|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | |  |  |  | | МИНОБРНАУКИ РОССИИ | | | | Федеральное государственное бюджетное образовательное учреждение  высшего образования  **«МИРЭА – Российский технологический университет»**  **РТУ МИРЭА** | | | | |
| Институт информационных технологий (ИИТ) | |
| Кафедра Прикладной Математики (ПМ) | |

|  |  |  |  |
| --- | --- | --- | --- |
| **ОТЧЁТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ №7** | | | |
| **по дисциплине «Технологии и инструментарий анализа больших данных»** | | | |
|  | | | |
| Выполнил студент группы ИКБО-14-20 | | Вежновец Ф. Ю. | |
|  | |  | |
| Принял: асистент | | Горячев А. А. | |
| Практические работы выполнены | «\_\_\_»\_\_\_\_\_\_\_2023 г. | | (подпись студента) | |
| «Зачтено» | «\_\_\_»\_\_\_\_\_\_\_2023 г. | | (подпись руководителя) | |
|  |  | |  | |

Москва 2023

Задание 1

Найти данные для задачи классификации или для задачи регрессии.

Реализация

Результат работы программы и код предоставлена на рисунке 1.

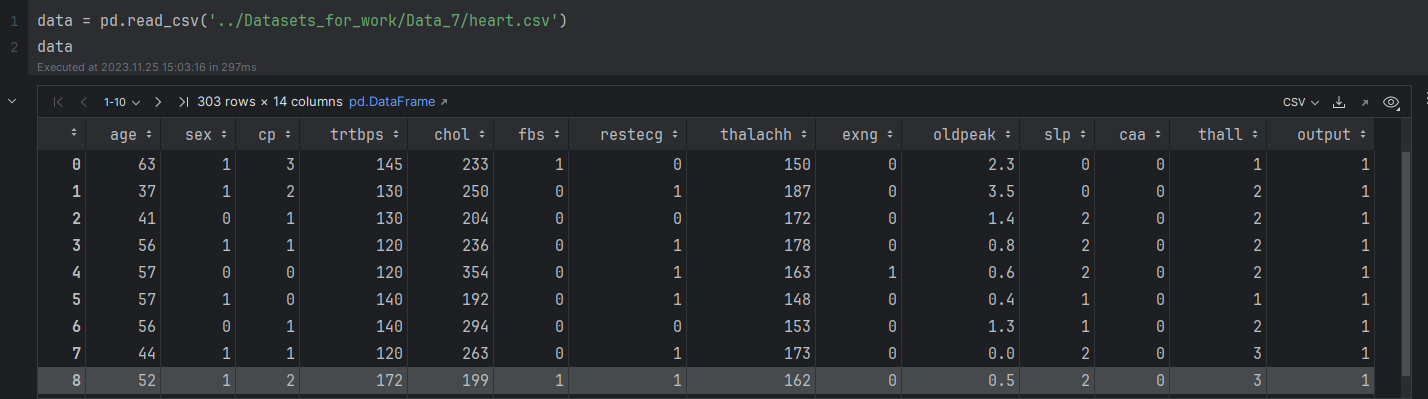


Рисунок 1 – Импорт данных

Задание 2

Реализовать баггинг.

Реализация

Результат работы программы и код предоставлена на рисунке 2 – 3.

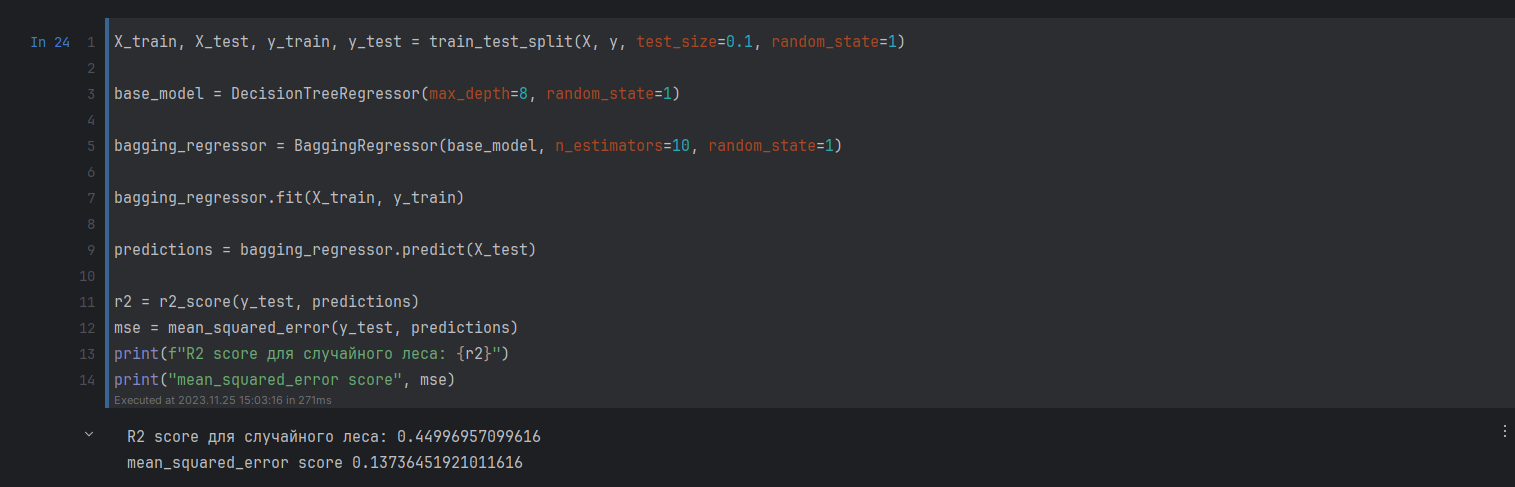


Рисунок 2 – Баггинг

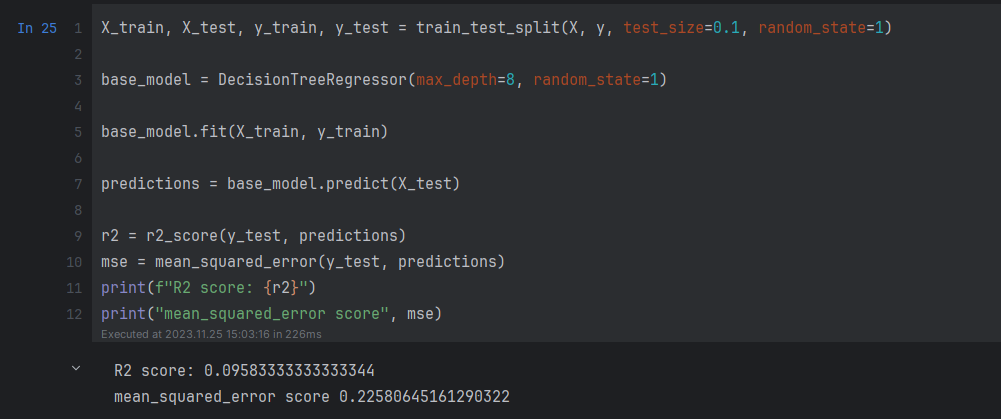


Рисунок 3 – Дерево решений

Задание 3

Реализовать бустинг на тех же данных, что использовались для баггинга.

Реализация

Результат работы программы и код предоставлена на рисунке 4 – 5.

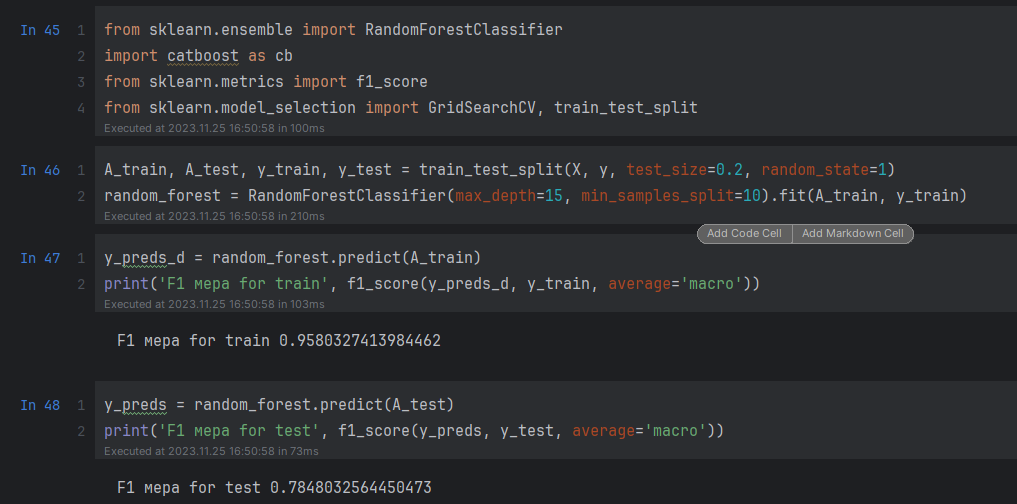


Рисунок 4 – Рандомный лес

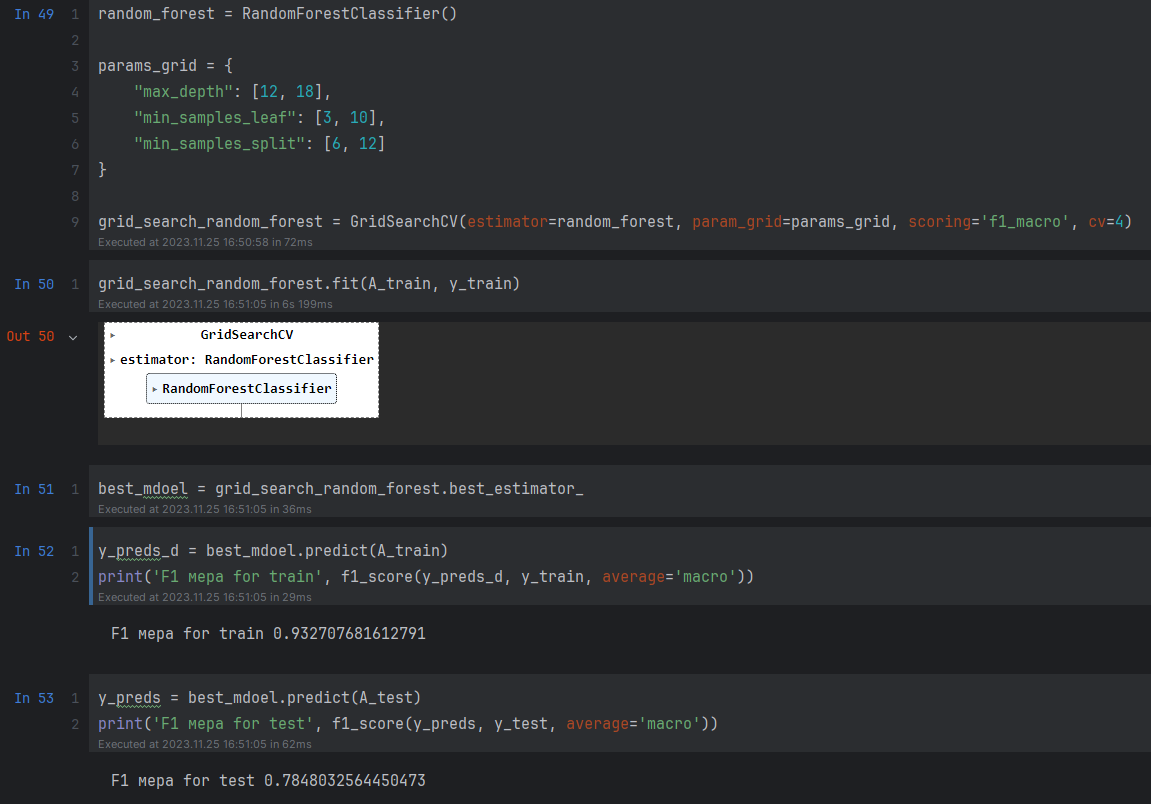


Рисунок 5 – Лучший рандомный лес

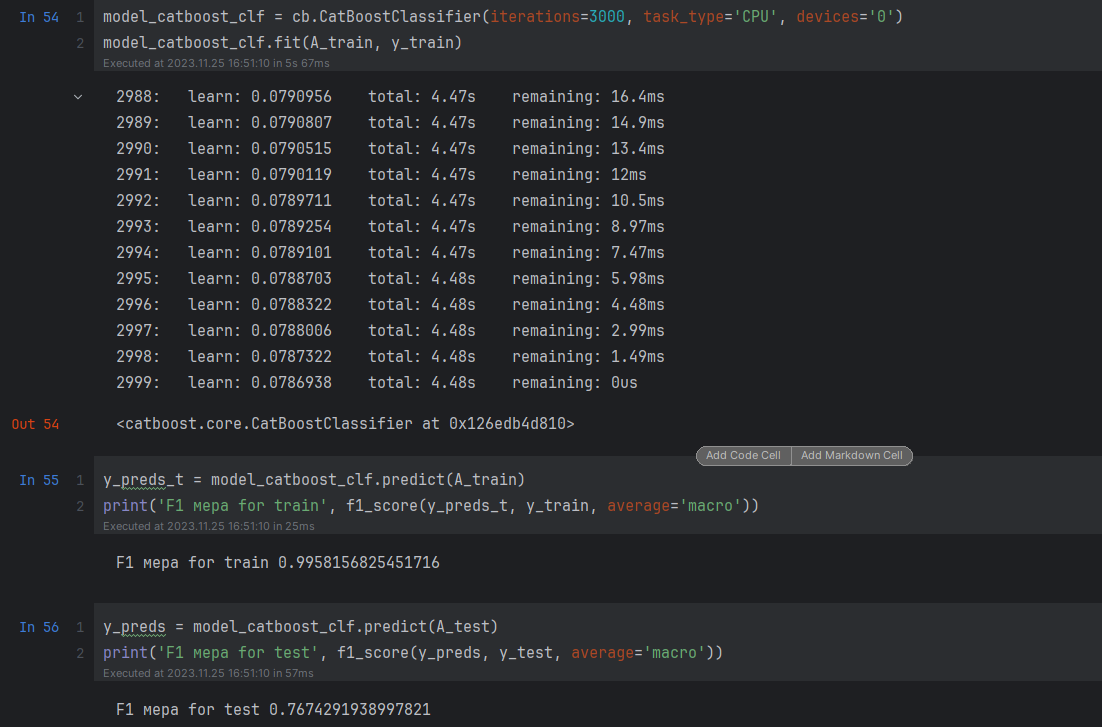


Рисунок 6 – Бустинг

Задание 4

Сравнить результаты работы алгоритмов (время работы и качество моделей). Сделать выводы.

Реализация

Результат работы программы (Таблица 1):

Таблица 1 – Результаты

|  |  |  |
| --- | --- | --- |
| Метод | Время (секунд) | Точность |
| Bagging | 0.027 | R2 = 0.44 |
| DecisionTree | 0.002 | R2 = 0.09 |
| RandomForest | 0.200 | F1 = 0.76 |
| RandomForest  (выбор лучшей модели) | 5.570 | F1 = 0.76 |
| CatBoost | 5.214 | F1 = 0.76 |