

Описание прогнозной модели

Прогнозная модель используется для оценки заказов, трудоемкости и распределения ресурсов. Модели и скрипты размещены на сервере izb-ally-nodered02 (IP: 10.1.241.244). Скрипты написаны на Python с использованием библиотеки ENTA.

Скрипты и команды для работы с моделью

```
python3.10 main.py --help
Usage: main.py [OPTIONS] COMMAND [ARGS]...

Options:
  --help  Show this message and exit.

Commands:
  copy    copy time_series from one DB to another
  fit     Fits model with collected lines from DB
  forecast  Makes forecast for the next 7 days and saves it to DB
  test    Evaluates forecast metrics
  view    Prints one segment used for forecast
```

Модель и обучение

Наилучшие результаты показывает модель CatBoost. Перед использованием модель необходимо обучить. После обучения она применяется для прогнозов. Данные для обучения поступают из различных источников, результаты сохраняются в гипертаблицу `time_series`.

Источники данных

- **Погода:**
 - Исторические данные (с 2005 года) и ежедневные обновления загружаются с `rp5.ru` в таблицы `weather` и `weather_stations` с помощью скрипта `rp5_weather` (настраивается через файл `static/cities.txt` с списком метеостанций).

- Актуальный прогноз погоды загружается ежедневно с api.met.no через Node-RED в таблицу weather.
 - **Календарь:** Учитываются данные производственного календаря из таблицы calendar и православных праздников. Включение мусульманских праздников не улучшило точность.
 - **База данных ВкусВилл**
-

Параметры прогноза

Прогноз строится на заданное количество дней вперед с различной дискретизацией (шагом по времени).

Прогноз для последней мили (сборка и доставка)

Строится с шагом 1 час на 14 дней вперед.

- **Обычный прогноз:** Основан на данных о сборке и доставке, где заказы привязаны к торговой точке (ТТ). Данные загружаются из MS SQL в БД Ally через Node-RED.
- **Зональный прогноз:** Учитывает данные о сборке/доставке и статистику распределения адресов по геозонам ТТ. На основе геозон формируются сводные буферы с вероятностями выполнения заказов (на основе статистики за 7 дней). История заказов сопоставляется с текущими геозонами для обучения модели, что позволяет прогнозировать для новых ТТ с ограниченной историей.

Прогноз для розницы

Строится с шагом 1 день на 21 день вперед. Основан на оценке трудоемкости, загружаемой во временной ряд 51 через Node-RED.

Revision #5

Created 19 August 2025 07:37:49 by Семен Долгов

Updated 24 September 2025 12:00:49 by Семен Долгов