

Галкин Илья Николаевич

магистр, аспирант

ФГБОУ ВО «Санкт-Петербургский государственный
экономический университет»

г. Санкт-Петербург

МЕТОДЫ АНАЛИЗА И ИНТЕРПРЕТАЦИИ ФИНАНСОВЫХ СИСТЕМ ПРИ ПОМОЩИ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

Аннотация: в статье проведено сравнение применимости эконометрических методов и методов машинного обучения для решения задач финансово-экономического анализа и прогнозирования, выделены преимущества и недостатки обоих классов методов, поставлена проблематика интерпретируемости моделей машинного обучения, которая является ограничением на пути их широкого применения в финансово-аналитическом инструментарии, и рассмотрены существующие решения в этой области. По итогам исследования сделан вывод о необходимости использования сразу нескольких методов интерпретации машинного обучения с целью перекрестной валидации результатов.

Ключевые слова: машинное обучение, финансово-экономическое прогнозирование, векторы Шепли, эконометрические методы.

Задачи анализа и прогнозирования финансово-экономических систем возникают достаточно часто в различных разделах экономической науки. К примерам таких задач можно отнести прогнозирование цен активов на финансовых рынках, расчет и анализ стоимости бизнеса компании, анализ и прогнозирование взаимосвязи макроэкономических показателей. Для решения таких задач исследователи и практики используют различные численные методы.

Традиционно такие задачи решаются при помощи эконометрических методов [7]. К преимуществам этих методов можно отнести их строгую математическую обоснованность, возможность учета при моделировании различных факторов, а также их прямую и прозрачную интерпретируемость. Последний пункт является крайне важным для исследователей и практиков, работающих с анализом

и прогнозированием финансовых систем, поскольку он позволяет в явном виде оцифровать и выделить отдельное влияние (или же его отсутствие) факторов модели на результирующую (объясняемую) переменную.

Тем не менее, в последние несколько десятилетий получили свое развитие и распространение методы машинного обучения и искусственного интеллекта, которые на данный момент широко используются в различных научных областях и отраслях народного хозяйства. Методы машинного обучения и искусственного интеллекта, как и эконометрические методы, решают одинаковую задачу – прогнозирование определенной величины с использованием информации о взаимосвязанных с ней величинах (факторах). Тем не менее, методы машинного обучения, по сравнению с эконометрическими методами, считаются более точными и позволяют работать с существенно большим объемом данных, но при этом обладают более низкой интерпретируемостью [2].

Часто об алгоритмах машинного обучения говорят как о «черном ящике» (black box), то есть как о системах с неизвестным механизмом работы. В то же время эконометрические методы, наоборот, позволяют математически точно и однозначно вывести формульные зависимости между факторами модели и исследуемой величиной. Таким образом, несмотря на преимущества методов машинного обучения, их применение сильно ограничено в финансовой сфере, поскольку теоретикам и практикам, решающим финансово-экономические задачи, крайне важна интерпретируемость моделей, то есть прозрачное отслеживание и объяснение результатов их работы. В то же самое время, с учетом возрастающего количества данных, а также повышения требований к точности и своевременности финансового анализа, применение методов машинного обучения в сфере финансового-экономического прогнозирования выглядит очень перспективным.

В связи с этим видится целесообразным рассмотреть некоторые существующие теоретические и практические наработки в области интерпретируемости моделей машинного обучения. Одним из методов интерпретации моделей машинного обучения является библиотека языка программирования python SHAP –

данная библиотека с применением подхода теории кооперативных игр – векторов Шепли – позволяет интерпретировать результаты моделей машинного обучения как на уровне отдельного наблюдения, так и на уровне всей модели [10]. Этот подход основан на том, что объясняющие факторы модели являются кооперативными «игроками», распределяющими между собой «результат» (в данном случае – прогноз) неким справедливым образом. В области прогнозирования цен акций компаний уже существуют наработки, использующие это решение для интерпретации прогнозных моделей [3–4].

Другим популярным решением в области интерпретации моделей машинного обучения является библиотека ELI5, которая позволяет исследовать веса влияющих факторов и получившиеся прогнозные значения [8]. Схожим образом работает и библиотека LIME [9]. Кроме того, исследователи также выделяют метод перестановки признаков, выполняемый следующим образом: значения одного из факторов модели перемешиваются случайным образом, и на основе этих данных обучается новая модель, в таком случае признак признается важным (значимым), если после перестановки качество модели сильно ухудшилось [1].

Указанные выше способы интерпретации моделей машинного обучения имеют свои недостатки, главный из которых заключается как раз-таки в том, что не существует единого стандарта интерпретации моделей. Так, применительно к одной и той же модели, построенной на одном и том же наборе данных, различные методы интерпретации могут давать различные результаты [6]. В таком случае можно рекомендовать исследователям и практикам прибегать к нескольким методам интерпретации моделей машинного обучения, что обеспечит кросс-валидацию результатов и повысит доверие к результатам моделирования.

Таким образом, несмотря на наличие методов интерпретации моделей машинного обучения, которые облегчают применение современных интеллектуальных моделей для анализа и прогнозирования финансовых систем, интерпретация таких моделей все еще видится затруднительной задачей, не имеющей однозначного решения. В связи с тем мы считаем актуальными дальнейшие теоре-

тические и практические исследования в области интерпретации моделей машинного обучения, в том числе применительно к финансово-экономическим прогнозным задачам.

Список литературы

1. Carvalho, D.V., Pereira, E.M., Cardoso, J.S. Machine Learning Interpretability: A Survey on Methods and Metrics // *Electronics*. – 2019. – P. 8.

2. Perez E., Domínguez J., Omatu S., Herrera-Viedma E., Corchado Rodríguez J. Machine Learning and Traditional Econometric Models: A Systematic Mapping Study // *Journal of Artificial Intelligence and Soft Computing Research*. – 2021. – pp. 79–100.

3. Галкин И.Н. Потенциал применения методов искусственного интеллекта для анализа долгосрочной динамики стоимости собственного капитала публичных компаний / И.Н. Галкин // *Повышение конкурентоспособности отечественной науки: развитие в условиях мировой нестабильности: материалы научной конференции аспирантов СПбГЭУ (Санкт-Петербург, 18 мая 2023 г.)*. – СПб.: Изд-во СПбГЭУ, 2023. – С. 9–15. – EDN NSXHQU

4. Коклев П.С. Оценка стоимости компании с использованием методов машинного обучения / П.С. Коклев // *Финансы: теория и практика*. – 2022. – №5. – С. 132–148. DOI 10.26794/2587-5671-2022-26-5-132-148. EDN ALHXXW

5. Неустроев Д.Д. Интерпретируемость моделей машинного обучения / Д.Д. Неустроев, Д.И. Курманова // *Язык в сфере профессиональной коммуникации: сб. матер. междунаrod. науч.-практ. конф. преподавателей, аспирантов и студентов (Екатеринбург, 18 апреля 2019 г.)*. – Екатеринбург: Ажур, 2019. – С. 484–488.

6. Перминов Н.К. Интерпретация результатов машинного обучения для задачи регрессии / Н.К. Перминов // *Информатика: проблемы, методы, технологии: материалы XXII Междунаrod. науч.-практ. конф. им. Э. К. Алгазинова, 2022*. – С. 1185–1196. EDN RAFERA

7. Суханова О.Н. Эконометрические модели как инструмент анализа в управлении экономическими системами / О.Н. Суханова, О.В. Ментюкова // Модели, системы, сети в экономике, технике, природе и обществе. – 2016. – №1 (17). – С. 125–134. EDN VTYLSV

8. Техническая документация библиотеки ELI5 [Электронный ресурс]. – Режим доступа: <https://eli5.readthedocs.io/en/latest/> (дата обращения: 22.12.2023).

9. Техническая документация библиотеки [Электронный ресурс]. – Режим доступа: <https://github.com/marcotcr/lime> (дата обращения: 22.12.2023).

10. Техническая документация библиотеки SHAP [Электронный ресурс]. – Режим доступа: <https://shap.readthedocs.io/en/latest/> (дата обращения: 22.12.2023).