*Федеральное государственное бюджетное образовательное учреждение* *высшего профессионального образования*

|  |  |
| --- | --- |
| **Gerb-BMSTU_01** | ***«Московский государственный технический университет  имени Н.Э. Баумана»***  ***(МГТУ им. Н.Э. Баумана)*** |

КАФЕДРА

Системы обработки информации и управления

**Отчет**

**по лабораторной работе № 3**

# Дисциплина: Технологии машинного обучения

Студент гр. **РТ5-61Б \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ А.Г. Обрезкова**

(Подпись, дата) (И.О. Фамилия)

Преподаватель  **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Ю.Е. Гапанюк**

(Подпись, дата) (И.О. Фамилия)

Москва, 2020

## Обработка пропусков в данных, кодирование категориальных признаков, масштабирование данных.

**Цель лабораторной работы:** изучение способов предварительной обработки данных для дальнейшего формирования моделей.

### Задание:

1. Выбрать набор данных (датасет), содержащий категориальные признаки и пропуски в данных. Для выполнения следующих пунктов можно использовать несколько различных наборов данных (один для обработки пропусков, другой для категориальных признаков и т.д.)
2. Для выбранного датасета (датасетов) на основе материалов [лекции](https://github.com/ugapanyuk/ml_course_2020/blob/master/common/notebooks/missing/handling_missing_norm.ipynb) решить следующие задачи:

* обработку пропусков в данных;
* кодирование категориальных признаков;
* масштабирование данных.











