

Программа курса «PYNN: Нейронные сети на Python»

О курсе: Вы начнете с самых основ, изучите математический аппарат, лежащий в основе нейронных сетей, и постепенно перейдете к созданию и настройке сложных архитектур для решения реальных задач в области компьютерного зрения, обработки естественного языка и многого другого.

Аудитория:

- Начинающие Data Scientists
- Разработчики на Python
- Студенты и исследователи
- IT-специалисты

Уровень подготовки: Опыт программирования, прохождение онлайн-курса FUNP «Основы языка Python для анализа данных и решения задач машинного обучения»

Продолжительность курса: 24 академических часа, 6 дней по 4 ак. часов дистанционно

Содержание программы

1. Введение в нейронные сети

Теоретическая часть:

- что такое нейронные сети и где они работают лучше классических методов
- как устроен процесс обучения: от данных к предсказаниям
- простые архитектуры для классификации изображений
- как оценивать качество моделей: ключевые метрики

Практическая часть:

- настраиваем окружение для экспериментов
- знакомство с PyTorch: основные операции с тензорами
- запускаем готовые модели (MobileNetV2, ResNet-50)
- осваиваем полный цикл: загрузка весов → предобработка → инференс → интерпретация результатов

Итог: ваш первый работающий классификатор изображений

2. Компьютерное зрение (CV)

Теоретическая часть:

- детекция, сегментация, трекинг объектов, задачи распознавания – что выбрать для вашей задачи
- эволюция архитектур: от сверточных сетей до современных решений
- метрики качества для разных задач компьютерного зрения
- когда использовать готовые модели, а когда обучать свои

Практическая часть:

- детекция объектов (YOLOv11)
- сегментация объектов (DeepLabV3, YOLOv11-seg)
- детекция ключевых точек (YOLOv11-pose)
- распознавание текста на изображениях с EasyOCR
- генерация изображений через предобученные GAN

Итог: набор инструментов для решения задач анализа изображений и видео

3. Обработка текста (NLP)

Теоретическая часть:

- подготовка текстовых данных: очистка, лемматизация, токенизация

- эволюция подходов: от Bag-of-Words к трансформерам
- основные виды нейронных сетей для задач NLP: RNN/GRU/LSTM, Transformer/BERT
- fine-tuning предобученных моделей под свои задачи

Практическая часть:

- препроцессинг данных
- классификация текстов с помощью TF-IDF и нейросетей
- работа с PyTorch: от простых моделей до трансформеров
- fine-tuning BERT для конкретной задачи

Итог: модель для классификации текстов с пониманием сильных сторон каждого подхода

4. Обработка звука

Теоретическая часть:

- цифровое представление звука: от waveform до спектрограмм
- детекция речи, разделение дикторов, транскрибация - когда что применять
- современные архитектуры для работы с аудио
- сравнение подходов: от Conformer до Wav2Vec/Whisper

Практическая часть:

- детекция речевых сегментов в записи
- разделение дикторов с помощью Pyannote
- транскрибация речи различными моделями ASR

Итог: готовый пайплайн для обработки аудио от очистки до расшифровки

5. Оптимизация моделей

Теоретическая часть:

- методы ускорения инференса: квантование, pruning, дистилляция
- выбор формата под целевую платформу: CPU, GPU, мобильные устройства
- сравнение фреймворков оптимизации
- баланс между скоростью и точностью

Практическая часть:

- конвертация моделей в кроссплатформенный ONNX
- оптимизация для CPU через OpenVINO
- ускорение на GPU с TensorRT

Итог: навыки адаптации моделей под требования продакшена

6. Внедрение в продакшен

Теоретическая часть:

- MLOps: жизненный цикл ML-проекта
- паттерны развертывания моделей
- мониторинг и обновление моделей
- выбор стратегии под масштаб проекта

Практическая часть:

- простой веб-сервис на FastAPI с моделью на борту
- настройка высокопроизводительного Triton Inference Server
- сравнение latency и throughput разных решений

Итог: опыт развертывания моделей от прототипа до продакшена