|  |
| --- |
| МИНОБРНАУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение высшего образования **«МИРЭА − Российский технологический университет»**  **РТУ МИРЭА** |

**Институт информационных технологий (ИИТ)**

**Кафедра практической и прикладной информатики (ППИ)**

**ОТЧЕТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ**

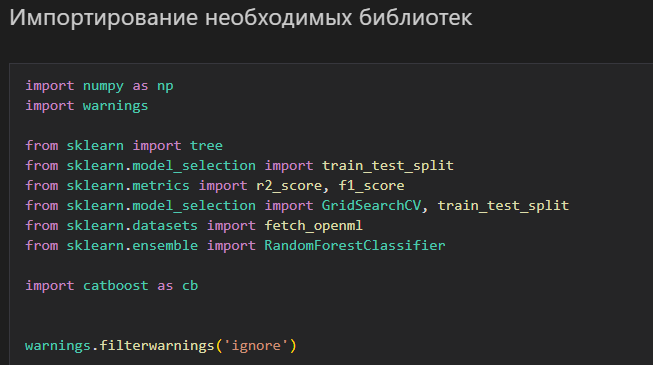
по дисциплине «Технологии и инструментарий анализа больших данных»

**Практическое занятие № 7**

|  |  |  |  |
| --- | --- | --- | --- |
| Студент группы ИНБО-01-17 | *ИНБО-03-20, Першутов Н.С.* | (подпись) | |
| Преподаватель | *Парамонов В.В* | (подпись) | |
| Отчет представлен | «\_\_\_»\_\_\_\_\_\_\_\_2023г. | |  | |

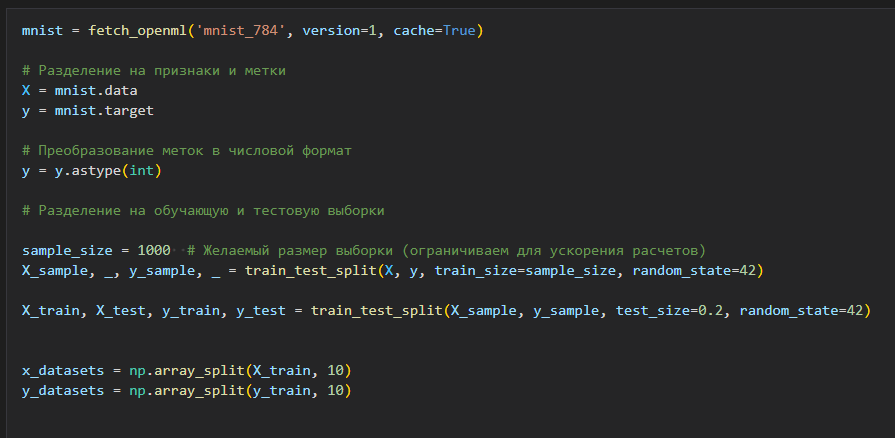
Москва 2023г.

**Результат работы:**

****

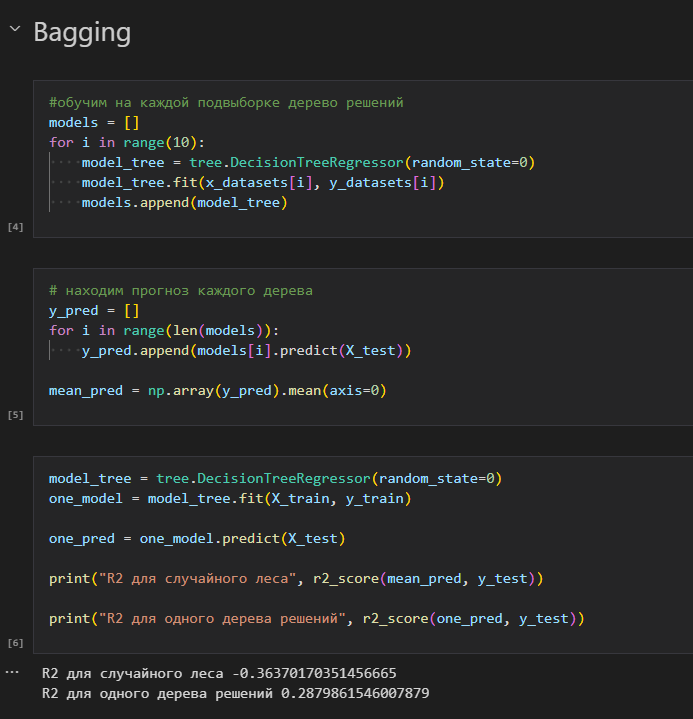
**Рисунок – Импорт необходимых библиотек**

**Задание 1**

****

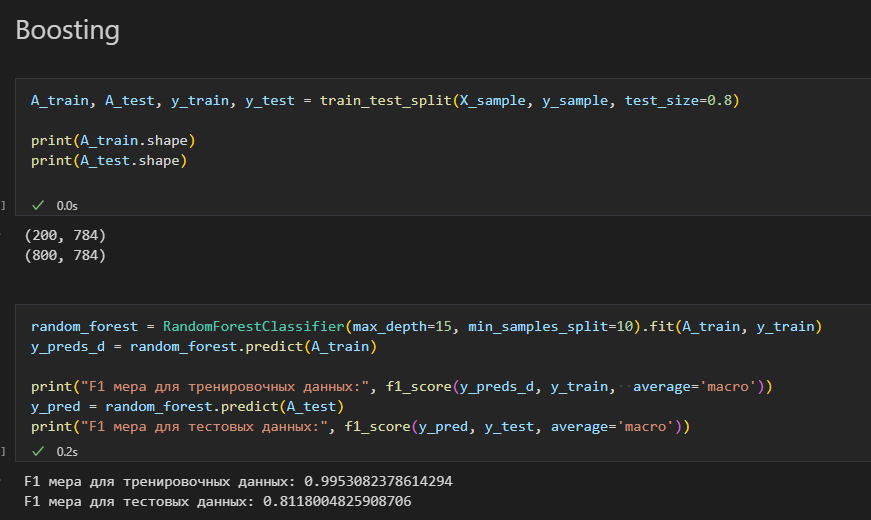
**Рисунок 1 – Импортирование и предобработка данных**

**Задание 2**

****

**Рисунок 2.1 – Реализация баггинга**

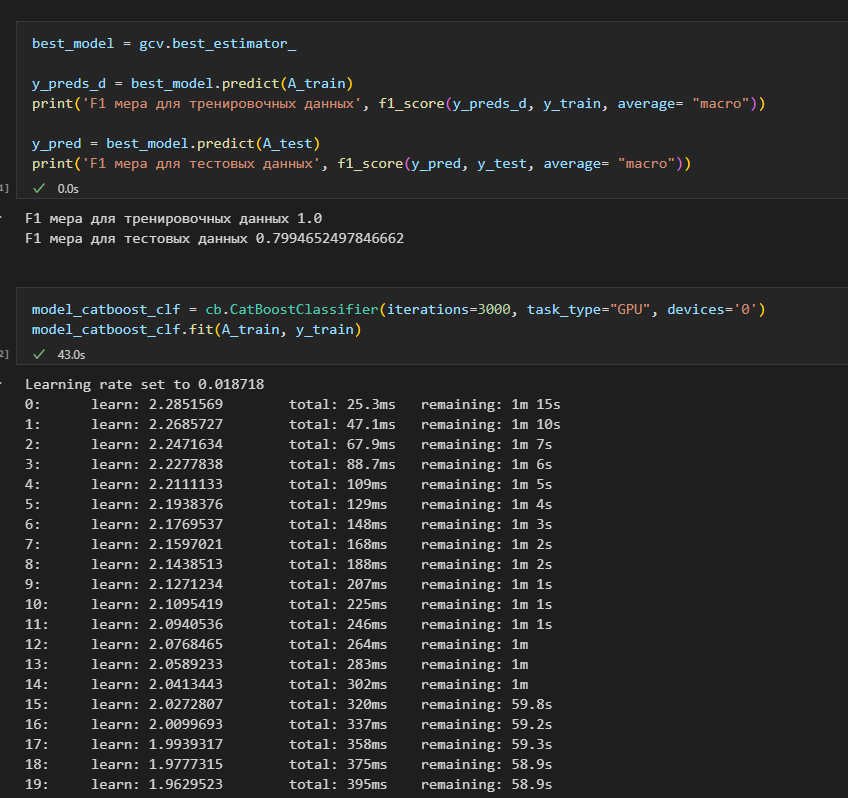
**Задание 3**

****

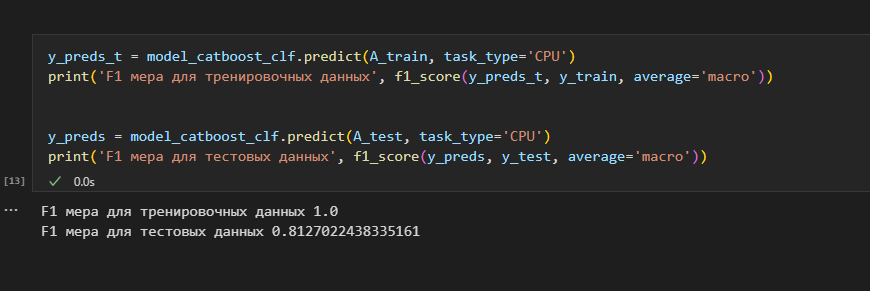
**Рисунок 3.1 – F1 мера для тестовых и тренировочных данных (случайный лес)**

****

**Рисунок 3.2 – Исправление проблемы случайного леса**

****

**Рисунок 3.3 – Бустинг**

****

**Рисунок 3.4 – F1 мера для тестовых и тренировочных данных (бустинг)**

**Задание 4**

Исходя из проделанной работы мы можем сравнить эти алгоритмы по двум параметрам:

* Время работы
* Качество моделей

Bagging – работает быстрее, но имеет меньшее качество моделей

Boosting – работает медленнее, требуя при этом больше вычислительных мощностей. Качество моделей гораздо выше