

DOI: <https://doi.org/10.32782/2524-0072/2022-43-60>

УДК 332.6

ЕКОНОМІКО-МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ВПЛИВУ МАКРОЕКОНОМІЧНИХ ФАКТОРІВ НА ЦІНОУТВОРЕННЯ РИНКУ НЕРУХОМОСТІ

ECONOMIC AND MATHEMATICAL MODELING OF THE INFLUENCE OF MACROECONOMIC FACTORS ON REAL ESTATE MARKET PRICING

Пишнограєв Іван Олександрович

кандидат фізико-математичних наук, доцент,
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: <https://orcid.org/0000-0002-3346-8318>

Коваль Анастасія Миколаївна

студентка,
Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
ORCID: <https://orcid.org/0000-0002-5007-4652>

Pyshnograiev Ivan, Koval Anastasiia

National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"

У статті розглянуто вплив різних факторів вартість житла та важливість дослідження даного питання для держави. Виділено потенційні макроекономічні та регіональні фактори впливу на ціноутворення ринку нерухомості в місті Київ. На основі статистичних даних державної служби України, проведено регресійно-кореляційний аналіз задля уникнення мультиколінійності в лінійній багатофакторній моделі. При цьому було застосовано такі критерії як коефіцієнт кореляції Пірсона, метод прямого відбору, критерій Фішера, значення показника p-value. В результаті була отримана модель та досліджена на адекватність. Також було обґрунтовано доцільність використання методу машинного навчання та побудована модель Random Forest на мові програмування R. Результати порівняння моделей показали, що машинне навчання, а саме метод Random Forest, є більш ефективним для дослідження складних питань, таких як вплив макроекономічних факторів на ціноутворення ринку нерухомості. Даний метод може бути корисним доповненням до гедонічної регресії, оскільки він більш адекватно фіксує складність або нелінійність. В ході дослідження було з'ясовано, що на ціну за квадратний метр у новобудовах впливають такі фактори як: облікова ставка, курс валюти, витрати населення на одну особу, валовий регіональний продукт, прожитковий мінімум та чисельність наявного населення.

Ключові слова: ціна, ціноутворення, нерухомість, макроекономічні фактори, машинне навчання.

The process of determining and forecasting the pricing of the real estate market has a high degree of complexity due to the dependence of market processes on many meso- and macroeconomic factors of the national economy. The purpose of the work is to create a model for assessing the impact of various factors on the price per square meter of the primary market in Kyiv. It has been identified potential macroeconomic and regional factors influencing the pricing of the real estate market in Kyiv. The construction of a linear multivariate model was based on statistical data from the civil service of Ukraine, This model allows us to estimate the correlation between two or more independent variables and one dependent variable. The regression-correlation analysis was conducted to avoid multicollinearity. The following criteria were applied: correlation coefficient Pearson, direct selection method, Fisher's criterion, and p-value. Some factors were excluded from the process. As a result, a model with fairly high accuracy was obtained. Also, the expediency of using the machine learning method was substantiated and the model of Random Forest was built by using the R programming language. The sample was divided into two parts, one of the artificial intelligence was trained (80%), and on the other, there was a model testing by forecasting (20%). The big advantage of

the random forest is that we can find out which factors are more important for our model. In the process of model building, the most insignificant factors were excluded and we got a model with a quite high accuracy. The results of the comparison of models showed that machine learning, namely the Random Forest method, is more effective for the study of complex issues, such as the impact of macroeconomic factors on the pricing of the real estate market. During the study, it was found that the price per square meter in new buildings is influenced by such factors as the discount rate, exchange rate, per capita expenditure, gross regional product, subsistence minimum, and the number of the available population.

Keywords: price, pricing, real estate, macroeconomic factors, machine learning.

Постановка проблеми. Підвищення ефективності функціонування ринку житлової нерухомості є масштабним економічним завданням, що обумовлено великою кількістю задіяних матеріальних, фінансових і трудових ресурсів.

Масова оцінка і прогнозування ринкової вартості об'єктів житлової нерухомості є одним із способів підвищення ефективності даного ринку. Однак процес визначення і прогнозування ціноутворення ринку нерухомості має високу ступінь складності, що обумовлено залежністю ринкових процесів від безлічі мезо- і макроекономічних факторів національної економіки. Тому ці економічні процеси доречно досліджувати за допомогою економіко-математичного моделювання.

Аналіз останніх досліджень і публікацій.

Тему впливу різних макроекономічних факторів на ціноутворення ринку нерухомості піднімали та досліджували чимало як вітчизняних так і зарубіжних науковців, а саме А. М. Іванченко, М. В. Кірносова, М. М. Білаш, С. В. Грибовський, Джозеф К. Еккерт, Л. А. Різва, М. А. Федотова, К. В. Загребельна та інші [1]. Проте питання впливу макроекономічних факторів на ціну саме в місті Київ є досить не дослідженим.

Формулювання цілей дослідження.

Метою даної роботи є побудова моделі оцінки впливу різних факторів на ціну за кв. м первинного ринку в місті Київ.

Виклад основного матеріалу дослідження. Поліпшення ефективності державного управління в сфері нерухомості можна досягти за рахунок надійних оцінок, які показали б, як справжній стан ринку, так і потенційно можливий. Ціна об'єктів нерухомості піддається впливу безлічі факторів, таким як характеристика конкретного об'єкта нерухомості, регіональні фактори, які безпосередньо не пов'язані з об'єктом, але впливають на ситуацію і тенденції. Наразі більшість досліджень, які проводяться на дану тему здійснені на рівні країни, проте важливу роль відіграє оцінка нерухомості на рівні регіонів, оскільки кожен з них відрізняється рівнем життя насе-

лення, рівнем середньої заробітної плати, рівнем безробіття та іншими факторами [2].

В такому розрізі мова може йти про розвиток ринку нерухомості за допомогою житлових програм, стимулюванні будівельної діяльності та інвестицій, розвитку доступності іпотечного кредитування.

За рахунок проведення заходів в зазначених напрямках, держава і регіональні інститути управління можуть успішно впливати на ринок нерухомості на регіональному рівні. Однак найчастіше можна спостерігати мінімальну ефективність проведених програм у зв'язку з недостатньою розвиненістю аналітики в даній сфері, яка б враховувала індивідуальні особливості та характеристики регіонів [3; 4].

Оскільки багато людей з різних куточків України прагнуть переїхати саме до столиці і придбати там квартиру, то попит на первинному ринку в Києві значно переважає попит у інших містах, тому доцільно окремо дослідити ціноутворення в столиці та визначити фактори впливу.

Побудова невизначеного майбутнього на ринку первинної нерухомості в місті Київ дозволила б знизити ризики як для органів державного управління, а також для пересічних громадян, які бажають придбати або продати квартиру, ріелторів та інвесторів.

Дослідження здійснимо за допомогою багатофакторної лінійної регресійної моделі на основі статистичних даних. Дана модель дозволяє оцінити зв'язок між двома або більше незалежними змінними та однією залежною змінною [5].

Множина регресія має такий загальний математичний вигляд (1) [5]:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \dots + \beta_n X_n + \mu, \quad (1)$$

де Y – залежна змінна;

$\beta_0, \beta_1, \beta_n$ – коефіцієнти регресії, які показують яким є вплив незалежної змінної на залежну;

X_i – пояснювальні змінні;

μ – випадкова величина.

У нашому випадку в якості залежної змінної виступатиме середня ціна за кв. м. на пер-

Таблиця 1

F-критерій для факторів

Фактор	Значення F-критерію
Обсяг прийняття в експлуатацію всього м кв. у місті Київ	0,204
Рівень безробіття (%)	0,357
Рівень інфляції (%)	0,181
Облікова ставка (% річних), за даними НБУ України	28,465
Капітальні інвестиції в будівельну діяльність (грн) по місту Київ	12,367
Чисельність наявного населення в місті Київ (осіб)	21,182
Прожитковий мінімум (грн.), з розрахунку на рік	26,498

Джерело: сформовано автором на основі відомостей [6; 8]

винному ринку у м. Київ, а в якості незалежних змінних – одинадцять різних соціально-економічних показників. Таких як: обсяг прийняття в експлуатацію всього м кв. у місті Київ, рівень безробіття (%), рівень інфляції (%), рівень середньої заробітної плати (грн) у місті Київ, у розрахунку на рік, курс гривні по відношенню до долара США (грн), валовий регіональний продукт м. Київ на душу населення (грн), облікова ставка (% річних), за даними НБУ України, чисельність наявного населення в місті Київ (осіб), прожитковий мінімум (грн), з розрахунку на рік, витрати населення на одну особу (грн) по місту Київ та капітальні інвестиції в будівельну діяльність (грн) по місту Київ.

Перед побудовою моделі необхідно провести аналіз парної кореляції незалежних факторів. Для цього було використано коефіцієнт кореляції Пірсона [7].

В результаті дослідження було виявлено, що найбільша кореляція, а саме 0,995 спостерігається між такими парами показників як значення валового регіонального продукту та середньої заробітної плати, а також між витратами населення та рівнем середньої заробітної плати. Також саме ці фактори складають з іншими показниками пари з кореляцією більшою ніж 0,95.

В наслідок цього, було прийнято рішення виключити вище перелічені змінні з аналізу для збільшення точності результатів і уникнення мультиколінійності. Оскільки лінійна залежність між пояснючими змінними не дасть змогу реально оцінити вплив кожного із факторів на ціну і може стати причиною не точної інтерпретації моделі.

Далі було використано метод прямого відбору та обчислено за допомогою критерію Фішера кореляцію кожного показника з залежною змінною (табл. 1).

При рівні значимості $\alpha = 0,05$ і степенях свободи $f_1 = 1$ та $f_2 = 13$, $F_{табличне}$ буде дорівнювати 4,67 [9]. Отже, згідно до проведеного прямого відбору до моделі необхідно включити значення курсу гривні по відношенню до долара США, облікову ставку (% річних), кількість капітальних інвестицій в будівельну діяльність, чисельність наявного населення та прожитковий мінімум.

Далі було виявлено, що значення p -value показника investment значно перевищує норму 0,05, тому було прийнято рішення виключити і даний чинник.

В результаті отримали модель зі значеннями: $R^2 = 0,990$, скоригований R^2 дорівнює 0,985, стандартна помилка оцінки= 678,168.

Значення R^2 означає, що варіація ціни за кв. метр на первинному ринку нерухомості в м. Київ на 99% обумовлена наведеними факторами, що є гарною оцінкою для побудованої моделі.

Побудована у ході дослідження модель представлена у формулі (2).

$$Price(m^2) = 181629,336 + 0,848 * subsistence_{min} - 0,066 * population + 562,634 * discount rate + 409,192 * currency. \quad (2)$$

Тобто при збільшенні прожиткового мінімуму на 1 гривню за рік ціна квадратного метра на первинну житлову нерухомість у місті Київ піднімається на 0,85 грн, а при зростанні чисельності фактичного населення в місті Київ на одну особу – знижується на 0,07 грн.. Зі зростанням облікової ставки на 1% ціна за квадратний метр збільшується на 562,63 грн, а ріст курсу гривні по відношенню до долара США провокує збільшення ціни на 409,19 гривень.

Лінійні методи мають ряд важливих переваг: швидко навчаються, здатні працювати з великою кількістю об'єктів та ознак, прості в оцінці та інтерпретації отриманих коефіцієнтів регресії. При цьому вони мають і значний недолік – враховують лише лінійні зв'язки між цільовою змінною та факторами [8].

Процес ціноутворення на ринку нерухомості є доволі складним явищем, тому між ціною та досліджуваними факторами впливу можуть виявитися і нелінійні зв'язки. Оскільки

лінійна багатофакторна регресійна модель не здатна їх врахувати, доцільно дослідити це питання та побудувати модель за допомогою машинного навчання.

Методи класу машинного навчання відносять до методів штучного інтелекту. Їх суть у створенні алгоритмів, які передбачають можливість самонавчання [10; 11]. Для вирішення поставленої задачі, було обрано метод *Random Forest*.

Побудову випадкового лісу будемо здійснювати на мові програмування *R*. Вибірку було поділено таким чином, що 80% вибірки пішло на навчання і 20% на прогнозування.

Великою перевагою випадкового лісу є те що можна дізнатися які із факторів є більш вагомими для нашої моделі (рис. 1).



Джерело: результати, отримані автором на основі побудови моделі *Random Forest* на мові програмування *R*

Найбільше відсотків загального результату визначають такі фактори: чисельність наявного населення, прожитковий мінімум та валовий регіональний продукт на душу населення. В цей же час рівень безробіття, обсяг прийняття в експлуатацію метрів квадратних та рівень інфляції мають найнижчий ступінь важливості, який складає близько нуля.

Було прийнято рішення поступово виключати з моделі найменш вагомні фактори та порівняти характеристики отриманих моделей.

В результаті було з'ясовано, що модель яка містить, такі фактори як облікова ставка, курс валюти, витрати населення на одну особу, валовий регіональний продукт, прожитковий мінімум та чисельність наявного населення є більш точною.

Наведемо характеристики кінцевої моделі. Її побудова виконувалася по 500 деревам, в яких було використано 92,74% змінних.

Середнє значення квадратних залишків складає 2400857. Мінімальна точка розташована на значенні кількості дерев ≈ 45 , тобто це є достатня кількість дерев для прогнозування. Відсоток в середньому відхилення від медіани нашої вибірки складає 0,14%. Найбільший відсоток відхилення складає 16%.

Оскільки ми ділили нашу вибірку на навчальну та тренувальну у розмірності 80% і 20% відповідно, а загальна кількість даних була взята за 13 років, то на тренування припали дані за три роки, а саме за 2010, 2011 та 2015. Згідно моделі ціна за 2010 рік складає 13479.64, що є на 2% менше за реальну. Ціна за 2011 рік була спрогнозована на 2% більше за реальну, а за 2015 рік на 6% менше.

Був проведений аналіз якості побудованих моделей. Середня абсолютна похибка в моделі *RF* склала 1196,75, що на 13% менше ніж у лінійній регресійній моделі. Середнє відсоткове відхилення між прогнозованою та фактичною ринковою ціною у моделі випадкового лісу склало близько 8%, тоді як у лінійній багатофакторній моделі цей показник дорівнював 9%. Середньоквадратична похибка моделі *Random Forest* менша на 21,8 відсотків ніж регресійної моделі, а *RMSE* менше на 11,6 відсотків.

Ці результати показують, що машинне навчання, а саме метод *Random Forest*, є більш ефективним для дослідження складних питань, таких як вплив макроекономічних факторів на ціноутворення ринку нерухомості. Даний метод може бути корисним доповненням до гедонічної регресії, оскільки він більш адекватно фіксує складність або нелінійність.

Проведене дослідження можна використовувати для відтворення аналізу впливу на ціноутворення нерухомості з іншими наборами даних. Більш точний прогноз розвитку ринку нерухомості повинен використовувати не тільки змінні, що стосуються місцевих особливостей ринку і економічних факторів, а і місця розташування і конкуренцію забудовників. Вдосконалена модель може показати практичні відгуки для всього суспільства. У міру зростання населення та урбанізації потреба у моделі, яка здатна вловити можливість еволюцію ринку нерухомості зростає та турбує все більшу кількість зацікавлених сторін, від пересічних громадян до компаній нерухомості та страхових компаній.

Висновки. Для досягнення поставленої мети було використано статистичні дані державної служби статистики України. Результатом дослідження є моделі, побудовані за

допомогою кореляційно-регресійного аналізу та машинного навчання. На основі їх апробації було виявлено, що машинне навчання, є більш ефективним для дослідження такого складного питання як ціноутворення на ринку нерухомості. В ході дослідження було

з'ясовано, що на ціну за квадратний метр у новобудовах впливають такі фактори як: облікова ставка, курс валюти, витрати населення на одну особу, валовий регіональний продукт, прожитковий мінімум та чисельність наявного населення.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Загребельна К. В. Динаміка кон'юнктури ринку нерухомості в Україні. *Економіка та держава*. 2018. № 11. С. 102–106.
2. Кобзан С. Формування ринку нерухомості. Практичні аспекти та особливості оцінки. 2019. 212 с.
3. Особливості ринку нерухомості. URL: https://stud.com.ua/44563/ekonomika/osoblivosti_rinku_neruhomosti.
4. Кучеренко В. Р., Заєць М. А., Захарченко О. В., Сментина Н. В., Улибіна В. О. Оцінка та управління нерухомістю : навчальний посібник. Одеса : Видавництво ТОВ «Лерадрук», 2013. 272 с.
5. Літнарівич Р. М. Побудова і дослідження математичної моделі за джерелами експериментальних даних методами регресійного аналізу. Навчальний посібник. Рівне : МЕРУ, 2011. 140 с.
6. Державна служба статистики України. URL: <http://www.ukrstat.gov.ua>.
7. Кореляційний аналіз. URL: https://pidru4niki.com/12461220/statistika/korelyatsiyiy_analiz.
8. Лінійний множинний регресійний аналіз. URL: <http://ekonometrika.ho.ua/4.htm>.
9. Таблиці функцій та критичних точок розподілів. Розділи: Теорія ймовірностей. Математична статистика. Математичні методи в психології. / Укладач: М. М. Горонескуль. Харків : УЦЗУ, 2009. 90 с.
10. Масова оцінка житлової нерухомості з використанням методів машинного навчання. URL: <https://www.ukrlogos.in.ua/10.11232-2663-4139.16.50.html>.
11. Clinton S. *Tree-based Machine Learning Algorithms: Decision Trees, Random Forests, and Boosting* / Sheppard Clinton., 2019.

REFERENCES:

1. Zagrebelska K. V. (2018) Dynamics of the real estate market in Ukraine. *Economy and state*, 11, 102–106. (in Ukrainian)
2. Kobzan S. (2019) Formation of the real estate market. Practical aspects and features of evaluation. (in Ukrainian)
3. Features of the real estate market. Available at: https://stud.com.ua/44563/ekonomika/osoblivosti_rinku_neruhomosti (in Ukrainian)
4. Kucherenko V. R., Zayets M. A., Zakharchenko O. V., Smentina N. V., Ulybina V. O. (2013) Real estate valuation and management: textbook. Odesa: Publishing house "Leradruk". (in Ukrainian)
5. Litnarovich R. M. (2011) Construction and research of mathematical model by sources of experimental data by methods of regression analysis. Rivne: Study guide, MEGU. (in Ukrainian)
6. State Statistics Service of Ukraine. Available at: <http://www.ukrstat.gov.ua>. (in Ukrainian)
7. Correlation analysis. Available at: https://pidru4niki.com/12461220/statistika/korelyatsiyiy_analiz.
8. Linear multiple regression analysis. Available at: <http://ekonometrika.ho.ua/4.htm>. (in Ukrainian)
9. Goroneskul M.M. (2009) Tables of functions and critical points of distributions. Sections: Probability theory. Mathematical statistics. Mathematical methods in psychology. Kharkiv. (in Ukrainian)
10. Residential real estate mass valuation using machine learning methods. Available at: <https://www.ukrlogos.in.ua/10.11232-2663-4139.16.50.html>. (in Ukrainian)
11. Clinton S. (2019) *Tree-based Machine Learning Algorithms: Decision Trees, Random Forests, and Boosting*.