

Оглавление

Предисловие	9
Проблема с зависимыми данными	9
Скольльзящее среднее и экспоненциальное сглаживание	12
Линейная регрессия	14
ARIMA	14
ARCH/GARCH	16
Градиентный бустинг	17
Нейронные сети	17
ЧАСТЬ 1. БИБЛИОТЕКА FACEBOOK PROPHET	19
1. Математический аппарат Facebook Prophet (на основе адаптированного перевода статьи «Forecasting at scale» от Шона Дж. Тейлора и Бенджамина Летама, входящих в команду разработчиков Facebook Prophet)	19
1.1. Введение	19
1.2. Особенности временных рядов в бизнес-задачах	21
1.3. Прогнозная модель Prophet	23
1.4. Автоматизация оценки качества прогнозов	37
1.5. Библиография	40
2. Установка Prophet	42
2.1. Установка в macOS	42
2.2. Установка в Windows	42
3. Построение простых моделей Prophet	43
3.1. Ежедневные данные	43
3.2. Ежемесячные данные	49
3.3. Почасовые данные	51
3.4. Данные с гэпами	55
4. Работа с сезонностью	61
4.1. Понимание аддитивной и мультипликативной сезонностей	61
4.2. Моделирование сезонности с помощью порядка Фурье	68
4.3. Добавление собственных сезонностей	73
4.4. Добавление условных сезонностей	77
4.5. Регуляризация сезонности	81

5. Добавление праздников.....	89
5.1. Добавление национальных праздников	89
5.2. Добавление праздников штатов/провинций.....	95
5.3. Добавление собственных праздников.....	96
5.4. Создание многодневных праздников	98
5.5. Регуляризация влияния праздников	101
6. Типы роста.....	106
6.1. Применение линейного роста.....	106
6.2. Знакомство с логистической функцией	108
6.3. Получение прогнозов для моделей логистического роста с насыщением.....	110
7. Точки изменения тренда	124
7.1. Автоматическое определение точек изменения тренда	124
7.2. Регуляризация точек изменения.....	130
7.3. Назначение пользовательских точек изменения	135
8. Добавление регрессоров.....	142
8.1. Добавление бинарных регрессоров	142
8.3. Интерпретация регрессионных коэффициентов	149
9. Выбросы и особые события.....	152
9.1. Обработка выбросов, вызывающих сезонные скачки.....	152
9.2. Обработка выбросов, являющихся причиной широких доверительных интервалов	157
9.3. Автоматическое обнаружение выбросов	159
9.4. Моделирование выбросов как особых событий.....	166
9.5. Моделирование шоков типа локдаунов во время пандемии COVID-19	168
10. Доверительные интервалы.....	178
10.1. Моделирование неопределенности тренда	178
10.2. Моделирование неопределенности сезонности	184
11. Перекрестная проверка	192
11.1. Перекрестная проверка с автоматически сгенерированными пороговыми точками	192
11.2. Перекрестная проверка с пользовательскими пороговыми точками	205
11.3. Обычная оптимизация гиперпараметров по сетке на основе перекрестной проверки	208
11.4. Байесовская оптимизация гиперпараметров на основе перекрестной проверки	211
12. Прогнозирование нескольких временных рядов.....	214

ЧАСТЬ 2. БИБЛИОТЕКА ETNA.....	241
1. Общее знакомство.....	241
1.1. Создание объекта TSDataset	248
1.2. Визуализация рядов объекта TSDataset	252
1.3. Получение сводки характеристик по объекту TSDataset	253
2. Модель наивного прогноза	255
2.1. Один временной ряд.....	255
2.2. Несколько временных рядов.....	260
3. Модель скользящего среднего.....	264
3.1. Один временной ряд.....	264
3.2. Несколько временных рядов.....	266
4. Модель сезонного скользящего среднего.....	268
4.1. Один временной ряд.....	268
4.2. Несколько временных рядов.....	270
5. Модель SARIMAX	272
5.1. Один временной ряд.....	272
5.2. Несколько временных рядов.....	280
6. Модель Хольта–Винтерса (модель тройного экспоненциального сглаживания, модель ETS).....	282
6.1. Один временной ряд.....	282
6.2. Несколько временных рядов.....	286
7. Модель Prophet	288
7.1. Один временной ряд	288
7.2. Несколько временных рядов	293
8. Модель CatBoost	300
8.1. Один временной ряд.....	300
8.2. Несколько временных рядов.....	318
9. Модель линейной регрессии с регуляризацией «эластичная сеть»	321
9.1. Один временной ряд.....	321
9.2. Несколько временных рядов.....	323
10. Объединение процедуры построения модели, оценки качества и визуализации прогнозов в одной функции.....	326
10.1. Один временной ряд.....	327
10.2. Несколько временных рядов.....	328

11. Перекрестная проверка нескольких моделей	330
11.1. Один временной ряд.....	330
11.2. Несколько временных рядов.....	331
12. Ансамбли.....	336
12.1. Один временной ряд.....	336
12.2. Несколько временных рядов.....	337
13. Стекинг.....	339
14. Создание собственных классов для обучения моделей ...	340
14.1. Использование функции <code>train_and_evaluate_model()</code> для быстрого построения базовых моделей.....	350
14.2. Ансамбли из собственных классов.....	355
15. Импутация пропусков.....	358
16. Работа с трендом и сезонностью	368
17. Обработка выбросов	382
18. Собираем все вместе	388
19. Модели нейронных сетей	402
20. Оптимизация гиперпараметров с помощью Optuna от разработчиков	410
21. Задача Райффайзен Банка (8 временных рядов)	415
21.1. Описание задачи.....	415
21.2. Построение прогнозной модели для горизонта в 90 дней (на продажи в конкретном магазине не влияют продажи и рекламная активность остальных магазинов)	415
21.3. Оптимизация гиперпараметров модели с горизонтом в 90 дней (на продажи в конкретном магазине не влияют продажи и рекламная активность остальных магазинов)	432
21.4. Построение прогнозной модели для горизонта в 729 дней (на продажи в конкретном магазине не влияют продажи и рекламная активность остальных магазинов)	475
21.5. Сравнение прогнозных конвейеров, созданных вручную и с помощью ETNA	489
22. Задача Store Sales – Time Series Forecasting (1782 временных ряда).....	524
22.1. Описание задачи.....	524