

**ОБЪЯСНИМЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ЗАДАЧЕ
ПРОГНОЗИРОВАНИЯ ЦЕН НА ЖИЛЬЕ****Зитун Ихсен (КФУ)****Научный руководитель – кандидат технических наук, доцент Арабов М.К.
(КФУ)**

Введение. Искусственный интеллект (ИИ) революционизирует различные отрасли, включая недвижимость, где он используется для прогнозирования цен. Однако по мере усложнения моделей ИИ становится все более важным обеспечивать их прозрачность и интерпретируемость. Объяснимый ИИ (ХАИ) решает эту проблему, предоставляя информацию о том, как модели делают прогнозы, что способствует доверию со стороны пользователей. В данной статье рассматривается роль ХАИ в прогнозировании цен на жилье с использованием российского набора данных о недвижимости (2018–2021) с Kaggle [1], демонстрируя, как модели машинного обучения могут быть одновременно точными и интерпретируемыми.

Основная часть. Набор данных включает ключевые атрибуты, такие как местоположение недвижимости, тип, размер, цена сделки и время продажи, что делает его подходящим для предсказательного моделирования [1]. Для разработки надежной модели прогнозирования мы используем алгоритмы градиентного бустинга, такие как XGBoost и LightGBM, обеспечивающие высокую точность в задачах регрессии [2]. Этап предобработки данных включает обработку пропущенных значений, нормализацию числовых признаков и кодирование категориальных переменных, а инженерия признаков улучшает производительность модели за счет добавления производных характеристик, таких как цена за квадратный метр и группировка по местоположению.

Для обеспечения прозрачности применяются методы объяснимого ИИ (ХАИ) для интерпретации предсказаний модели. SHAP (SHapley Additive exPlanations) используется для оценки значимости признаков, выявляя влияние таких факторов, как местоположение и размер недвижимости, на прогнозируемые цены [3]. Метод LIME (Local Interpretable Model-agnostic Explanations) предоставляет локальную интерпретацию, объясняя отдельные предсказания [4], а графики частичной зависимости (PDP) иллюстрируют, как изменения конкретных признаков влияют на стоимость жилья [5]. Дополнительно встроенные инструменты оценки значимости признаков в деревьях решений, таких как XGBoost, помогают выявить ключевые факторы ценообразования [2,6].

Интеграция этих методов позволяет получить ценные аналитические выводы о рыночных трендах. Выявляются временные закономерности, демонстрирующие влияние экономических циклов на цены на жилье. Эффект местоположения показывает различия в стоимости недвижимости в зависимости от инфраструктуры и региональной политики. Понимание этих факторов способствует принятию более обоснованных решений для покупателей, продавцов, государственных регуляторов и специалистов по недвижимости.

Выводы. Объяснимый искусственный интеллект (ХАИ) обеспечивает то, что предсказательные модели в сфере недвижимости не являются просто «черными ящиками», а представляют собой ценные инструменты, сочетающие точность и интерпретируемость. Применяя ХАИ-методы к прогнозированию цен на жилье, заинтересованные стороны могут принимать обоснованные решения на основе прозрачных аналитических данных. Данное исследование демонстрирует потенциал объяснимых моделей в сфере недвижимости, способствуя укреплению доверия и внедрению стратегий, основанных на данных. Дальнейшие исследования, такие как работа Чжана Юйи по объясняемому ИИ в прогнозировании цен на недвижимость [7], подчеркивают важность разработки интерпретируемых моделей, повышающих эффективность рынка.

Список использованных источников:

1. Kaggle (n.d.). World's Real Estate Data (147k) // Dataset link: <https://www.kaggle.com/datasets/toriqulstu/worlds-real-estate-data147k>.
2. Zitoun I., Arabov M.K. Comparative Analysis of Ensemble and Linear Machine Learning Models in the Task of House Price Prediction // Proceedings of the 2024 International Russian Automation Conference (RusAutoCon), Sochi, Russian Federation. – 2024. – C. 50–55. – DOI: 10.1109/RusAutoCon.2024.00011.
3. Lundberg S.M., Lee S.-I. A Unified Approach to Interpreting Model Predictions // Advances in Neural Information Processing Systems (NeurIPS). – 2017. – № 30. – C. 4765–4774.
4. Ribeiro M.T., Singh S., Guestrin C. "Why Should I Trust You?" Explaining the Predictions of Any Classifier // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – 2016. – C. 1135–1144. – DOI: 10.1145/2939672.2939778.
5. Friedman J.H. Greedy Function Approximation: A Gradient Boosting Machine // Annals of Statistics. – 2001. – № 29(5). – C. 1189–1232. – DOI: 10.1214/aos/1013203451.
6. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – 2016. – C. 785–794. – DOI: 10.1145/2939672.2939785.
7. Zhang Yuyi. Explainable Artificial Intelligence Methods for Real Estate Price Prediction // PhD Dissertation, St. Petersburg State University. – 2024. – C. 1–120.