* би допринесло значително за оптимизиране на процеса, чрез изследването на по-големи обеми информация.

Основните критерии за селектиране на пазарни свидетелства при прилагане на сравнителния подход за оценка са основно техни технически характеристики, като размер, категория имот, местонахождение, изложение, години, конструкция на сградата, разположение на имота във височина, дата на сделката и др. (Касърова, 2015). Информация за тези характеристики може да бъде намирана в популярните портали за публикуване на обяви за продажба на имоти. Често тази информация е налична в текстовото описание на публикуваните оферти, което прави извличането ѝ изключително труден процес, който отнема значително количество време. За изпълнението на тази задача биха могли да бъдат използвани възможностите на *AI* за анализ и обработка на данни.

В научните издания има много разработки за приложението на *AI* в сферата на недвижимите имоти. Основно се разглежда приложението на този тип технологии при автоматизираното изготвяне на оценки и прогнози за развитието на пазара (Rossini, 2000), развитието на „Проптек“, или дигитализацията на бранша на недвижимата собственост и набор от технологии, променящи подходите за покупка, продажба и управление на имоти (Бобчев, 2020), глобални пазарни ана-

лизи и прогнози насочени към големите инвеститори в имоти (Караиванов, 2021), повишаване на операционната ефективност, потребителския интерес, продуктивността и усъвършенстването на заетите в бранша на недвижимите имоти (Siniak, 2020) и др. Обикновено, авторите разглеждат *AI* и прилагането на модела на Хедоничната стойност, като инструменти използвани при изготвянето на комплексни пазарни оценки и изготвяне на ценови индекси за региони, или дори цели държави (Selim, 2008).

Технология на изследването

В настоящото изследване се поставя фокус върху възможностите за използване на модела на Хедоничната стойност за извеждане на зависимости, които оказват влияние върху косвените цени на различните характеристики включени в модела. След това, на база тези зависимости могат да бъдат определени обективно, коефициенти за корекция на пазарните свидетелства по отношение на изследваните характеристики, които да бъдат прилагани от оценителя при изготвянето на пазарни оценки по сравнителния подход. За да бъде събрана и обработена необходимата за изграждане на модела информация, ще бъдат използвани възможностите на *AI*. Целта на автора е да обърне внимание на технологията, не като заместител на човека в процеса на изготвяне на пазарни оценки, а като инструмент за подпомагане на дейността на оценителите.

Източници на информация и методи на изследването

Използвани са данни за 614 оферти за продажба на апартаменти в 11 Софийски квартала: Дианабад, Докторски паметник, Дървеница, Зона Б-19, Иван Вазов, Изгрев, Мусагеница, Обеля (1 и 2), Оборище, Разсадника и Яворов. Информацията е извлечена от базата данни на интернет сайта www.imot.bg на 10 февруари 2025. Сайтът е избран по метода на основния масив, като един от най-големите източници на информация, съдържащ оферти за продажби и отдаване под наем. Кварталите са избрани така, че да имат различни локационни характеристики по отношение на инфраструктурата, отдалечеността от центъра, транспортната достъпност и не на последно място категорията „престиж“. На базата на тези различия, за отделните локации са определени различни рангове, като с най-ниска стойност са най-отдалечените от центъра

и с най-нисък престиж квартали, а с най-висок районите около Докторски паметник и ул. Оборище, които се отличават с изключително висок престиж, отлична инфраструктура и разположение в централната градска част.

Останалите характеристики, за които е събрана информация са: площ на жилищата, етаж от сградата, на който са разположени, височина на сградата (в етажи), последен етаж (фиктивна (*dummu*) променлива равна на 1, ако етажът е последен), наличие на централно отопление (ТЕЦ), вид на сградата (панел, тухла и ЕПК с рангове от едно до три, като с най-нисък ранг е панела, а с най-висок – тухла), възраст на сградата, състояние на апартамента и наличие на обзавеждане.

Първоначално, данните са извлечени във вид на текстова информация в pdf файлове. Част от характеристиките на имотите, като възраст, степен на завършеност, състояние на апартамента и наличие на обзавеждане се съдържа в свободното описание на офертите. За да бъде извлечено това съдържание в удобен за използване вид е формулирана конкретна задача към *AI*, и в резултат информацията е систематизирана от програмата в табличен вид. Извършена е проверка за достоверност и валидиране на данните.

За постигане на целта на изследването се използва множествена регресия по модела на Хедоничната стойност (1). За зависима променлива е избрана цената за 1 кв.м, на жилищата от извадката, защото именно тази цена се използва, като основа в процеса на оценяване на недвижими имоти. Моделът се описва с формулата:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n + \varepsilon \quad (1)$$

където:

Y е зависимата променлива, в случая цена за 1 кв.м. за апартамент;

b_0 е константа;

X_i са отделните фактори (независими променливи);

b_i са коефициентите на регресия, които показват с колко ще се промени цената на жилището при промяна с единица на съответния фактор, при равни други условия;

ε е остатъчният компонент, или факторите които влияят върху образуването на цената, но не са включени в модела.

Като втора стъпка в анализа е проведен тест за мултиколинearност сред независимите променливи, за да се провери в дали е налична ко-

релационна зависимост между две, или повече променливи, което би поставило под съмнение тяхното участие в модела. За целта са изчислени коефициент на Пиърсън (2) и коефициент на инфлиране VIF (3), по формулите:

$$r_{X,Y} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2} \sqrt{\sum(Y_i - \bar{Y})^2}} \quad (2)$$

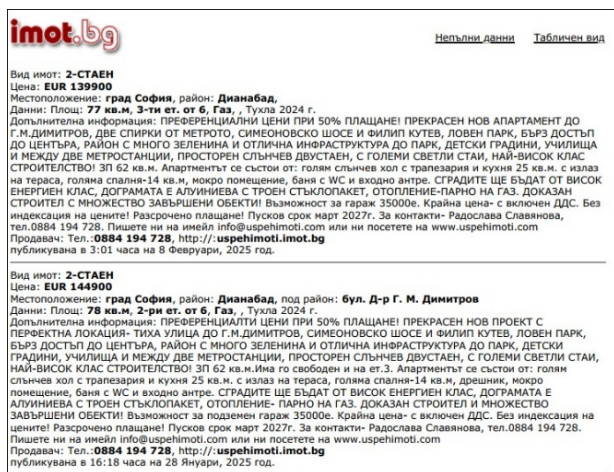
$$VIF[X_i] = \frac{1}{1 - R^2[X_i]} \quad (3)$$

За обработка на данните за пазарните свидетелства е използвана платформата Chat GPT 4o, платен абонамент с конкретно формулирано задание за начин на обработване на информацията от PDF файловете и съставянето на таблици по определените критерии.

За провеждане на статистическите анализи е използван специализиран софтуер IBM SPSS, версия 26.

Анализ и резултати от изследването

При изпълнение на търсене в базата данни на сайта www.imot.bg, за оферти за продажба на едно-, дву- и тристаини апартаменти в 11 избрани квартала на София бяха получени 25 файла във формат pdf, съдържащи данни и описания на 986 броя оферти.



Фиг. 1. PDF файл с оферти

Източник: www.imot.bg

Информацията във файловете е структурирана, като еднотипно описание в което последователно са посочени вида на имота (1-стаен, 2-стаен, 3-стаен), цената му в евро, местоположение (град, район, подрайон), кратки данни (площ, вид строителство, година на построява-

не), допълнителна информация- свободен текст, данни за продавача (телефонен номер и уебсайт) и дата и час на публикуване на офертата (фиг.1).

Характеристиките, които ще бъдат включени в Хедоничния модел, като независими променливи са: брой стаи, локация, площ, етаж, последен етаж, вид сграда, възраст на сградата и състояние на апартамента. Специализирания софтуерен продукт за статистически анализ на данни, който е използван от автора (IBM SPSS), обработва коректно събрани данни, представени в подходящ табличен вид и форматиран така, че да бъдат разбираеми за програмата. За да бъдат обработени събраните в pdf файловете данни е необходимо да бъде разчетена цялата текстова информация и данните да се трансформират в табличен вид. При наличието на голям обем информация, което е задължително изискване за представителност на извадката, това е изключително трудоемка и времеемка задача.

В таблица 1 е представен наборът от променливи, избрани да бъдат включени в конструирания модел.

Таблица 1
Променливи включени в модела

променлива	описание
PPS	Цена за 1 кв.м (зависима променлива Y)
NUMR	Брой стаи (1-;2-;3-стаен)
LOC	Локация по рангови характеристики на квартала (най-добър = 100)
SIZE	Площ на апартамента
FLOOR	Етаж в сградата
LSTFL	Последен етаж (1 = да, 0 = не)
TYPEBLD	Вид сграда (1 = панел; 2 = ЕПК; 3 = тухла)
BLDAGE	Възраст на сградата (2025 минус година на построяване)
COND	Състояние на имота (незавършен = 1; за основен ремонт = 2; за ремонт = 3; за лек ремонт = 4; освежен = 5; след ремонт = 6; луксозен = 7)

Съвременните технологии предоставят възможност, чрез използване на функциите на AI, събраната в процеса на изследването информация да бъде обработена бързо и качествено. Необходимо е да се формулира недвусмислено и конкретно задание към софтуера, за да бъдат извлечени необходимите данни от текста на офертите и за бъдат нанесени коректно в електронна таблица. В резултат програмата генерира xls файл за всеки отделен pdf, в който всеки ред

е различна оферта а в колоните са нанесени данните за характеристиките на имота, които ще бъдат изследвани (фиг.2).

В процеса на обработка на информацията бяха премахнати дублирани оферти, както и тези за които не е налична информация за повече от две от основните характеристики. Така броят на офертите, подлежащи на анализ беше редуциран до 614 оферти, приведени във вид, който позволява да бъдат обработени в IBM SPSS.

Пореден номер	Местоположение	Цена (EUR)	Площ (кв.м)	Етаж	Етажност	ТЕЦ	Тип строителство	Година на построяване	Състояние	Обявен
1	София, Дланбад	139900	77	3	6	Не	Тухла	2024	Незавършен	Не
2	София, Дланбад	144900	78	2	6	Не	Тухла	2024	Незавършен	Не
3	София, Дланбад	149400	86	N/A	N/A	Не	Тухла	2024	Незавършен	Не
4	София, Дланбад	162600	56	4	7	Да	Тухла	2025	Незавършен	Не
5	София, Дланбад	173595	60	2	7	Да	Тухла	2026	Лускозен	Не
6	София, Дланбад	174005	59	2	6	Да	Тухла	2026	Лускозен	Не
7	София, Дланбад	178343	62	2	7	Да	Тухла	2026	Лускозен	Не
8	София, Дланбад	178911	60	2	6	Да	Тухла	2026	Лускозен	Не
9	София, Дланбад	199000	73	3	18	N/A	Тухла	N/A	Незавършен	Не
10	София, Дланбад	196000	70	8	20	Да	Тухла	N/A	След ремонт	Да
11	София, Дланбад	200000	91	3	5	Да	Тухла	2023	Незавършен	Не
12	София, Дланбад	200000	89	4	5	Да	Тухла	2024	Незавършен	Не
13	София, Дланбад	200000	89	3	5	Да	Тухла	2024	Незавършен	Не
14	София, Дланбад	205000	73	13	16	Да	ЕПК	1980	Обявен	Да
15	София, Дланбад	205000	73	13	16	Да	ЕПК	1981	Обявен	Да
16	София, Дланбад	208998	75	13	16	Не	ЕПК	1981	За лек ремо	Не
17	София, Дланбад	209900	72	2	7	Да	Тухла	2025	Незавършен	Не
18	София, Дланбад	210000	61	4	4	Да	Панел	1975	След ремонт	Да
19	София, Дланбад	214338	73	2	7	Да	Тухла	2026	Лускозен	Не
20	София, Дланбад	217000	94	6	8	Да	Тухла	2010	Обявен	Да
21	София, Дланбад	217420	74	2	7	Да	Тухла	2026	Лускозен	Не
22	София, Дланбад	182000	75	7	8	N/A	Тухла	N/A	Незавършен	Не
23	София, Дланбад	220000	75	1	5	Да	Тухла	2021	След ремонт	Да
24	София, Дланбад	230000	80	6	8	Да	Тухла	2003	Обявен	Да

Фиг. 2. Макет за съдържанието на XLS файл с данни за имоти включени в извадката

Източник: www.chatgpt.com

След обработката на данните от специализирания софтуер за статистически анализ (бяха получени следните резултати:

Коефициентът на корелация R е 0.708, което показва значителна степен на зависимост на цената на жилищата от факторите, използвани при конструирането на модела.

Коефициентът на детерминация R^2 е 0.502, което означава, че наборът от характеристики е „отговорен“ за формирането 50,2% от цената на апартаментите. Останалите 49,8% се формират от други характеристики, които не са попаднали в модела. Корикираният спрямо големината на извадката коефициент на детерминация е 0,495, не се различава значително от R^2 , което показва, че моделът е стабилен (таблица 2).

Анализът на дисперсията (ANOVA тест) показва F-характеристика 76,083 и значимост (sig.) $0,000 < 0,05$, което показва, че поне една от независимите променливи оказва статистически значимо влияние върху цената на апартаментите и моделът, като цяло е статистически значим (таблица 2).

Таблица 2

Параметри на изчисления множествен линеен регресионен модел и проверка на статистическата значимост на параметрите

Y	Min.	Max.	Средна	Ст.откл.
PPS	1265.94	6600.00	2694.08	920.78
резюме на модела				
R	R ²	кор.R ²		
0,708	0,502	0,495		
ANOVA				
sig.	0,000	<α (0,05)		
регресионен анализ				
характеристика	b коеф.	ст. β	sig.	
(Константа)	485,03		0,010	
NUMR	146,63	,091	0,027	
LOC	18,441	,454	0,000	
SIZE	-6,942	-,284	0,000	
FLOOR	15,169	,062	0,053	
LSTFL	-183,45	-,067	0,026	
TYPEBLD	230,645	,150	0,000	
BLDAGE	4,408	,199	0,000	
COND	101,466	,292	0,000	

Източник: изчисления на автора

Статистическо значимо ниво на влияние върху образуването на зависимата променлива „Цена на апартамент“ оказват факторите, чието равнище на значимост (sig.) е по малко от вероятността от грешка ($\alpha=0.05$). Анализът на получените след изчисляване на множествената линейна регресия коефициенти води до следните заключения:

1. На първо място по значимост е факторът „Локация“ със sig.=0,000, и стандартизиран коефициент $\beta=0,454$. По-високият ранг на локацията води до по-висока цена за 1 кв.м на апартамент. В случая единица повишение в йерархичната характеристика на локациите добавя 18,44 евро в цената за един квадратен метър застроена площ.

2. Площ: стандартизиран коефициент $\beta=-0,284$ и sig.=0.000. Този фактор очаквано има много високо ниво на влияние върху цената на апартаментите. В случая влиянието е отрицателно. Всеки квадратен метър увеличение в площта на жилището води до по-ниска с 6,94 евро цена за един квадратен метър застроена площ.

3. Факторът „състояние на апартамент“ със sig.=0,000 и стандартизиран коефициент $\beta=0,292$ е третият по значимост в модела. По-доброто състояние на апартамент увеличава неговата цена за квадратен метър със 101,47 евро.

4. Видът на сградата и нейната възраст също имат висока статистическа значимост при формирането на стойността на зависимата променлива sig. и за двете променливи е 0,000, като стандартизираните коефициенти β са съответно 0,150 и 0,199.

7. Сравнително висока статистическа значимост имат и факторите „последен етаж“ и „брой стаи“, като очаквано влиянието на последния етаж върху стойността за един квадратен метър е отрицателна, т.е. когато жилището е разположено на последен етаж в сградата, неговата цена е по-ниска.

8. Факторът „етаж“ има много близка до граничната стойност на sig., 0,053 и стандартизирана $\beta=0,062$. Това означава, че този фактор има минимално влияние върху образуването на стойността на зависимата променлива, което се доближава до това да бъде статистически значимо. Възможно е при оптимизиране на модела и изключване и коригиране на някои от останалите фактори, да се окаже че етажът на който е разположен апартаментът, има значение за обра-

зуването на неговата цена.

Извършена е проверка за наличието на мултиколинеарност между някои от факторите, които са включени в модела (таблица 3).

Резултатите от теста за наличие на корелация по-двойки от участващите в модела променливи не показаха наличието на значителна корелация между факторите включени в модела.

За допълнителна проверка на това, дали между някои от променливите съществува корелационна зависимост, която би могла да окаже влияние върху резултатите от анализа и следва да бъдат отстранени от модела поради наличието

на мултиколинеарност са изчислени вариационно-инфлиращите фактори (VIF).

За нито една от променливите стойността на VIF не надхвърля 10. Най-високата стойност на коефициента е за фактора „площ“, но стойността му 2,499 е доста под приетото за проблемно ниво 10. Това означава, че в модела няма проблем с колинеарност, който би довел до изкривяване на резултатите от анализа (таблица 3).

Таблица 3

Проверка за наличие на мултиколинеарност

Променлива	Най-висок коефициент на корелация			VIF
NUMR	коэф. на Пиърсън	0,675	(площ)	2,041
	sig (2-tailed)	0,000		
LOC	коэф. на Пиърсън	0,128	(площ)	1,316
	sig (2-tailed)	0,001		
SIZE	коэф. на Пиърсън	0,675	(брой стаи)	2,499
	sig (2-tailed)	0,000		
FLOOR	коэф. на Пиърсън	0,233	(площ)	1,246
	sig (2-tailed)	0,000		
LSTFL	коэф. на Пиърсън	0,183	(етаж)	1,110
	sig (2-tailed)	0,000		
TYPEBLD	коэф. на Пиърсън	0,329	(възр.сгр.)	1,470
	коэф. на sig (2-tailed)	0,000		
BLDAGE	коэф. на Пиърсън	0,329	(вид сгр.)	1,630
	sig (2-tailed)	0,000		
COND	коэф. на Пиърсън	0,296	(възр.сгр.)	1,277
	sig (2-tailed)	0,000		

Източник: изчисления на автора

На база получените от регресионния анализ резултати, могат да бъдат получени корекционни коефициенти за привеждане на характеристиките на пазарни аналози към обект на оценка по сравнителния подход. Един от възможните начини за това е да бъде приложен регресионният коефициент за съответната характеристика

(b_i) към средната цена за един квадратен метър за офертите от извадката по формулата (4):

$$C_c = 1 + \frac{b_i}{\bar{Y}} \quad (4)$$

Ако разгледаме променливата LSTFL (пос-

леден етаж), регресионният коефициент е отрицателен, -183,45 (табл.2), т.е. последният етаж означава цена за квадратен метър, по ниска със 183,45 евро за квадратен метър в сравнение с жилище, което не е на последен етаж, при равни други условия. Средната стойност на зависимата променлива PPS (цена за квадратен метър) в извадката е 2694,08 евро (табл.2). По формулата (4) за изчисляване на коефициента на корекция:

$$C_c = 1 - \frac{183,45}{2694,08} = 1 - 0,068 = 0,932$$

Така полученият коефициент би бил валиден при изготвяне на оценка на имот, който по качествен и териториален признак, попада в обхвата на извадката, използвана при извършването на регресионния анализ.

Пример: изготвяне на оценка на жилище на последен етаж, находящо се в териториалния обхват на извадката използвана в Хедоничния регресионен модел. При прилагане на сравнителния подход на оценяване са открити пазарни свидетелства, които представляват подобни жилища на не-последни етажи. За приравняване на цената за един квадратен метър от аналозите по този показател, следва тя да бъде умножена с корекционен коефициент 0,932. Така, при равни други условия се получава стойност на оценявания обект, на база сравнение с подобни на него пазарни свидетелства по отношение на останалите им характеристики.

За разлика от коефициента на регресия (b_i), който се отнася към средната стойност на зависимата променлива, корекционният коефициент изчислен по-горе, се прилага към цената на аналозите, открити в процеса на оценяване. По същия начин, биха могли да бъдат определени корекционни коефициенти и за останалите характеристики, които според оценителя биха били релевантни към формирането на стойността на оценявания обект.

Корекционни коефициенти, изведени по подобен метод, могат да бъдат прилагани в определен период от време, в който не са наблюдавани значителни промени в пазарните условия, в териториален обхват, съвпадащ с ограниченията на извадката заложен в Хедоничния модел, и по отношение на типовете имоти, включени в модела.

Заклучение

Въз основа на това базово проучване, след извършване на по-широкообхвтен анализ, с включване на по-голям брой данни в извадката, прилагане на различни модели на множествена регресия и провеждане на различни изследователски експерименти с цел апробация, би могло да бъде съставена методология за определяне на обективни, емпирично тествани корекционни коефициенти, приложими при изготвянето на оценки на недвижими имоти по сравнителния подход.

Получените резултати в проведеното изследване изцяло потвърждават основната изследователска теза. Изкуственият интелект позволява да се формира представителна извадка от големи съвкупности и тя да се приведе във вид подходящ за обработка с помощта на специализиран статистически софтуер.

В конкретния случай с помощта на селектираните данни чрез използването на изкуствен интелект с общ достъп, какъвто е ChatGPT бяха изчислени параметрите на множествен линейен регресионен модел като приложение на Хедонична регресия и беше апробирана методика за изчисляване на корекционен коефициент за уточняване на пазарна цена на имот, валиден за изучаваната съвкупност. Беше установено, че колкото по-добре е формулирано заданието към AI, толкова по-качествени резултати ще бъдат получени.

Беше установено, че локацията, площта и състоянието на апартаента в най-голяма степен влияят върху образуването на цената за един квадратен метър застроена площ в изследваните квартали на София. Локацията и състоянието имат положително влияние, докато площта влияе отрицателно, т.е. апартаментите с по-голяма площ имат по-ниска цена за един квадратен метър. В значителна степен във формирането на изследваната зависима променлива участват и факторите вид и възраст на сградата, брой стаи и разположение на последен етаж. Последният етаж влияе отрицателно върху цената, докато останалите фактори са с положителен ефект. Факторът „етаж“ има гранична стойност на коефициента за значимост, което означава, че в конкретния модел неговото влияние върху зависимата променлива е много малко от статистическа гледна точка. Това би могло да се тълкува и като сигнал за несъвършенство на модела, като за по-задълбочен анализ би могло да се приложи клъстерен анализ по отношение

на етажа, като бъдат формирани групи от етажи (напр. от 2-ти до 4-ти, от 5-ти до 7-ми, над 8-ми, над 12-ти и т.н.) за да се провери дали има статистически значима разлика в цените за квадратен метър площ за отделните клъстери.

Тестът за мултиколинearност потвърди, че между включените в модела променливи няма зависимост, която би станала причина за изкривяване на резултатите от анализа и би довела до необходимост за изключване на някой от факторите от модела.

Получените чрез метода на линейния регресионен анализ коефициенти b_i за отделните фактори, могат да послужат за разработване на модел за определяне на корекционни коефициенти при прилагане на сравнителния подход за оценка на недвижими имоти. За целта е необходимо да бъде изготвен подобен анализ на извадка, която да включва данни за по-голям брой оферти в повече квартали с различни характеристики, които могат да бъдат обединявани в групи, чрез клъстерен анализ и други статистически методи.

Използването на помощта на *AI* ще доведе до значително съкращаване на времето за подготовка на необходимите данни и ще направи възможно по нататъшното прецизиране на модела. На база посоченото би могъл да бъде направен изводът, че приложението на изкуствения интелект в оценителската практика, значително би улеснило и ускорило извършването на пазарни оценки за стойностите на недвижимите имоти по сравнителния метод и би подобрило тяхната точност.

Литература

Бобчев, К. (2020). Proptech предизвикателствата на технологиите в сферата на недвижимите имоти. *Недвижими имоти & Бизнес*, Том IV (2).

(Bobchev, K. (2020). Proptech predizvikelstvoto na tehnologiite v sferata na nedvijimite imoti. *Nedvijimi imoti & Bisnes*, Tom IV (2).

Гоев, В. и колектив (2019). Статистически анализ в социологически, икономически и бизнес изследвания. София, ИК на УНСС

(Goev, V. et al. (2019). Statisticheski analiz v sociologicheski, Ikonomicheski biznes izsledvania. Sofa, IK na UNSS)

Кабаиванов С., Венета Марковска (2021), Изкуственият интелект в анализите на пазара на недвижими имоти, "AIP Conference Proceedings 2333, 030001" (2021), налично в интернет на адрес: <https://doi.org/10.1063/5.0041806>

(Kabaivanov, S., Veneta Markovska (2021), Izkustveniat intellect v analizite na pazara na nedvijimi imoti, "AIP Conference Proceedings 2333, 030001" (2021). Available at: <https://doi.org/10.1063/5.0041806>

Калинков, К., (2007), „Недвижима собственост“. ИК „Геа-принт“, Варна)

(Kalinkov, K. (2007), "Nedvijima sobstvenost". IK "Gea-print", Varna).

Калчев, Щ. (2023), „Аспекти на автоматизираните модели за оценка и приложения в сферата на недвижимите имоти“, Списание *Недвижими имоти&Бизнес*, том VII (3), УНСС София, 2023

(Kalchev, Sht. (2023), "Aspekti na avtomatiziranite modeli za ocenka i prilozhenie v sferata na nedvizhimite imoti", *Spisanie Nedvizhimi imoti&Biznes*, tom VII(3), UNSS Sofia, 2023)

Касарова, В., Нихгос Канарян, (2015), „Приложение на пазарни множители при оценяване на недвижими имоти по сравнителния метод“. Доклад от форума *Оценителската дейност в Гърция и България-практика и перспективи*, КНОБ, А.ВА.Г&ПК София-град, София област, Солун, Гърция, 2015, достъпно на адрес: [www://eprints.nbu.bg/id/eprint/2573/1/Kasarova_Kanaryan-16%20May%202015.pdf](http://eprints.nbu.bg/id/eprint/2573/1/Kasarova_Kanaryan-16%20May%202015.pdf)

(Kasarova, V., Nihgos Kanaryan, (2015), "Prilozhenie na pazarni mnozhiteli pri ocenivane na nedvizhimi imoti po sravnitelnia metdo" Doklad ot foruma Otsenitelskata deynost v Gurcia I Bulgaria-praktika I perspektivi, KNOB, A.VA.G&RK Sofia-grad, Sofia-oblast, Thessaloniki, Greece, 2015. Available at: [www://eprints.nbu.bg/id/eprint/2573/1/Kasarova_Kanaryan-16%20May%202015.Pdf](http://eprints.nbu.bg/id/eprint/2573/1/Kasarova_Kanaryan-16%20May%202015.Pdf)

Пелов, Т. (2011). „Основи на оценяването на недвижими имоти“ в *Икономика и управление на недвижимата собственост. Сборник студии* (стр.44-57). София, Университетско издание "Стопанство".

(Pelov, T. (2011). Osnovi na ocenivaneeto na nedvizhimi imoti, "Ikonomika I upravlenie na nedvizhimata sobstvenost. Sbornik studii" (pp.44-57). Sofia, universitetsko izdanie "Stopanstvo").

Стоенчева, Я. (2020). „Привлекателност и достъпност на инвестицията в жилищни имоти в териториален аспект“ в *Научни трудове на УНСС*, том,2, София (стр.279-300)

(Stoencheva, Y. (2020). Privlekatelnost i dostapnost na investiziatia v zhlishtni imoti v teritorialen aspekt [Attractiveness and Accessibility of Investment in Residential Real Estate from a Territorial Perspective]. *Nauchni trudove na UNSS*

[Research papers UNWE], vol. 2, Sofia (279-300)

Court, A.T. (1939). "Hedonic Price Indexes with Automotive Examples", *The Dynamics of Automobile Demand*, 99-119, General Motors Corporation. Available at: https://allengoodman.wayne.edu/Research/PUBS/Court_Hedonic.pdf

Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition.", *Journal of Political Economy*, Vol. 82, No. 1 (Jan. - Feb., 1974), pp. 34-55. Available at: <http://www.jstor.org/stable/1830899>

Rossini, P., (2000). Using Expert Systems and Artificial Intelligence For Real Estate Forecasting, *Sixth Annual Pacific-Rim Real Estate Society Conference Sydney, Australia, 24-27 January 2000*. Available at: www.citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7f48a4af041d4992704390651f8cfcb04903528e

Selim, H. (2009). Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. *Expert Systems with Application*. Vol. 36, Issue 2, Part 2 (2009) pp.2843-2852. Elsevier. Available at: <https://doi.org/10.1016/j.eswa.2008.01.044>

Siniak, N., et al. (2020). "The impact of proptech on real estate industry growth". *IOP Conference Series: Materials Science and Engineering*. Available at: <https://iopscience.iop.org/article/10.1088/1757-899X/869/6/062041>

Stoenchev, N., Hrischeva, Y. (2018) Methodology for conducting a market analysis and valuation of a residential property subject to sale, using a comparative approach. *The Valuation Journal, National Association of Romanian Valuers*, vol. 13 (1), pp. 79-99, ISSN:1842-3787

Assoc. Prof. Dr. Yana Georgieva Stoencheva
University of National and World Economy,
Real Estate Department
1700 Sofia, Student Town, UNWE
E-mail: hhrischeva.y@abv.bg

Borislav Stanoev Borisov
University of National and World Economy,
Real Estate Department
1142, Sofia, 62 Gen. Gurko str, entr. A ap.№3
E-mail: bstanoeff@gmail.com