

Шихов Андрей Сергеевич, магистрант
ФГБОУ ВО «Вятский государственный университет»

ИСПОЛЬЗОВАНИЕ МАКРОЭКОНОМИЧЕСКИХ ПАРАМЕТРОВ В ЗАДАЧЕ ПРОГНОЗИРОВАНИЯ СТОИМОСТИ ЖИЛОЙ НЕДВИЖИМОСТИ

Аннотация. В статье рассматривается использование макроэкономических параметров в обучающем наборе данных для повышения точности прогнозирования стоимости жилой недвижимости.

Ключевые слова: Машинное обучение, прогнозирование стоимости недвижимости, алгоритмы прогнозирования, методы оценки стоимости, математические модели.

Оценка и прогнозирование цен на недвижимость играют важную роль в экономике и финансах, затрагивая различные сферы – от инвестиционных решений до формирования государственной жилищной стратегии. Цена объекта недвижимости складывается под влиянием множества факторов, включая географическое положение, развитость инфраструктуры, экономические условия региона и демографические тенденции. Умение прогнозировать стоимость объектов позволяет не только анализировать текущую рыночную конъюнктуру, но и предвидеть ее изменения, что крайне важно для частных вкладчиков, коммерческих структур и органов власти при разработке эффективных управленческих решений.

Машинное обучение активно применяется для оценки недвижимости, такие модели имеют существенные недостатки [1]. Главная проблема – чувствительность к качеству данных: пропуски, выбросы и субъективные описания объектов снижают достоверность прогнозов. Кроме того, модели, разработанные для одной страны, часто не учитывают специфику экономики других государств, что ограничивает их применение [2]. Еще один недостаток – игнорирование пространственных закономерностей, например, локальных ценовых трендов в отдельных районах, которые крайне важны для недвижимости. Наконец, из-за высокой волатильности рынка модели быстро устаревают и требуют постоянного обновления данных для сохранения точности прогнозирования [3].

Формирование стоимости недвижимости зависит от множества факторов, включая макроэкономические условия, географические факторы и индивидуальные характеристики жилой единицы [4]. Среди ключевых макроэкономических параметров – уровень инфляции, ставки по ипотечным кредитам, курс доллара, цена экспортируемой марки нефти [5]. Например, низкие процентные ставки стимулируют спрос, повышая цены, а экономический кризис, напротив, может привести к их снижению. К географическим факторам можно отнести расстояние до центра города или до метро. Основными характеристиками единицы жилой недвижимости являются: общая площадь, жилая площадь, количество комнат, этаж и наличие ремонта.

Для оценки влияния использования макроэкономических параметров в обучающем наборе данных на точность прогнозирования стоимости жилой недвижимости был проведен эксперимент, в котором модель машинного обучения Catboost была обучена на двух наборах данных. Первый набор данных содержит в себе только физические и географические характеристики объекта жилой недвижимости, а второй набор данных содержит не только физические и географические характеристики объекта жилой недвижимости, но и такие макроэкономические показатели: процентная ставка по вкладам и ипотеке, курс доллара и курс нефти. Оба набора данных содержат в себе 4995 строк об объектах жилой недвижимости.

CatBoost – алгоритм градиентного бустинга [6], специально разработанный для работы с категориальными признаками. Он автоматически обрабатывает категориальные переменные без необходимости их предварительного кодирования, что упрощает подготовку данных. CatBoost также использует упорядоченный boosting для борьбы с переобучением и улучшения обобщающей способности модели. Преимуществами данного алгоритма являются:



эффективная работа с категориальными признаками, автоматическая обработка пропущенных значений, устойчивость к переобучению благодаря упорядоченному бустингу, высокая точность на задачах с разнородными данными.

В качестве сравнительных метрик использовались коэффициент детерминации [7] и средняя абсолютная ошибка [8]. R-квадрат – коэффициент детерминации. Этот показатель оценивает долю дисперсии зависимой переменной, которая может быть объяснена моделью. Высокое значение R-квадрата указывает на то, что модель хорошо объясняет вариацию цен на недвижимость. MAE (Mean Absolute Error) – средняя абсолютная ошибка. Этот показатель измеряет среднее абсолютное отклонение между прогнозируемыми и фактическими ценами на недвижимость. Он позволяет оценить точность модели в прогнозировании цен.

В результате эксперимента были получены следующие результаты. У модели, обученной на наборе данных только с жилищными характеристиками на тренировочной выборке коэффициент детерминации составил 0,882, а значение средней абсолютной ошибки составило 393210, а на тестовой коэффициент детерминации составил 0,727, а значение средней абсолютной ошибки составило 545230. Модель, обученная на наборе данных с дополнительными макроэкономическими параметрами, показала лучшие результаты. Коэффициент детерминации на тренировочной выборке составил 0,895, а значение средней абсолютной ошибки 360100. На тестовой выборке коэффициент детерминации составил 0,793, а значение средней абсолютной ошибки составило 502944.

Сравнение моделей показало, что включение макроэкономических параметров существенно улучшает качество прогнозирования стоимости жилой недвижимости. Модель, учитывающая как жилищные, так и макроэкономические показатели, достигла лучших результатов, что свидетельствует о более высокой объясняющей способности. Таким образом, макроэкономические факторы оказывают статистически значимое влияние на ценообразование в данном сегменте рынка. Их учет особенно важен для среднесрочных и долгосрочных прогнозов, где изменения экономической ситуации могут существенно менять динамику цен.

В заключение можно отметить, что из-за постоянно меняющегося рынка недвижимости и экономической ситуации, данный тип моделей быстро теряет актуальность. Решением данной проблемы является своевременное обновление обучающего набора данных и переобучение модели на обновленном наборе данных. Проведенное исследование подтвердило, что добавление макроэкономических признаков в задачи с прогнозированием повышает точность оценки стоимости недвижимости. Результаты работы могут быть применены в системах оценки стоимости.

Список литературы:

1. Методы машинного обучения в исследовании рынка жилой недвижимости. – URL: <https://cyberleninka.ru/article/n/metody-mashinnogo-obucheniya-v-issledovanii-rynka-zhiloy-nedvizhimosti/viewer> – электронный.
2. Стерник, Г. М. Развитие и совершенствование методов прогнозирования на рынке жилой недвижимости / Г. М. Стерник, С. Г. Стерник, А. В. Свиридов // Урбанистика и рынок недвижимости. – 2014. – № 1. – С. 53–93.
3. Стерник Г.М., Стерник С.Г. Пять макроэкономических законов функционирования рынка недвижимости как неотъемлемой составляющей глобального финансового рынка в транзитивной экономике // Финансовая аналитика: проблемы и решения. – 2010. – № 11 (35). – С. 15–29.
4. Массовая оценка объектов недвижимости на основе технологий машинного обучения. Анализ точности различных методов на примере определения рыночной стоимости квартир. – URL: <https://cyberleninka.ru/article/n/massovaya-otsenka-obektov-nedvizhimosti-na-osnove-tehnologiy-mashinnogo-obucheniya-analiz-tochnosti-razlichnyh-metodov-na-primere/viewer> – Текст: электронный. – (Дата обращения 18.03.2025)



5. Статистический анализ динамики стоимости квартир на вторичном рынке жилой недвижимости города Москвы. – URL: <https://cyberleninka.ru/article/n/statisticheskiy-analiz-dinamiki-stoimosti-kvartirna-vtorichnom-rynke-zhiloy-nedvizhimosti-goroda-moskvy/viewer> – Текст: электронный. – (Дата обращения 16.03.2025)

6. Методы машинного обучения в исследовании рынка жилой недвижимости. – URL: <https://cyberleninka.ru/article/n/metody-mashinnogo-obucheniya-v-issledovanii-rynka-zhiloy-nedvizhimosti/viewer> – электронный.

7. R2_score – Gallery examples: L1-based models for Sparse Signals Linear Regression Example Non-negative least squares Failure of Machine Learning to infer causal effects Effect of transforming the targets in r... – URL: https://scikitlearn/stable/modules/generated/sklearn.metrics.r2_score.html – Текст: электронный. (Дата обращения 12.03.2025)

8. Mean_squared_error – Gallery examples: Early stopping in Gradient Boosting Gradient Boosting regression Prediction Intervals for Gradient Boosting Regression Model Complexity Influence Linear Regression Example Poisson... – URL: https://scikit-learn/stable/modules/generated/sklearn.metrics.mean_squared_error.html – Текст: электронный. (Дата обращения 15.03.2025)

