



Программа курса «Машинное обучение»

EcoAcademy

www.econophysica.ru

+7 (3822) 900-601

Томск, пр-т Ленина 60/1, 3 этаж

Модуль 0. Теоретическая база

Данный модуль знакомит слушателя с ключевыми математическими концепциями и методами, необходимыми для понимания машинного обучения, а также с языком Python – наиболее популярным языком в области data science.

Основы математических методов

- ❖ Основы теории вероятности и статистики
- ❖ Линейная алгебра
- ❖ Методы оптимизации
- ❖ Основные понятия машинного обучения

Основные понятия машинного обучения

Основы программирования на Python для машинного обучения

- ❖ Основы обработки данных (структуры данных, операции над данными)
- ❖ NumPy
- ❖ Pandas

Практическая часть:

- ❖ Основы Python
- ❖ Интерактивные редакторы
- ❖ Инструменты для построения графиков
- ❖ Scikit-learn
- ❖ Чтение, запись и основные операции с датафреймами
- ❖ Основные операции с NumPy



Модуль 1. Обучение с учителем

В этом модуле вводятся понятия, связанные с алгоритмическим обучением на данных. Слушателям даётся начальное понимание приложений контролируемого машинного обучения, прежде чем проходить основные блоки, начиная с линейных моделей регрессии и классификации до моделей с нелинейными ядрами и теории, лежащей в основе метода опорных векторов, и заканчивая мощными ансамблями методов.

Часть 1. Классификация:

Обзор круга решаемых задач

- ❖ Приложения машинного обучения
- ❖ Обучение на данных

Задача классификации

- ❖ Основные понятия задачи классификации
- ❖ Примеры применения

Алгоритмы классификации

- ❖ Линейные модели и модели с нелинейными ядрами

Практическая часть:

- ❖ Реализация логистической регрессии с помощью Numpy
- ❖ Реализация логистической регрессии с помощью Sklearn
- ❖ Оценка качества моделей

Часть 2. Регрессия:

Задача регрессии

- ❖ Аппроксимация, интерполяция, экстраполяция
- ❖ Заполнение пропусков, прогнозирование

Линейные модели и модели с нелинейными ядрами

Практическая часть:

- ❖ Прогноз с помощью методов регрессионного анализа пакета Sklearn

Модуль 2. Обучение без учителя

Важным классом задач машинного обучения является обучение без учителя – обучение при отсутствии “учителя”, т.е. без размеченных примеров.

В данном модуле слушатель курса познакомится с алгоритмами обучения без учителя, а также с задачами кластеризации и снижения размерности, которые данные алгоритмы позволяют решать.

Задача кластеризации

- ❖ Особенности кластерного анализа

Алгоритмы кластеризации

Задача отбора признаков и снижения размерности

Практическая часть:

- ❖ Кластеризация данных, Sklearn
- ❖ Визуализация данных с применением алгоритмов снижения размерности

Модуль 3. Практический подход к машинному обучению

Этот модуль фокусируется на практических задачах и сложностях, возникающих при развёртывании моделей машинного обучения в реальном контексте.

Каждая сессия в этом модуле охватывает конкретную практическую проблему и предоставляет кандидатам руководство и представление о том, как подходить к различным этапам цикла разработки модели: от сбора данных и экзаменов до тестирования моделей, проверки и интерпретации результатов и связи.

Алгоритмы и методы машинного обучения

- ❖ Фазы обучения (тренировки) и применения моделей

Проблемы и ограничения машинного обучения, способы их преодоления

- ❖ Разделение выборки данных на обучающую, валидационную и тестовую части
- ❖ Домены применимости алгоритмов. Выбор алгоритмов

Извлечение признаков

Многоуровневое машинное обучение

- ❖ Ансамблирование
- ❖ Примеры алгоритмов

Практическая часть:

- ❖ Использование XGBoost

Модуль 4. Нейронные сети

Нейросетевые модели – важный инструмент при решении множества современных задач машинного обучения.

В данном модуле слушатели научатся понимать алгоритмы и решать основные проблемы, возникающие при обучении нейронных сетей. Модуль начинается с модели перцептрона и метода обратного распространения ошибки, объясняет методы регуляризации и оптимизации. Курс покрывает теоретические основы нейронных сетей, однако приоритет отдается практическим работам, позволяющим напрямую экспериментировать с нейронными сетями.

Модель перцептрона

Задача оптимизации в алгоритмах машинного обучения: градиентный спуск

Метод обратного распространения ошибки

Алгоритмы оптимизации гиперпараметров ML алгоритмов

Практическая часть:

- ❖ Реализация перцептрона с помощью Numpy
- ❖ Реализация перцептрона с помощью Sklearn

Модуль 5. Глубокое обучение

Глубокое обучение является центральным и наиболее важным компонентом множества впечатляющих IT-продуктов.

Данный модуль идеологически продолжает предыдущий модуль (Нейронные сети). Объясняется причина перехода от традиционных сетей к более глубоким архитектурам, современные глубокие архитектуры сетей, регуляризация в глубоких нейронных сетях, продвинутые способы оптимизации весов нейронных моделей.

В практической части модуля делается упор на использование современных фреймворков машинного обучения.

Часть 1. Глубокое обучение:

Введение в глубокое обучение

- ❖ Основные понятия глубокого обучения
- ❖ Области применения

Алгоритмы глубокого обучения

- ❖ Архитектуры нейронных сетей

Практическая часть:

- ❖ Оптимизация гиперпараметров моделей по различным методам

Часть 2. Оптимизация глубоких сетей:

Регуляризация глубоких сетей

- ❖ Недообучение и переобучение моделей
- ❖ Методы решения

Продвинутые стратегии оптимизации

Практическая часть:

- ❖ Построение графиков обучения

Модуль 6. Продвинутое машинное обучение

В данном модуле слушатели узнают о самых современных методах машинного обучения. В частности, слушатели узнают об обработке временных рядов и естественного языка с помощью рекуррентных и LSTM (Long Short-Term Memory, долгая краткосрочная память) сетей, а также о способах решения множества задач компьютерного зрения (такие, как детекция, сегментация, классификация...) с помощью современных архитектур нейронных сетей.

Часть 1. Работа с глубокими сетями:

Архитектуры нейронных сетей

- ❖ Рекуррентные нейронные сети
- ❖ LSTM (долгая краткосрочная память)
- ❖ Autoencoder, Encoder-Decoder, Kohonen map (самоорганизующаяся карта Кохонена), Generative Adversarial Networks (генеративно-сопоставительные сети)

Обработка естественных языков

Инструменты обработки естественных языков

Практическая часть:

- ❖ Построение и обучение LSTM / GRU сетей в Keras
- ❖ Распознавание сущностей

Часть 2. Обработка изображений при помощи глубокого обучения:

Компьютерное зрение

- ❖ Классификация изображений
- ❖ Сегментация
- ❖ Детекция и трекинг объектов

Инструменты компьютерного зрения

Практическая часть:

- ❖ Сегментация изображений
- ❖ Детекция

Общая продолжительность курса – **40 часов**

Курс состоит из 6 модулей.

Занятия в вечернее время два раза в неделю по 4 часа

Стоимость: **40 000 Р**

Преподаватели



Алексей Кульневич

Data Scientist с богатым практическим опытом в интеллектуальном анализе данных, машинном обучении и задачах оптимизации, победитель хакатонов по машинному обучению регионального и всероссийского масштаба.



Сергей Зюбин

Head of Practice, кандидат физико-математических наук, доцент кафедры ВМиМФ ФТИ ТПУ с двадцатилетним практическим и преподавательским опытом.



Роман Чугунов

Data Scientist, имеет опыт решения сложных бизнес-задач по машинному обучению и анализу данных, победитель хакатонов по машинному обучению регионального и всероссийского масштаба.

Зарегистрироваться и узнать подробности можно у специалиста компании

Электронная почта: olga.bragina@econophysica.com

Телефон: +7 (3822) 90-06-01, доб. 1002