



Методы машинного обучения

ИУ-5, магистратура, 2 семестр

Объём дисциплины по видам учебных занятий (в часах) – 3 з.е.

Виды учебной работы	Объём по семестрам, ч	
	Всего	1 семестр
Объём дисциплины	108	108
Аудиторная работа	68	68
Лекции (Л)	34	34
Лабораторные работы (ЛР)	34	34
Самостоятельная работа (СР)	40	40
Проработка материала лекций	6	6
Подготовка к лабораторным работам	20	20
Подготовка к рубежным контролям	6	6
Подготовка реферата	8	8
Вид промежуточной аттестации	распр. экзамен	

Содержание дисциплины, структурированное по модулям

№ п/п	Тема (название) модуля	Виды занятий, часы		
		Л	ЛР	СР
1 семестр				
1.	Методы обработки данных	18	17	20
2.	Методы построения моделей машинного обучения	16	17	20
	ИТОГО	34	34	40

Структура курса

- Лекции:
 - СРЕДА, 17.25, 424ю
- Лабораторные работы – 6 лр.
 - ПЯТНИЦА (по расписанию)
- Реферат (статья) по машинному обучению.
- НИРС по обработке и анализу данных (3 семестр):
 - Решение комплексной задачи, объединяющей различные технологии анализа данных и машинного обучения –
[https://github.com/Yorko/mlcourse.ai/wiki/Individual-projects-and-tutorials-\(in-Russian\)](https://github.com/Yorko/mlcourse.ai/wiki/Individual-projects-and-tutorials-(in-Russian))
- Репозиторий курса:
 - https://github.com/ugapanyuk/ml_course_2020/
- Бонусы:
 - участие в соревнованиях по машинному обучению, подготовка статей, выступление на студенческой весне.

Модуль 1. Методы обработки данных

1. Вводная лекция. Цели и задачи курса. Содержание курса. Обзор существующих источников по анализу данных и машинному обучению.
2. Методология CRISP-DM. Описательный анализ данных. Обзор библиотек для описательного анализа данных.
3. Использование библиотек Pandas и Pandasql для обработки данных.
4. Визуализация данных с помощью библиотек Matplotlib и Seaborn.
5. Использование визуальных средств анализа данных: KNIME, RapidMiner, Orange
6. Применение методов анализа данных в гибридных интеллектуальных информационных системах.

Модуль 2. Методы построения моделей машинного обучения

1. Основные задачи машинного обучения: обучение с учителем, обучение без учителя, обучение с подкреплением, другие виды обучения. Задачи классификации и регрессии.
2. Методы обучения с учителем. Метод ближайших соседей. Методы на основе градиентного спуска с регуляризацией. Машина опорных векторов. Деревья решений.
3. Методы обучения с учителем. Ансамбли моделей. Случайный лес. Бэггинг. Бустинг. Метрики оценки качества моделей.
4. Методы обучения без учителя. Методы и метрики кластеризации. Методы понижения размерности.

Курсы по машинному обучению

- <https://netology.ru/programs/data-scientist>
- <http://digitaltech.school/courses/data-science-and-python>
- <https://skillbox.ru/course/profession-machine-learning/>
- https://geekbrains.ru/geek_university/data-science
- Видеолекции курса в ШАД - <https://yandexdataschool.ru/edu-process/courses/machine-learning>
- <https://www.coursera.org/learn/vvedenie-mashinnoe-obuchenie> - краткий курс
- <https://www.coursera.org/specializations/machine-learning-data-analysis> - специализация из 6 курсов
- <https://mlcourse.ai/> - курс по машинному обучению от OpenDataScience. Статьи на хабре: <https://habr.com/ru/company/ods/blog/322626/>
- Открытый курс «Data Mining in Action» - <http://www.dmia.space/>

Интересные ссылки:

- <http://www.machinelearning.ru> - основной ресурс по машинному обучению на русском
- <https://neurohive.io/ru/> - статьи на русском по ИИ
- <https://www.kaggle.com/>
- <https://towardsdatascience.com/>
- <https://machinelearningmastery.com/>
- <https://paperswithcode.com/sota> - публикации, сгруппированные по основным направлениям ИИ
- <https://arxiv.org/> - открытая библиотека научных публикаций Корнеллского университета

Книги

- Одна из лучших книг по машинному обучению.



Книги

- Хорошая вводная книга непосредственно по Python.



Книги

- Хорошая практическая книга с описанием библиотек Python. Меньше примеров задач, больше описания библиотек.

Дж. Вандер Плас

Python

для сложных задач
наука о данных:
и машинное обучение

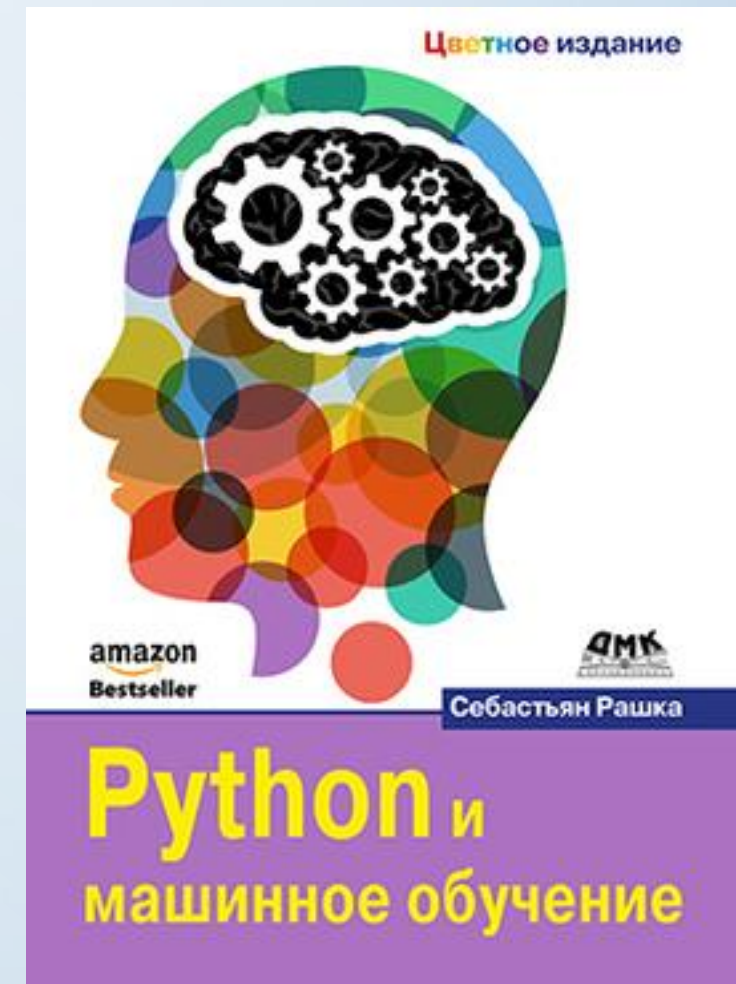


Санкт-Петербург • Москва • Екатеринбург • Воронеж
Нижний Новгород • Ростов-на-Дону • Самара • Минск

2018

Книги

- Рассматривается весь жизненный цикл анализа данных – выделение признаков, оценка качества моделей.
- Рассматриваются много различных задач машинного обучения, в том числе довольно специфических.



Книги

- Хорошая практическая книга. Исторически была переведена первой.
- Содержит в основном примеры решения задач. В меньшей степени содержит описание библиотек.



Книги

- Рассматриваются как задачи обучения с учителем, так и задачи обучения без учителя.
- Рассматривается весь жизненный цикл анализа данных – выделение признаков, оценка качества моделей.

Андреас Мюллер, Сара Гвидо

Введение в машинное обучение с помощью Python

Руководство для специалистов по работе с данными



Москва
2016-2017

Книги

- Хорошая вводная книга в основном по методам машинного обучения. Методы разбираются достаточно детально.
- Примеров кода относительно немного.



Книги

- Теоретический учебник. Разбираются теоретические основы машинного обучения на основе большого количества примеров.
- Не привязан к конкретному языку программирования.

