

**ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ФАКТОРОВ ЦЕНООБРАЗОВАНИЯ В
МОДЕЛЯХ МАШИННОГО ОБУЧЕНИЯ****Берзин Д.В. (ИТМО)****Научный руководитель – кандидат социологических наук, Чупин Р.И.(ИТМО)**

Введение. Ценообразование является одним из важнейших компонентов любой бизнес модели. Его характерной чертой является то, что из-за конкуренции данный процесс не имеет конечной точки и нуждается в постоянном совершенствовании. В связи с этим, необходимость оптимизации процесса ценообразования является очевидной. В настоящее время методы машинного обучения помогают эффективно изучать зависимости, влияющие на ценообразование. Проблему интерпретации полученных таким образом результатов можно решить при помощи вектора Шепли, суть применения которого состоит в том, что позволяет измерить количественный вклад каждого фактора в итоговый результат. Использование вектора Шепли уже широко применяется при интерпретировании результатов, выдаваемых моделями машинного обучения, однако данный метод не является идеальным и в случае, когда функция коррелирует с предсказанием модели, использование вектора Шепли демонстрирует ложный высокий уровень важности фактора, который может не иметь никакого отношения к функционированию модели [1].

Основная часть. Оптимизировать процесс ценообразования предлагается при помощи реализации системы принятия решений, основанной на ML-моделировании и имитационной модели. Для достижения данной цели, предполагается адаптивное интуитивное моделирование теории игр.

Алгоритм моделирования устойчивости ценообразования имеет следующие этапы:

- 1) Построение количественной модели, которая будет основана на применении ML-моделирования и вектора Шепли. В результате данного этапа будет получена матрица значимости факторов ценообразования.
- 2) Оценка эволюционно устойчивого состояния, где главным методом выступает применение эволюционной теории игр. Результатом данного этапа будет определение эволюционно устойчивого состояния факторов.
- 3) Имитационное моделирование, цель которого определение вариантов устойчивости.

В общем виде алгоритм выглядит следующим образом. Имитационная модель предназначена для нахождения устойчивых точек покоя системы дифференциальных уравнений, описывающих динамику системы ценообразования. В данном приближении используется модель бесконечного набора факторов ценообразования, которые взаимодействуют друг с другом в процессе подготовки данных и обучения модели, через методы класса ColumnHandler. Каждый метод класса ColumnHandler представляет собой операцию агрегации данных по определенным категориям или условиям. В коде используется экстремальный бустинг для построения модели прогнозирования. При создании экземпляра модели, указываются оптимальные гиперпараметры, найденные с помощью решетчатого поиска (GridSearchCV). Затем модель обучается на обучающих данных с помощью метода `fit()`, а затем используется для предсказания на тестовых данных с помощью метода `predict()`. После этого оценивается качество модели с помощью метрик, таких как коэффициент детерминации (R^2 score), среднеквадратичная ошибка (MSE) и RMSE. В коде используются следующие коэффициенты для экстремального бустинга: `learning_rate`, `max_depth`, `n_estimators` и `random_state`. Для повышения устойчивости и точности, к модели был применен метод SSMS (Selecting Stable Model SHAP), позволяющий разрешить проблему ошибочной интерпретации важности фактора [2]. В качестве имитационной модели ЭБР выбрана базовая EGT.

Выводы. Реализована модель прогнозирования. К данной модели был применен метод

SSMS, после чего произведена оценка эффективности модели прогнозирования с применением указанного метода и без его использования.

Список использованных источников:

1. Merrick L., Taly A. The Explanation Game: Explaining Machine Learning Models Using Shapley Values // Fiddler Labs, Palo Alto, USA. – 2020. arXiv:1909.08128. DOI10.1007/978-3-030-57321-8_2.
2. Воробьев А.В. Метод выбора модели машинного обучения на основе устойчивости предикторов с применением значения Шепли // Экономика. Информатика. – 2021. - №48 (2). - С. 350-359.