

Обработка данных с использованием Python (PYTHON_DATA)

ID PG-PYTHON_DATA Цена 89 250,- руб. Длительность 5 дней

Кому следует посетить

Инженеры по работе с данными, а также разработчики ПО, желающие изучить основы машинной обработки данных

Предварительные требования

Знания языка Python на уровне курса [Программирование на языке Python. Базовый уровень \(PYTHON_BASIC\)](#).

Содержание курса

Данный курс рассматривает основные возможности языка Python и вспомогательных библиотек для работы с большими объемами структурированных и неструктурированных данных с использованием локальных ресурсов и облачных сред. Изучаются основы машинного обучения как способа работы с неструктурированными данными на основе библиотеки Tensorflow. Также рассматриваются распространённые инструменты визуализации данных.

Программа курса

День 1. Python для работы с данными.

1. Обработка данных с помощью Python
2. Введение в Jupyter Lab
3. Собственные структуры данных Python
4. Библиотека NumPy (цель и возможности)
5. Библиотека Matplotlib (цель и возможности)

Лабораторная работа 1. Работа с типами данных

День 2. Оценка данных, введение в машинное обучение

1. Библиотека Pandas, понятие датафреймов и операции над ними
2. Введение в машинное обучение

- Терминология
- Виды и возможности
- Обучение моделей
- Градиентный спуск
- Hyperparameters, Losses, Learning Rate
- Generalization
- Сеты данных

Лабораторная работа 2. Работа с датафреймами

Лабораторная работа 3. Изучение зависимостей в машинном обучении

День 3. Tensorflow и Feature Engineering

1. Введение в Tensorflow и Keras
2. Возможности и способы использования
3. Готовые модели
4. Инструменты для создания собственных моделей
5. Feature Selection and Engineering

Лабораторная работа 4. Использование Tensorflow

Лабораторная работа 5. Feature Engineering

День 4. Глубокое обучение и среды для обучения моделей

1. Глубокое обучение
2. Применение и возможности
3. Ресурсы для обучения моделей
4. Облачные ресурсы для обучения моделей

Лабораторная работа 6. Глубокое обучение

Лабораторная работа 7. Обучение моделей в подготовленной среде

День 5. Визуализация данных.

1. Библиотеки Python для визуализации
2. Облачные среды для визуализации данных
3. Статистическая оценка данных

Лабораторная работа 8. Визуализация и статистическая оценка данных