

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ЭКСПОРТА КАКАО-ПРОДУКТОВ

Проведён сравнительный анализ LASSO и XGBoost для прогнозирования экспорта какао-продуктов Республики Беларусь. Выявлен структурный сдвиг с 2022 года. LASSO, обеспечивает более точный прогноз, благодаря автоматическому отбору признаков.

Резкий рост мировых цен на какао-бобы весной 2025 года (свыше 10 тыс. долл. США за тонну) оказывает прямое влияние на стоимость шоколадной продукции [1]. Несмотря на отсутствие собственного производства какао-бобов, Республика Беларусь имеет место в международной торговле переработанными какао-продуктами. Цель работы – сравнительный анализ методов LASSO и XGBoost для прогнозирования стоимости экспорта какао-продуктов и выявление факторов, определяющих точность прогноза.

Исходные данные – месячные показатели экспорта товарной позиции «Шоколад и прочие готовые пищевые продукты, содержащие какао» (код ТН ВЭД 1806) за январь 2010 – декабрь 2024 г. (180 наблюдений) с сайта Национального статистического комитета РБ [2]. Для применения машинного обучения ряд преобразован в формат «признак-цель»: лаговые значения (lag_1 , lag_3 , lag_6 , lag_{12}), скользящие средние (ma_3 , ma_6 , ma_{12}), временные признаки и фиктивные переменные для структурного сдвига. Горизонт прогноза – 1 месяц.

Тест Чоу подтвердил структурный сдвиг с 2022 года ($F=22,87$, $p=0,00$). Оценка качества выполнена на тестовой выборке (2021–2024 гг., 48 наблюдений). Модель LASSO показала $MAPE = 8,42\%$, $MAE = 443,65$ тыс. долл., $RMSE = 722,23$ тыс. долл. XGBoost показал $MAPE = 19,77\%$, $MAE = 1362,20$ тыс. долл., $RMSE = 1702,65$ тыс. долл.



Рис. 1 – Динамика экспорта и прогнозы моделей LASSO и XGBoost (2021–2024 гг.)

Визуальный анализ подтверждает, что LASSO точно воспроизводит сезонные колебания и тенденцию снижения после 2022 года. Анализ коэффициентов LASSO выявил доминирующую роль ma_3 (1322,62). SHAP-анализ для XGBoost подтвердил наибольший вклад ma_3 (1382,11 тыс. долл.), lag_{12} (290,77 тыс. долл.) и lag_6 (156,93 тыс. долл.) [3].

Преимущество LASSO обусловлено L1-регуляризацией, автоматически обнуляющей неинформативные признаки, устойчивостью к мультиколлинеарности и отсутствием переобучения на стабильный период 2010–2020 гг [4]. Двухэтапная оптимизация XGBoost снизила $MAPE$ лишь до 18,67%, что не превзошло LASSO. В условиях структурного сдвига LASSO показал более высокую точность благодаря механизму автоматического отбора признаков, обеспечивающему устойчивость к изменению структуры данных. Результаты могут быть использованы при разработке инструментов поддержки управленческих решений в сфере планирования экспортной деятельности белорусских предприятий.

Список литературы

1. РБК Компании [Электронный ресурс]. – Режим доступа: <https://companies.rbc.ru/news/wqIBjUn1fb/mirovoj-ryinok-kakao-sostoyanie-tsenyi-i-trendyi-2025-goda/>. – Дата доступа: 13.03.2026.
2. Национальный статистический комитет Республики Беларусь [Электронный ресурс]. – Режим доступа: <https://www.belstat.gov.by/> – Дата доступа: 13.03.2026.
3. Сурина, А. В. Анализ временных рядов: учеб. пособие / А. В. Сурина. – СПб. : Питер Ком, 2025. – 90 с.
4. Разин, Н. Применение методов машинного обучения в задаче прогнозирования региональной инфляции / Н. Разин [и др.]. – М. : Банк России, 2025. – 38 с.

Татарчук Вероника Алексеевна, студент Инженерно-экономического факультета Белорусского государственного университета информатики и радиоэлектроники, veronik4578@gmail.com.

Научный руководитель: Мозоль Алеся Александровна, кандидат экономических наук, доцент кафедры экономической информатики Белорусского государственного университета информатики и радиоэлектроники, a.mozol@bsuir.by.