

ML Систем Дизайн для Прогнозирования Спроса FMCG-компаний

Джаводов Санджар

Содержание

1	Сбор и хранение данных	2
1.1	Источники данных	2
1.2	Хранение	2
2	Подготовка данных	2
2.1	Очистка данных	2
2.2	Преобразование данных	2
2.3	Извлечение признаков	2
3	Обучение модели	3
3.1	Выбор алгоритмов	3
3.2	Оптимизация гиперпараметров	3
3.3	Ансамблирование	3
4	Оценка и валидация	3
4.1	Метрики	3
4.2	Бэктестинг	3
5	Развертывание модели	3
5.1	Инфраструктура	3
5.2	API	4
6	Мониторинг и обновление	4
6.1	Мониторинг производительности	4
6.2	Переобучение	4
7	Интерпретация результатов	4
7.1	Объяснение прогнозов	4
7.2	Визуализация	4

1 Сбор и хранение данных

1.1 Источники данных

- Система ERP компании (продажи, запасы)
- POS-терминалы в магазинах (транзакции)
- Внешние API (цены на нефть, погодные данные)
- Календарь праздников и маркетинговых акций

1.2 Хранение

- СУБД ClickHouse для хранения исторических данных
- Kafka для потоковой передачи данных в реальном времени

2 Подготовка данных

2.1 Очистка данных

- Обработка пропущенных значений
- Удаление дубликатов
- Обработка выбросов

2.2 Преобразование данных

- Агрегация данных о продажах по различным временным интервалам
- Создание лаговых переменных
- Кодирование категориальных переменных

2.3 Извлечение признаков

- Временные признаки (день недели, месяц, сезон)
- Признаки продукта (категория, цена, промо-акции)
- Признаки магазина (локация, тип, размер)

3 Обучение модели

3.1 Выбор алгоритмов

- ARIMA для базовой модели
- Градиентный бустинг (LightGBM, XGBoost) для учета нелинейных зависимостей
- Нейронные сети (LSTM, Transformer) для сложных временных паттернов

3.2 Оптимизация гиперпараметров

- Байесовская оптимизация
- Кросс-валидация с учетом временной структуры данных

3.3 Ансамблирование

- Взвешенное усреднение прогнозов разных моделей

4 Оценка и валидация

4.1 Метрики

- MAPE (Mean Absolute Percentage Error)
- RMSE (Root Mean Square Error)
- MAE (Mean Absolute Error)

4.2 Бэктестинг

- Скользящее окно для оценки производительности модели на разных временных интервалах

5 Развертывание модели

5.1 Инфраструктура

- Контейнеризация с использованием Docker
- Оркестрация с помощью Kubernetes

5.2 API

- REST API для получения прогнозов
- gRPC для высокопроизводительного взаимодействия между сервисами

6 Мониторинг и обновление

6.1 Мониторинг производительности

- Отслеживание дрейфа данных
- Мониторинг точности прогнозов в реальном времени

6.2 Переобучение

- Автоматическое переобучение моделей при снижении производительности
- A/B тестирование новых версий моделей

7 Интерпретация результатов

7.1 Объяснение прогнозов

- SHAP (SHapley Additive exPlanations) для интерпретации влияния признаков
- Partial Dependence Plots для визуализации зависимостей

7.2 Визуализация

- Интерактивные дашборды с прогнозами и ключевыми факторами влияния