

Разработка стратегии автоследования

Чекпоинт №5

Модель машинного обучения

- На прошлом чекпоинте была ошибка, из-за которой происходила утечка данных из целевой переменной. Поэтому первым делом эта ошибка была исправлена
- После исправления ошибки метрики качества модели машинного обучения значительно упали, настолько, что качество модели незначительно превышает качество константного прогноза

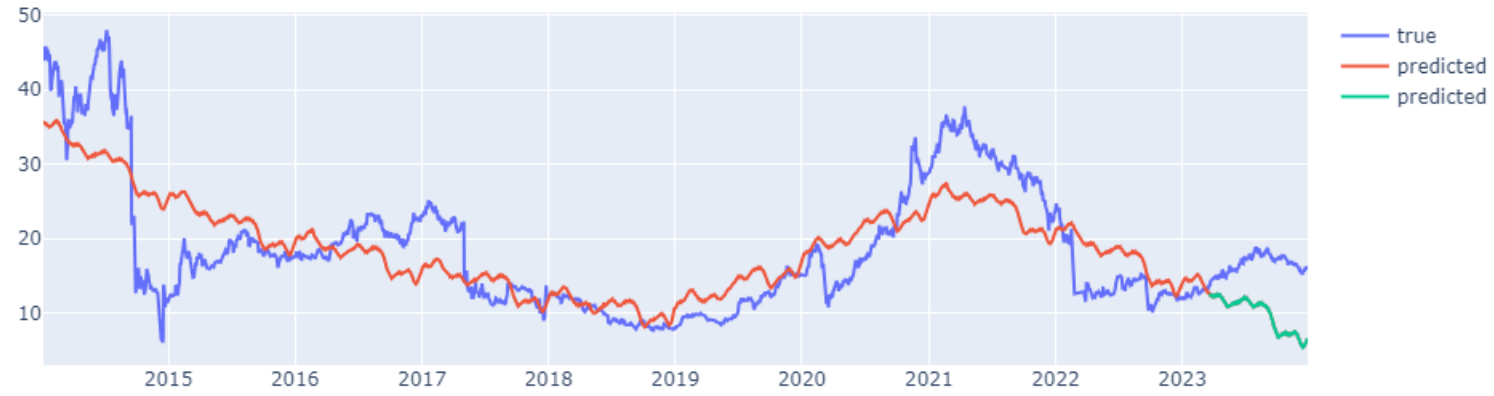
SARIMAX

- Был реализован алгоритма SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous factors)
- Наилучшие результаты показал алгоритм с параметрами :
 $p=1, d=1, q=1,$
 $P=2, D=1, Q=3, S=7$
- На загруженных временных рядах алгоритм показал низкие метрики качества MSE

Prophet

- Был реализован алгоритм Facebook Prophet с гиперпараметрами, подобранными по сетке
- Данным алгоритмом тоже не удалось улучшить метрики качества

Prophet



SARIMAX



Анализ текста

- Была выдвинута идея использовать chat-gpt для прогнозирования временных рядов
- Идея заключается в том, чтобы текстовые данные, полученные с помощью парсера, загружать в chat-gpt так, чтобы он выдавал оценку о потенциале акции к росту, основываясь на предложенном тексте
- В качестве источника данных была выбрана социальная сеть для инвесторов Tinkoff Pulse
- Был реализован парсер постов Tinkoff Pulse

Итог

- К сожалению, все реализованные на данный момент алгоритмы не показали достойных результатов
- И прогноз по SARIMAX, и прогноз алгоритма Prophet показали плохие метрики качества

Прогресс

- Исправлен алгоритм машинного обучения AdaBoost с базовым алгоритмом DecisionTree
- Реализован алгорит SARIMAX
- Реализован алгоритм Prophet
- Парсер Tinkoff Pulse доработан под задачу анализа текста chat gpt

Ближайшие цели проекта

- Реализовать алгоритм прогноза на основании анализа постов в Tinkoff Pulse чатом gpt
- Добавить парсер новостных телеграм каналов и реализовать алгоритм прогноза на основании новостей