

## Прогнозирование потребительского спроса и стоимости объектов недвижимости на основе данных ЦИАН и поведенческой аналитики пользователей

Андрющенко В.В., Загуменная В.В., Лазаренко Е.Ю.

Научный руководитель – Федоров Д.А.

Национальный исследовательский университет ИТМО

[408154@niuitmo.ru](mailto:408154@niuitmo.ru), [vvzagumennaya@mail.ru](mailto:vvzagumennaya@mail.ru), [lazarenkolizka@yandex.ru](mailto:lazarenkolizka@yandex.ru)

### Введение

Цифровизация экономики сопровождается активным развитием интеллектуальных сервисов анализа данных, основанных на методах машинного обучения. Рынок недвижимости представляет собой сложную социально-экономическую систему, характеризующуюся значительным количеством факторов, влияющих на формирование стоимости объектов.

Современные онлайн-платформы недвижимости аккумулируют большие массивы информации о характеристиках объектов и активности пользователей. Помимо структурированных данных о параметрах недвижимости такие платформы содержат поведенческие данные пользователей, включая просмотры объявлений, добавление объектов в избранное и динамику интереса к различным районам. Анализ подобных данных позволяет выявлять текущие тенденции потребительского спроса и использовать их для прогнозирования рынка недвижимости [6].

### Основная часть

В работе рассматривается подход к прогнозированию стоимости объектов жилой недвижимости на основе данных платформы ЦИАН. Используемый набор данных включает характеристики объектов недвижимости, такие как площадь квартиры, этаж, тип жилья, год постройки здания, транспортная доступность и другие параметры, влияющие на рыночную стоимость объектов.

Методологической основой исследования выступает теория гедонического ценообразования, согласно которой стоимость недвижимости определяется совокупностью её характеристик [2]. Однако современные исследования показывают, что методы машинного обучения позволяют более эффективно выявлять сложные взаимосвязи между признаками и ценой объекта [3].

Для решения задачи прогнозирования стоимости недвижимости применяются алгоритмы машинного обучения, включая множественную линейную регрессию, случайный лес (Random Forest) и градиентный бустинг (Gradient Boosting). Данные алгоритмы широко используются в задачах интеллектуального анализа данных и демонстрируют высокую точность при обработке больших наборов разнородных признаков.

На этапе подготовки данных проводится предварительная обработка информации, включающая очистку данных от выбросов, нормализацию числовых признаков и кодирование категориальных переменных. Дополнительно выполняется извлечение признаков из текстовых описаний объявлений, что позволяет учитывать дополнительные характеристики объектов недвижимости. Особое внимание уделяется использованию поведенческих показателей пользователей цифровой платформы. Показатели активности пользователей, такие как количество просмотров объявлений и добавление объектов в избранное, могут рассматриваться как индикаторы интереса аудитории к объектам недвижимости. Интеграция таких данных в модели машинного

обучения позволяет учитывать текущие тенденции потребительского спроса и повышает точность прогнозирования стоимости объектов.

Научная новизна исследования заключается в интеграции поведенческих метрик пользователей цифровых платформ с традиционными характеристиками объектов недвижимости в рамках единой модели машинного обучения. Предложенный подход позволяет учитывать цифровые индикаторы интереса пользователей и расширяет возможности интеллектуальных сервисов анализа рынка недвижимости.

### **Выводы**

Применение методов машинного обучения и поведенческой аналитики пользователей позволяет повысить точность прогнозирования стоимости объектов недвижимости и анализа потребительского спроса. Интеграция структурированных данных объявлений и показателей пользовательской активности обеспечивает более полное представление о текущем состоянии рынка.

Разработанный подход может быть использован при создании интеллектуальных аналитических сервисов для онлайн-платформ недвижимости, а также в системах поддержки принятия решений для девелоперов и инвесторов.

### **Литература**

1. Naan M., Hewitt C., Chuatico G. By the numbers: the construction industry in Canada from 1986 to 2016 // *Labour and Industry*. – 2020. – No 30(4). – P. 299–320.
2. Власенко Т. В., Белокобыленко С. В. Рынок недвижимости и его структура // *Экономика и экология территориальных образований*. – 2015. – № 1. – С. 50–53.
3. Abdallah S. An intelligent system for identifying influential words in real-estate classifieds // *Journal of Intelligent Systems*. – 2018. – No 27(2). – P. 183–194.
4. Eickmeier S., Hofmann B. Monetary policy, housing booms and financial (im)balances // *Macroeconomic Dynamics*. – 2013. – No 17(4). – P. 830–860.
5. Ahmed I., Soggi C., Severini F., Pretaroli R., Al Mahdi H. K. Unconventional monetary policy and real estate sector: a financial dynamic computable general equilibrium model for Italy // *Economic Systems Research*. – 2020. – No 32(2). – P. 221–238.
6. Slemrod J., Weber C., Shan H. The behavioral response to housing transfer taxes: evidence from a notched change in D.C. policy // *Journal of Urban Economics*. – 2017. – No 100. – P. 137–153.
7. Cocconcelli L., Medda F. R. Boom and bust in the Estonian real estate market and the role of land tax as a buffer // *Land Use Policy*. – 2013. – No 30(1). – P. 392–400.
8. Stevens D. Predicting Real Estate Price Using Text Mining Automated Real Estate Description Analysis: дис. ... канд. комп. наук. – Tilburg, 2014. – 70 p.
9. Abdallah S., Khashan D. A. Using text mining to analyze real estate classifieds. – Springer. – 2015. – P. 193–202.