|  |  |
| --- | --- |
|  | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

**Факультет «Информатика и системы управления»**

**Кафедра ИУ5 «Системы обработки информации и управления»**

Отчет по лабораторной работе

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

## по дисциплине «Технологии машинного обучения»

Выполнила:

студентка группы РТ5-71 Каркавина А.А.

подпись, дата

Проверил:

к.т.н., доц., Гапанюк Ю.Е.

подпись, дата

2019 г.

**Цель работы:**

изучение способов предварительной обработки данных для дальнейшего формирования моделей.

**Описание задания:**

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/blob/master/common/notebooks/missing/handling_missing_norm.ipynb) решить следующие задачи:

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

**Текст программы с экранными формами:**























