

## Проект

В рамках курса предусмотрена защита проекта. Проект представляет из себя программный комплекс, решающий задачу извлечения и использования знаний из реальных данных с использованием машинного обучения.

Проект готовится в течение месяца и включает в себя следующие этапы:

- Постановка задачи. Предлагается самостоятельно найти предметную область и обосновать применение в ней машинного обучения.
- Разработка данных. Одно из требований к проекту - использование данных из открытых источников. Необходимо разработать процесс сбора и очистки данных.
- Поиск алгоритма и модели для решения задачи. Необходимо выполнить подготовку данных, выбрать алгоритм и подобрать параметры для построения модели.
- Использование модели для достижения поставленной цели. Необходимо реализовать применение разработанной модели.
- Построение процесса. Решение задачи необходимо оформить в единый процесс по обработке данных от источника до предсказания, не требующий участия эксперта.
- Обоснование процесса. Необходимо показать работоспособность модели, предоставить отчет о проделанной работе. Описать компоненты и архитектуру системы, описать найденные в данных факты, описать характеристики построенной модели и показать, что ее качество является приемлемым.

Проект будет оцениваться субъективно по применению пройденных в курсе тем.

Обязательно наличие модели и применение практик по настройке модели.

Использование хранилищ данных и фреймворков для управления процессов по обработке данных будет плюсом, но не является обязательным (требует много времени по настройке, поэтому можно это делать при наличии заинтересованности в технологиях)