

Homework 2.

Предсказание временных рядов.

1. (без баллов) Доделать функционал класса TimeSeriesPredictor, который не успели сделать на занятии.
2. (1 балла) Опишите из каких компонент состоит модель ARIMA, за что отвечает каждая из компонент.
3. (4 балла) Постройте модель ARIMA временного ряда alcohol_sales.csv из датасет. Распишите как вы искали гиперпараметры p, d, q.
4. (5 балла) Используя класс TimeSeriesPredictor:
 - a) Добавьте в функционал метод validate_residuals, который бы принимал на вход реальные и предсказанные значения за определенный период и выдавал бы вердикт о стационарности остатков.
 - b) Попробуйте следующие модели для прогнозирования временного ряда alcohol_sales.csv: LinearRegression и GradientBoostingRegressor из sklearn. Подберите оптимальные гиперпараметры по отложенной выборке. Какая модель дает лучшее качество по метрике RMSE. Почему? Попробуйте сделать простейший model blending двух моделей по формуле:
$$Y_{final} = \alpha Y_{linear} + (1 - \alpha) Y_{boosting}$$
Улучшилось ли качество модели?