

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Уральский федеральный университет имени первого Президента России
Б. Н. Ельцина»



УТВЕРЖДАЮ

Директор по образовательной деятельности

 С.Т. Князев

2021 г.

Спортивный анализ данных

Учебно-методические материалы по направлению подготовки
09.04.01 Информатика и вычислительная техника
Образовательная программа «Инженерия искусственного интеллекта»

Екатеринбург

2021

РАЗРАБОТЧИКИ УЧЕБНО-МЕТОДИЧЕСКИХ МАТЕРИАЛОВ

Доцент, канд.техн.наук



Созыкин Андрей
Владимирович

Спортивный анализ данных

Семестр 1

1. [Введение](#). Работа с соревнованиями kaggle из [google colab](#)
2. Знакомство с [numpy](#). [Задание](#) по теме
3. Знакомство с [pandas](#). [Задание](#) по теме
4. [Pandas](#). Часть 2. [Задание](#) по теме
5. Exploratory data analysis ([EDA](#)) - разведывательный анализ данных. [Пример](#) построения диаграмм. [Задание](#) по теме
6. [Преобработка данных](#): шкалирование, выбросы, пропуски в данных, дискретизация данных, категориальные признаки
7. [Метрики](#) качества моделей. [Реализация метрик](#) в sklearn
8. [Анализ качества моделей: типы выборок, валидация, перекрестная валидация. Значимость признаков для модели. Конвейеры. Подбор гипер параметров](#)
9. Знакомство с первой моделью. Линейная [регрессия](#). Задание: применить алгоритм линейной регрессии для соревнования актуального на момент прохождения занятия (по выбору преподавателя). [Пример применения линейной регрессии и создание новых признаков для линейной регрессии](#). При решении всех заданий необходимо придерживаться [шаблона](#)
10. [Логистическая регрессия](#). Задание: применить алгоритм логистической регрессии для соревнования актуального на момент прохождения занятия (по выбору преподавателя)
11. [К ближайших соседей](#). Применить алгоритм регрессии к ранее выбранному соревнованию с регрессионной целевой меткой. Применить алгоритм классификации к ранее выбранному соревнованию с классификационной целевой меткой
12. [Деревья решений](#). Применить алгоритм регрессии к ранее выбранному соревнованию с регрессионной целевой меткой. Применить алгоритм классификации к ранее выбранному соревнованию с классификационной целевой меткой
13. [Наивный Байес](#). Применить алгоритм классификации к ранее выбранному соревнованию с классификационной целевой меткой
14. [Машины опорных векторов](#). Применить алгоритм регрессии к ранее выбранному соревнованию с регрессионной целевой меткой. Применить алгоритм классификации к ранее выбранному соревнованию с классификационной целевой меткой
15. [Ансамбли моделей](#). Применить алгоритм регрессии к ранее выбранному соревнованию с регрессионной целевой меткой. Применить алгоритм классификации к ранее выбранному соревнованию с классификационной целевой меткой

Семестр 2

1. Выбор пары актуальных на данный момент соревнований: на регрессию, на классификацию, на NLP, на временной ряд. Подбор датасета, который подойдет для решения задач кластеризации
2. Реализации ансамблей моделей: [XGBoost](#), [LightGBM](#), [CatBoost](#)
3. [Методы понижения размерности](#): сжатие информации, визуализация
4. [Задачи обучения без учителя](#): кластеризация, поиск аномалий. Задачи кластеризации: сегментация, понижение размерности, частичное обучение, поисковые механизмы, сегментация изображений
5. Задачи обработки естественного языка (natural language processing, [NLP](#)). Реализация методов на [практике](#). Применение продвинутых [моделей](#)
6. [Временные ряды](#): [классические методы](#), методы на основе [генерации временных признаков и регрессионных моделей](#)
7. [Введение](#) в нейронные сети. [Инициализация](#) весов нейронной сети. [Пример](#) реализации нейронных сетей на Tensor Flow. Задание: выбрать актуальное соревнование на kaggle и применить на нем навыки полученные на занятии
8. [Свёрточные нейронные сети](#) (convolutional neural network, CNN). Реализация [CNN](#), применение [API keras](#), [аугментация](#) данных. Задание: выбрать актуальное соревнование на kaggle и применить на нем навыки полученные на занятии
9. [Предварительно обученные нейронные сети](#). [Пример](#) использования
10. [Сложные свёртки](#). Продвинутые архитектуры: [Inception](#), [ResNet](#), [DenseNet](#), [NASNet](#), [EfficientNet](#). Применить все продвинутые архитектуры для решения и улучшения результата на ранее выбранном соревновании
11. Другие темы распознавания изображений: [детектирование](#) объектов, [распознавание лиц](#), [сегментация](#). Реализация [UNet](#)
12. Задачи обработки естественного языка (natural language processing, [NLP](#)). Реализация методов на [практике](#). Задание: выбрать актуальное соревнование на kaggle и применить на нем навыки полученные на занятии
13. [Рекуррентные нейронные сети](#). [Реализация](#). Применить все архитектуры: RNN, LSTM, GRU - для решения и улучшения результата на ранее выбранном соревновании
14. [Нейронные сети во временных рядах \(серия блокнотов\)](#). Задание: выбрать актуальное соревнование на kaggle и применить на нем навыки полученные на занятии