



Программа онлайн-интенсива для 7–11 классов
«Подготовка к муниципальному этапу ВсОШ и перечневым олимпиадам по искусственному интеллекту»

№	Тема занятия	Формат занятия	Содержание занятия
1	Теория	Лекция + семинар	Типы задач: регрессия, классификация, кластеризация, обучение с учителем и без. Жизненный цикл ML-проекта.
2	Предобработка и анализ данных – 1	Лекция + семинар	Обзор библиотек Python (pandas, scikit-learn, matplotlib). Установка и настройка окружения. Первое знакомство с pandas: загрузка CSV-файла (например, Iris или Titanic), вывод базовой статистики.
3	Предобработка и анализ данных — 2	Лекция + семинар	Работа с пропусками: удаление, заполнение средним/медианой. Кодирование категориальных признаков (One-Hot, Label Encoding). Нормализация и стандартизация. Визуализация для анализа (гистограммы, scatter plot, box plot). Полный анализ и предобработка датасета. Построение 3–5 различных графиков для понимания данных.
4	Линейные модели — 1. Регрессия	Лекция + семинар	Математические основы линейной регрессии (метод наименьших квадратов). Переобучение и недообучение. Разделение данных на train/valid/test. Метрики качества: MSE, MAE, R ² . Реализация линейной регрессии в scikit-learn. Сравнение качества на разных выборках. Расчет метрик вручную и с помощью библиотеки.



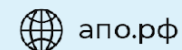


5	Линейные модели. Часть 2: Классификация	Лекция + семинар	Логистическая регрессия. Порог классификации. Метрики качества: Accuracy, Precision, Recall, F-мера, AUC-ROC. Кросс-валидация. Решение задачи бинарной классификации. Построение ROC-кривой и подбор оптимального порога.
6	Ансамбли моделей и деревья решений	Лекция + семинар	Принцип работы деревьев решений. Критерии (энтропия, Джини). Bagging (Random Forest) и Boosting (Gradient Boosting, XGBoost, LightGBM). Сильные и слабые стороны разных подходов. Сравнение точности одного дерева решений и случайного леса на одном датасете.
7	Нейронные сети	Лекция + семинар	Базовые принципы: искусственный нейрон, функции активации, архитектура полносвязной сети (FNN), алгоритм обратного распространения ошибки (backpropagation). Фреймворки: Keras, TensorFlow, PyTorch. Построение и обучение простой нейросети для классификации рукописных цифр (MNIST).
8	Обработка естественного языка (NLP)	Лекция + семинар	Основные этапы обработки текста: токенизация, стемминг, лемматизация. Векторизация: Bag of Words, TF-IDF. Word Embeddings (Word2Vec, GloVe) как основа современных подходов. Создание классификатора тональности отзывов с использованием TF-IDF и логистической регрессии.
9	Задачи оптимизации и генетические алгоритмы	Лекция + семинар	Формальная постановка задачи оптимизации. Понятие о NP-трудных задачах. Принципы работы эволюционных алгоритмов: селекция, кроссовер, мутация.





Интенсивные курсы



10	Сверточные нейронные сети (CNN) и компьютерное зрение	Лекция + семинар	Свертки и пулинг. Основы архитектуры CNN для распознавания изображений. Примеры современных архитектур (LeNet, VGG).
11	Заключительная олимпиада интенсива, приближенная к муниципальному этапу	Олимпиада	Написание олимпиады, приближенной к муниципальному этапу.
12	Разбор олимпиады	Лекция	Разбор ошибок, повторение пройденного материала.

