

**Московский авиационный институт  
(Национальный исследовательский университет)**

Институт: «Информационные технологии и прикладная математика»

Кафедра: 806 «Вычислительная математика и программирование»

Дисциплина: «Искусственный интеллект»

**Лабораторная работа № 0**

Студент: Шараковский Ю. Д.

Группа: 80-306

## 1. Цель работы

Вы выступаете в роли предприимчивого начинающего стартапера в области машинного обучения. Вы заинтересовались этим направлением и хотите предложить миру что-то новое и при этом неплохо заработать. От вас требуется определить задачу которую вы хотите решить и найти под нее соответствующие данные. Так как вы не очень богаты, вам предстоит руками проанализировать данные, визуализировать зависимости, построить новые признаки и сказать хватит ли вам этих данных, и если не хватит найти еще. Вы готовитесь представить отчет ваши партнерам и спонсорам, от которых зависит дальнейшая ваша судьба. Поэтому тщательно работайте:) И главное, день промедления и вас опередит ваш конкурент, да и сплагиченная работа отразится на репутации

## 2. Результат

Выбрал датасет с помощью веб-сервиса kaggle, загрузил датасет с использованием библиотеки pandas, убрал лишние данные, проанализировал, восстановил пропущенные значения (NaN), построил гистограмму распределения числовых признаков, отнормировал данные, проверил сбалансированность датасета, проверил отсутствие корреляции между признаками.

## 3. Вывод

В результате выполнения данной лабораторной работы я познакомился применил pandas для анализа с работы с данными (DataMining) и получил практический опыт препроцессинга данных.