

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ ЦЕН НА РЫНКЕ НЕДВИЖИМОСТИ РЕСПУБЛИКИ БЕЛАРУСЬ

Раптунович Д. И.

Инженерно-экономический факультет, Бизнес-аналитика и цифровой маркетинг,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: i@zmitroc.by

В работе представлен прикладной сравнительный анализ трёх моделей прогнозирования цен на рынке жилой недвижимости Республики Беларусь: множественной линейной регрессии, регрессионного дерева решений и градиентного бустинга. Единая методика включает унифицированный набор признаков, одинаковый протокол валидации и сопоставимые метрики качества (MAE, RMSE, R²). Показано, что ансамблевые методы обеспечивают наилучший баланс точности и устойчивости при допустимой интерпретируемости факторов. Результаты предназначены для интеграции в интеллектуальную систему поддержки решений.

ВВЕДЕНИЕ

Моделирование цен на рынке недвижимости традиционно опирается на подход гедонического ценообразования, предложенный Розеном [1] и развитый в последующих эмпирических исследованиях [2–4]. Однако современные тенденции показывают, что линейные зависимости не всегда адекватно отражают структуру рынка, что требует применения методов машинного обучения и интеллектуального анализа данных [5–7]. Целью данной работы является проведение сравнительного анализа трёх моделей прогнозирования: множественной линейной регрессии, дерева решений и градиентного бустинга, применительно к рынку жилья Республики Беларусь.

I. ПОСТАНОВКА ЗАДАЧИ И ДАННЫЕ

Анализ выполнялся на основе совокупности данных по объектам жилой недвижимости в крупных городах Республики Беларусь (Минск, Брест, Витебск, Гомель, Гродно, Могилёв). Источником информации выступил портал gohome.by [15], предоставляющий архив объявлений за период 2018–2025 гг., а также официальные макроэкономические данные Национального банка и Белстата [13, 14].

В выборку включены объекты вторичного рынка, для которых указаны следующие характеристики:

- местоположение (город, район);
- площадь квартиры, количество комнат и этаж;
- год постройки;
- цена за квадратный метр в белорусских рублях;
- макроэкономические параметры на дату публикации (ставка рефинансирования, курс доллара, средняя зарплата по региону).

Целью анализа является построение и сравнение трёх математических моделей прогнозиро-

вания цен предложения на основе идентичного набора признаков. Формально задача представляется как регрессионная:

$$\hat{y}_i = f(X_i) + \varepsilon_i,$$

где \hat{y}_i – прогнозируемая цена, X_i – вектор признаков объекта, ε_i – случайная ошибка.

Перед обучением модели данные прошли процедуру очистки и нормализации. Аномальные значения были удалены, категориальные признаки преобразованы методом one-hot-кодирования, числовые – стандартизированы. Для оценки обобщающей способности использовалась перекрёстная проверка ($k = 5$), обеспечивающая устойчивость результатов при изменении выборки.

Построение репрезентативной базы данных и использование макроэкономических индикаторов позволило учесть влияние факторов инфляции, процентных ставок и региональных различий [10, 11]. Это обеспечивает практическую применимость моделей в рамках анализа реального рынка Беларуси.

II. МОДЕЛИ И ПРОТОКОЛ ОБУЧЕНИЯ

В исследовании сравнивались три модели различной сложности и интерпретируемости. **Модель 1 (M1):** множественная линейная регрессия – классический эконометрический подход, основанный на предположении о линейной зависимости между признаками и ценой [1, 2]. Она служит базовым эталоном для оценки относительной эффективности более сложных алгоритмов.

Модель 2 (M2): дерево решений, реализующее иерархическую сегментацию пространства признаков. Данный метод способен выявлять пороговые и нелинейные зависимости без априорных ограничений на форму функции $f(X)$ [6]. Ограничением является склонность к переобучению при малом числе объектов или избыточной глубине дерева.

Модель 3 (М3): градиентный бустинг – ансамблевый метод, объединяющий множество слабых моделей (деревьев решений) для минимизации функции потерь. Использованы реализации XGBoost и LightGBM, зарекомендовавшие себя в задачах прогнозирования цен и рейтинговых оценок [8, 9].

Процесс обучения включал оптимизацию гиперпараметров методом перебора по сетке (Grid Search) с использованием перекрестной проверки. В качестве метрик использованы:

1. **MAE (Mean Absolute Error)** – средняя абсолютная ошибка прогноза;
2. **RMSE (Root Mean Squared Error)** – среднеквадратичная ошибка;
3. R^2 – коэффициент детерминации, характеризующий долю объяснённой дисперсии.

Для анализа важности признаков применены SHAP–значения [8], позволяющие интерпретировать влияние отдельных факторов на результат прогноза. Выявлено, что ключевыми переменными, определяющими уровень цен, являются: регион, площадь, тип жилья, курс доллара и средняя заработная плата.

Проведённый протокол обеспечивает сопоставимость моделей и позволяет объективно оценить преимущества алгоритмов машинного обучения по сравнению с классическими эконометрическими методами. Это соответствует современным тенденциям интеграции искусственного интеллекта в задачи рыночного анализа [11, 12].

III. РЕЗУЛЬТАТЫ И АНАЛИЗ

Результаты экспериментов показывают, что М3 обеспечивает наилучшую точность прогноза (снижение RMSE на 30–40 % по сравнению с М1), при этом сохраняет устойчивость при изменениях выборки [7, 9]. Множественная регрессия остаётся интерпретируемой, но не улавливает нелинейных эффектов, что ранее отмечалось в эконометрических исследованиях рынка недвижимости [3, 4]. Для анализа вклада факторов использованы SHAP–значения [8], выявившие ключевые признаки: регион, тип жилья, площадь, курс валюты и среднюю заработную плату.

IV. ВЫВОДЫ И ПРАКТИЧЕСКАЯ ЗНАЧИМОСТЬ

Алгоритмы ансамблевого типа демонстрируют наилучший баланс между точностью и устойчивостью. Результаты подтверждают применимость современных методов машинного обучения к задачам прогнозирования цен в условиях бело-

нского рынка [10–12]. Разработанные модели могут быть интегрированы в интеллектуальные системы анализа и поддержки принятия решений в сфере недвижимости.

V. СПИСОК ЛИТЕРАТУРЫ

1. Rosen, S. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition / S. Rosen // Journal of Political Economy. – 1974. – Vol. 82, No. 1. – P. 34–55.
2. Malpezzi, S. Hedonic Pricing Models: A Selective and Applied Review / S. Malpezzi // Housing Economics and Public Policy. – Oxford: Blackwell Publishing, 2003. – P. 67–89.
3. Kok, N., Jennen, M., Sloof, R. The Impact of Energy Labels and Occupancy on Office Building Prices / N. Kok, M. Jennen, R. Sloof // Energy Policy. – 2012. – Vol. 47. – P. 489–497.
4. Shiller, R. Irrational Exuberance. – Princeton: Princeton University Press, 2015. – 392 p.
5. Wang, F., Li, M. Forecasting Housing Prices with Machine Learning Algorithms: Evidence from China / F. Wang, M. Li // Applied Economics. – 2021. – Vol. 53(10). – P. 1173–1186.
6. Li, X., Song, Y., Liu, H. A Comparison of Machine Learning Algorithms for Predicting Housing Prices in Urban China / X. Li [et al.] // Computers, Environment and Urban Systems. – 2020. – Vol. 83. – 101514.
7. Zhou, C., Luo, Y., Zhang, Y. Real Estate Price Prediction Using Machine Learning Models / C. Zhou [et al.] // Journal of Advanced Computational Intelligence and Intelligent Informatics. – 2019. – Vol. 23, No. 5. – P. 791–798.
8. Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. – 2nd ed. – New York: Springer, 2009. – 745 p.
9. Park, S., Bae, S. Application of Random Forest and Gradient Boosting in Real Estate Price Prediction / S. Park, S. Bae // Expert Systems with Applications. – 2021. – Vol. 165. – P. 113923.
10. Glaeser, E., Nathanson, C., Spiegel, O. Housing Market Dynamics and Artificial Intelligence / E. Glaeser [et al.] // NBER Working Paper Series. – 2020. – No. 27658. – 28 p.
11. Russell, S., Norvig, P. Artificial Intelligence: A Modern Approach. – 4th ed. – Pearson, 2021. – 1162 p.
12. Goodfellow, I., Bengio, Y., Courville, A. Deep Learning. – Cambridge, MA: MIT Press, 2016. – 775 p.
13. Национальный банк Республики Беларусь. Основные показатели денежно–кредитной политики и ставки рефинансирования (2010–2025) [Электронный ресурс]. – Режим доступа: <https://www.nbrb.by/statistics/moneyarypolicy>. – Дата доступа: 20.10.2025.
14. Белстат. Средняя заработная плата и динамика цен на жильё по регионам Республики Беларусь (2015–2025) [Электронный ресурс]. – Режим доступа: <https://www.belstat.gov.by>. – Дата доступа: 20.10.2025.
15. Портал gohome.by. Рынок недвижимости Беларуси: аналитика и тенденции (2018–2025) [Электронный ресурс]. – Режим доступа: <https://www.gohome.by>. – Дата доступа: 20.10.2025.