|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **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 №3

по дисциплине «Методы машинного обучения»

по теме «Обработка признаков. Часть 2»

Выполнила:

студентка группы № ИУ5-21М

Саватюгина Ю.Б.

подпись, дата

Проверил:

Гапанюк Ю.Е.

подпись, дата

2021 г.

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

**Задание.**

1. Выбрать набор данных (датасет), содержащий категориальные и числовые 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 лекций решить следующие задачи:
3. Масштабирование признаков (не менее чем тремя способами);
4. Обработка выбросов для числовых признаков (по одному способу для удаления выбросов и для замены выбросов);
5. Обработка по крайней мере одного нестандартного признака (который не является числовым или категориальным);
6. Отбор признаков:

* Один метод из группы методов фильтрации (filter mothods);
* Один метод из группы методов обертывания (wrapper methods);
* Один метод из группы методов вложений (embedded methods).

**Текст программы и экранные формы с примерами работы программы:**

                                              