

Практический курс по временным рядам

Организационная лекция

О себе

- Закончил МФТИ (ФИВТ)
- Tinkoff
 - time series, классический ML, anti-fraud
 - разработчик сервиса для мониторинга данных на основе анализа временных рядов
 - maintainer in-house фреймворка для анализа временных рядов
- 4 года опыта работы с разными задачами анализа временных рядов
- Преподаю python на физтехе и в сообществе moscow python

Прerequisites к курсу

- Уверенное знание python
- Базовые знания ML
- Git + shell

О курсе

- Курс с упором на практику
- Сжатая теория “на базовое понимание”
- Множество практики на написание кода как на семинарах, так и в домашних
- Покрывает большинство задач анализа time series: прогнозирование, поиск выбросов, классификация

Почему упор на практику?

1. Анализ временных рядов можно условно разделить на классический и современный подход.
2. Классический подход к анализу - отдельная наука эконометрика.
3. Современный (ML) подход к задачам по большей части не содержит специфичной теории, выходящей за рамки общего курса машинного обучения.
4. Зато область содержит множество отдельных задач и кейсов, интуицию для решения которых можно выработать на практике.

Расписание и нагрузка

- Удаленно (zoom)
 - ссылку на созвон пришлют организаторы
 - запись созвонов
- 10 занятий по 3 часа
 - понедельник, среда
 - 18:30 - 21:30

Программа курса

1. Базовые понятия для анализа временных рядов. Классификации, преобразования, статистические критерии.
2. Компоненты временных рядов. STL разложение.
3. Эконометрический подход к прогнозированию. Статистические и авторегрессионные модели: MA, AR, ARIMA, SARIMA и т.п.
4. Современный подход к прогнозированию: получение матрицы лагов, обогащение доп признакам. Обучения произвольной модели регрессии.
5. Метрики качества.
6. Поиск аномалий во временных рядах, метод доверительных интервалов, метод population stability index.

Программа курса

7. Нейронные сети для анализа временных рядов. Часть 1.
8. Нейронные сети для анализа временных рядов. Часть 2.
9. Обзор задачи классификации временных рядов.
Dynamic time warping, метод скользящего окна и др.
Использование DTW для классификации датасета Beetle Fly.
Использование CNN для задачи классификации.

Система оценивания

- Дз каждую неделю: либо отдельное, либо дорешивание задач с семинара.
- Формат сдачи - juruter тетрадка
- Дипломный проект либо соревнования на Kaggle по желанию
- Оценка трех типов
 - < 40% - незачет
 - >= 40% - зачет
 - >= 90% - зачет с отличием

Ссылка на репозиторий с материалами курса

- https://github.com/technogleb/ts_summer
- Рекомендуется прочитать README и подготовить окружение
- Озаботиться доступом из корпоративного контура

Ссылки

■ Преподаватель - Глеб Синяков

⚡ t.me/technogleb

■ Куратор (СБЕР) - Дарья Клименко

⚡ t.me/DAKlimenko

■ Куратор (МФТИ) - Анастасия Иванова

⚡ t.me/AnastasiaIvanovaMarketing

■ Ссылка на инфо канал

⚡ https://t.me/joinchat/DulUmkLZY_xlZWQ6

■ Ссылка на чат

⚡ https://t.me/joinchat/djm_q62xl5sxOWNi