

Команда 87

Предсказание стоимости акций

5 чекпоинт

Состав команды:
Станислав Тюлягин

Куратор:
Мария Кофанова

Задача:

Разработка и тестирование различных нелинейных моделей для предсказания цен акций

Данные:

- Источник: Финансовые данные (CLOSE, TICKER,TRADEDATE)
- Отобраны акции первого эшелона.
- Данные предобработаны:
 - Фильтрация по тикерам.
 - Преобразование дат.
 - Удаление пропущенных значений.

Используемые модели

- **Exponential Smoothing** – сглаживание временных рядов.
- **Random Forest** – ансамблевый метод на основе деревьев решений.
- **Gradient Boosting** – метод последовательного обучения.
- **XGBoost** – улучшенный градиентный бустинг.
- **LGBM** – быстрый и эффективный бустинг.
- **Decision Tree** – одиночное дерево решений.
- **k-NN** – метод ближайших соседей.
- **Ансамбль моделей** – комбинация Exponential Smoothing, Random Forest, Gradient Boosting и XGBoost.

Метрики оценки

Основная:

- **MAE (Mean Absolute Error)** – средняя абсолютная ошибка.

Второстепенные:

- **MSE (Mean Squared Error)** – среднеквадратичная ошибка.
- **RMSE (Root Mean Squared Error)** – корень из среднеквадратичной ошибки.
- **MAPE (Mean Absolute Percentage Error)** – средняя абсолютная процентная ошибка.
- **R² (R-squared)** – коэффициент детерминации.

Результаты моделей

Model	MSE	MAE	RMSE	MAPE	R ²
Decision Tree	13356.55	39.75	52.59	0.0199	0.3932
Ensemble	6438.44	27.93	36.90	0.0152	0.6525
Exponential Smoothing	6142.96	27.29	35.40	0.0146	0.6235
Gradient Boosting	7327.06	30.49	40.83	0.0174	0.5523
LGBM	8035.96	33.18	44.51	0.0218	0.4149
Random Forest	6907.29	29.26	38.40	0.0165	0.6357
XGBoost	8567.84	31.50	42.18	0.0171	0.5694
k-NN	12135.87	39.03	50.35	0.0230	0.2408

Результаты моделей

- Exponential Smoothing показывает наименьшие ошибки, но уступает по R^2 .
- Random Forest, Gradient Boosting, XGBoost и LGBM демонстрируют хорошие результаты.
- Decision Tree и k-NN показали низкую предсказательную способность.
- Ансамбль моделей достигает лучшего баланса между ошибками и точностью.

Итоговый выбор модели

- Ансамбль моделей продемонстрировал **лучший R^2 (0.6525)** и **второе место по ошибкам**.
- Exponential Smoothing минимизирует ошибки, но хуже объясняет вариации данных.
- Простые модели (Decision Tree, k-NN) не рекомендованы к использованию.

Выводы

- Использование ансамбля моделей позволяет повысить точность прогнозов.
- Exponential Smoothing можно применять как вспомогательный или отдельный инструмент в условиях ограничения ресурсов.
- Для долгосрочного прогнозирования целесообразно строить ансамбли и гибридные модели.

Спасибо за Ваше время и внимание!