**Лабораторная работа по методам машинного обучения**

Эту лабораторную работу можно сдавать на протяжении всего семестра, она является повторением курса Методы машинного обучения, которые являются одной из обязательных тем дисциплины Интеллектуальная обработка и анализ данных.

На **diskstation** в папке дисциплины **ТИОиАД/Datasets** находятся датасеты, которые вы анализировали на нашем первом занятии и проделайте с ними то, о чем вы рассказывали устно. Там же находится файл с описанием параметров датасетов. Датасет про собак заменен на датасет Personality Types.

Вот примерный план работы:

1. Произведите предобработку датасета, если это необходимо.
2. При помощи матрицы корреляции (или другими известными вам способами) найдите зависимости и визуализируйте их в виде графиков.
3. Выберите целевой столбец (как правило это тот параметр, который зависит от остальных параметров и который можно спрогнозировать на основе остальных, к тому же прогноз этого столбца должен представлять какую-то ценность и смысл, *например, не логично определять пол человека по количеству времени проведенном с гаджетом, очевидно, что цель исследования должна быть другой.*).
4. Обучите 2-3 модели машинного обучения известные вам. Примените параметры, которые позволяют избежать переобучение моделей (в каждой модели это разные параметры).
5. На основе метрик качества и ROC-кривой в случае моделей классификации выберите лучшую модель.