|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | |  |  |  | | 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 №5** | | | |
| **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 классификации. Предобработать данные, если это необходимо.

Реализация

Результат работы программы и код предоставлена на рисунке 1 – 20-.

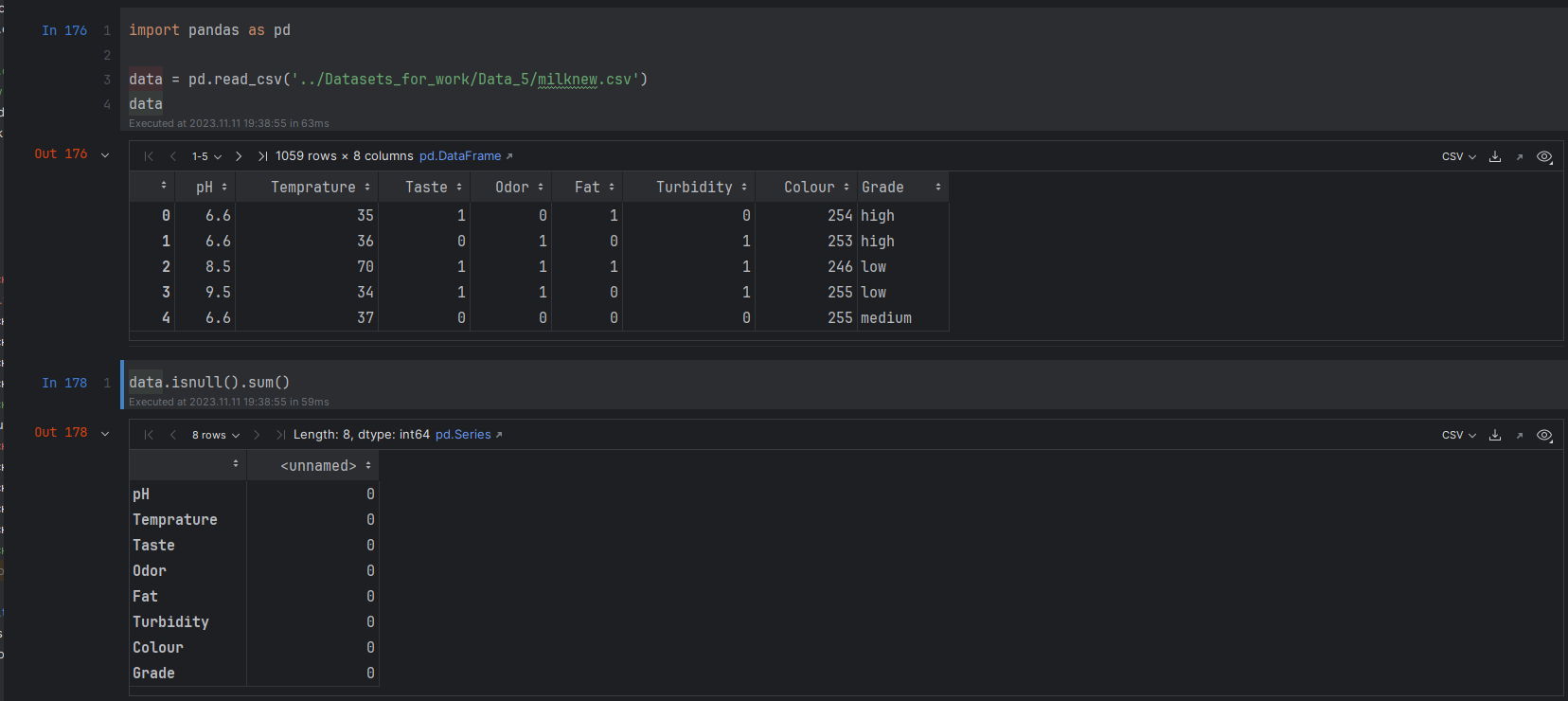


Рисунок 1 – Импорт данных

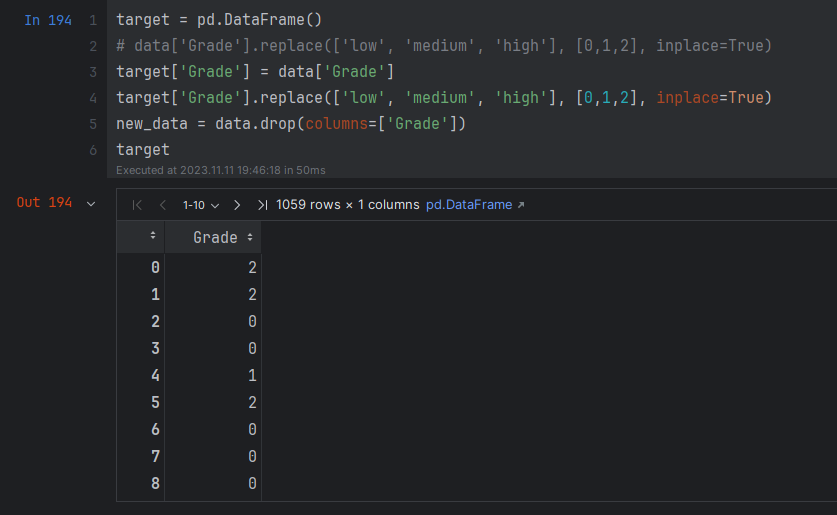


Рисунок 2 – Предобработка данных

Задание 2

Изобразить гистограмму, которая показывает баланс классов. Сделать выводы.

Реализация

Результат работы программы и код предоставлена на рисунке 3 - 4.

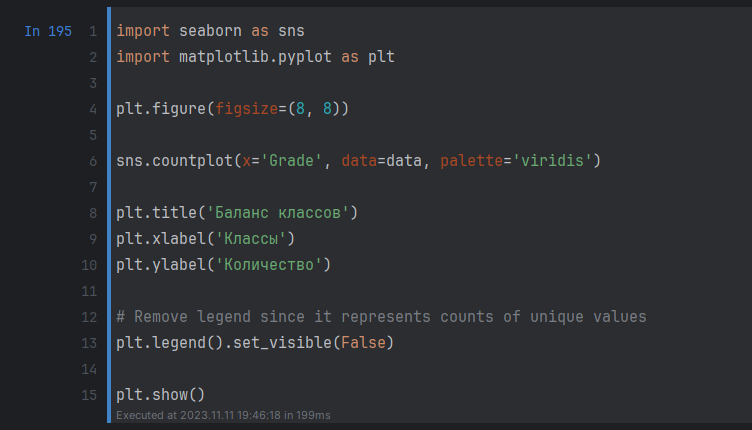


Рисунок 3 – Гистограмма код

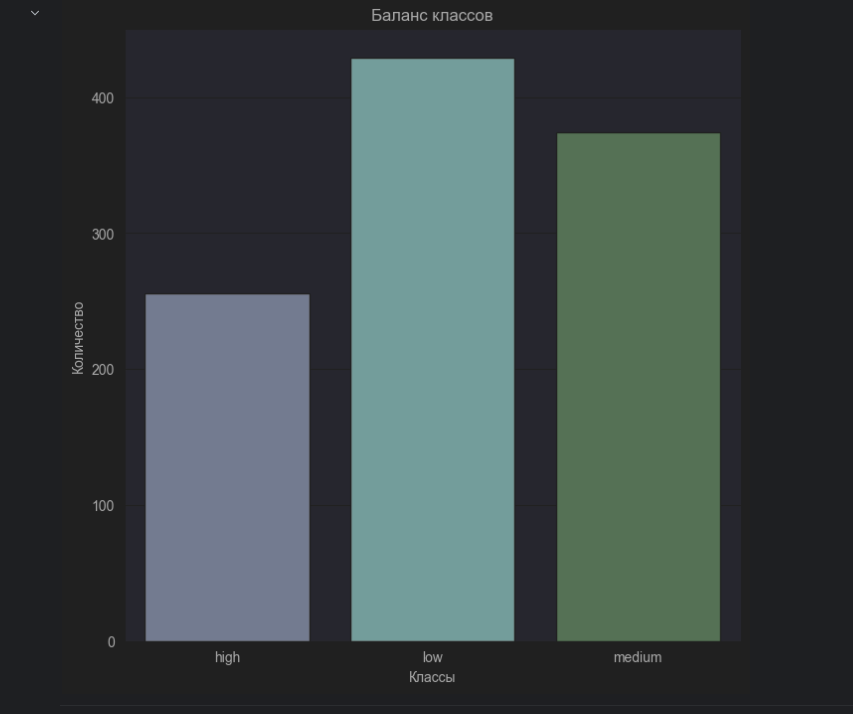


Рисунок 64 – Гистограмма

**Вывод:** самые распространённые объёмы оперативной памяти 8, 32, 16 гигабайт оперативной памяти.

Задание 3

Разбить выборку на тренировочную и тестовую. Тренировочная для обучения модели, тестовая для проверки ее качества.

Реализация

Результат работы программы и код предоставлена на рисунке 5.

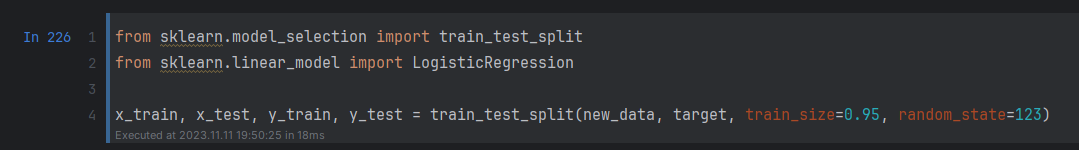


Рисунок 5 – Выборки

Задание 4

Применить алгоритмы классификации: логистическая регрессия, SVM, KNN. Построить матрицу ошибок по результатам работы моделей (использовать confusion\_matrix из sklearn.metrics).

Реализация

Результат работы программы и код предоставлена на рисунках 6 - .

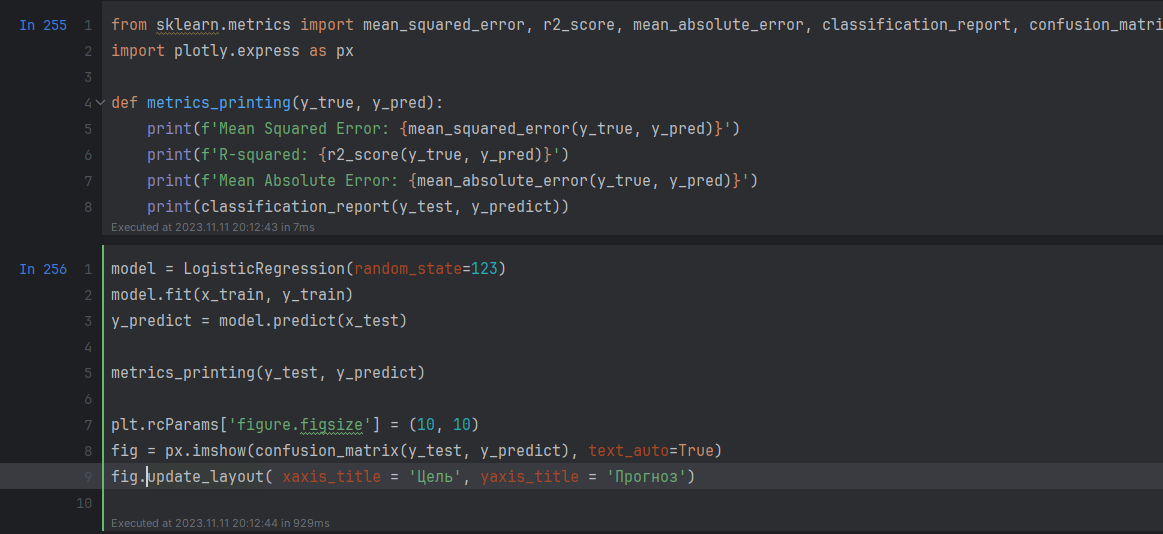


Рисунок 6 – Логистическая регрессия

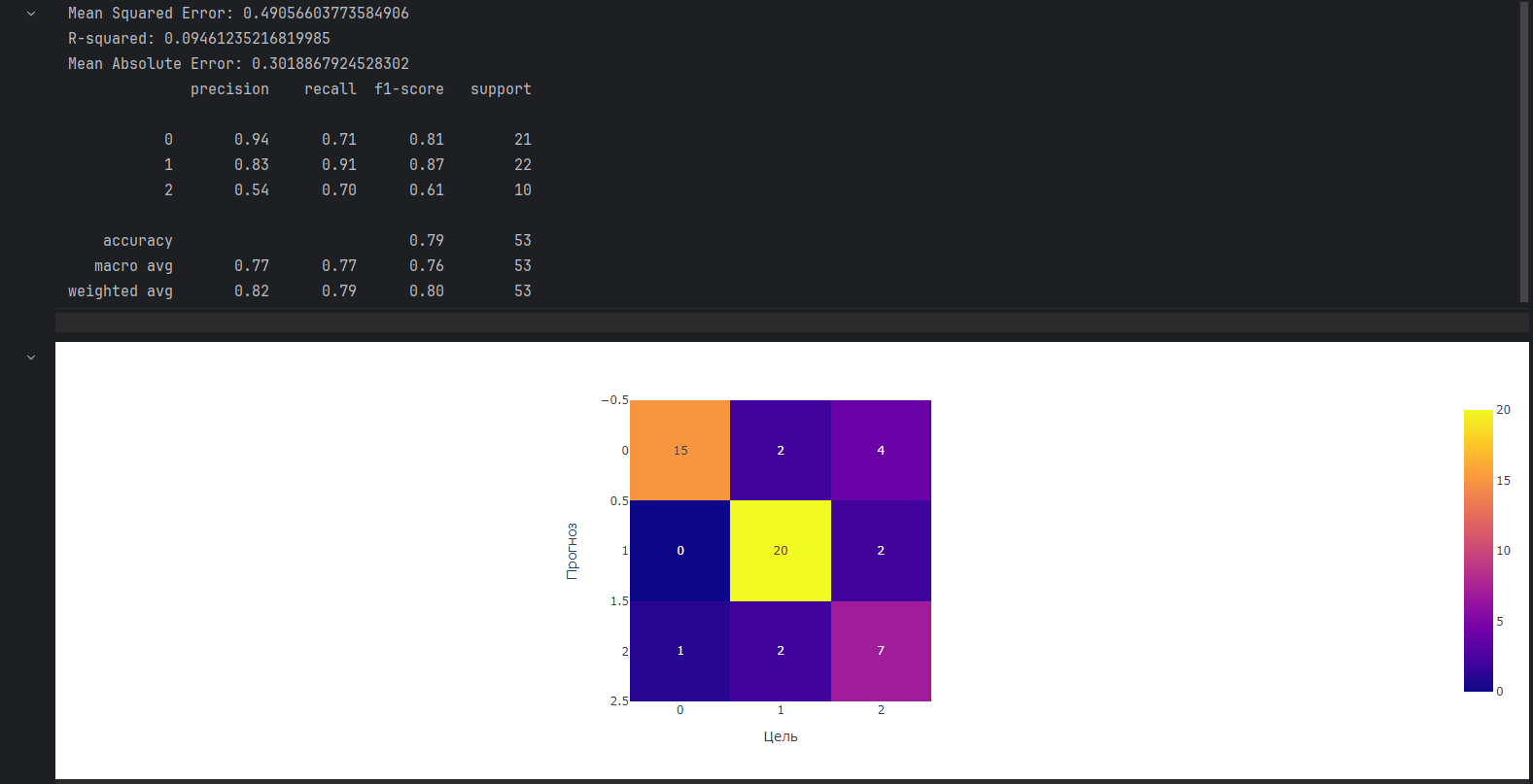


Рисунок 7 – Логистическая регрессия

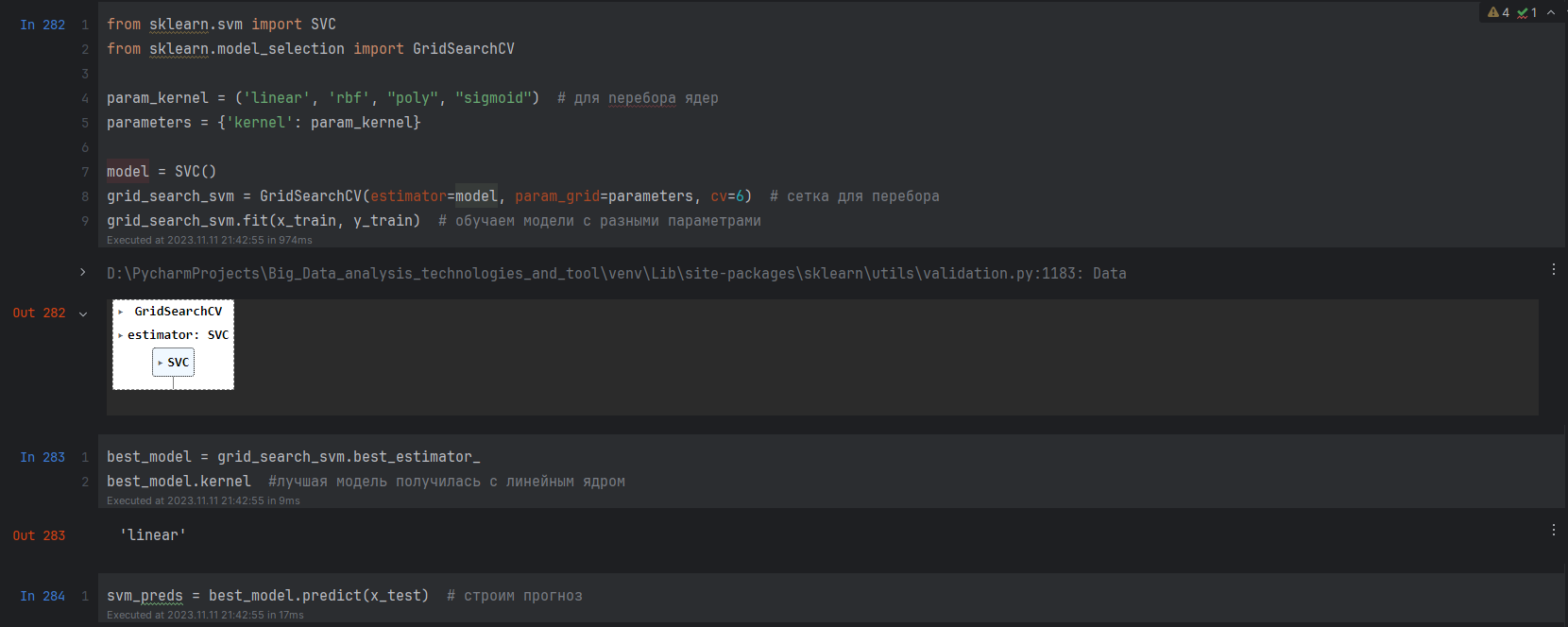


Рисунок 8 – SVC



Рисунок 9 – SVC

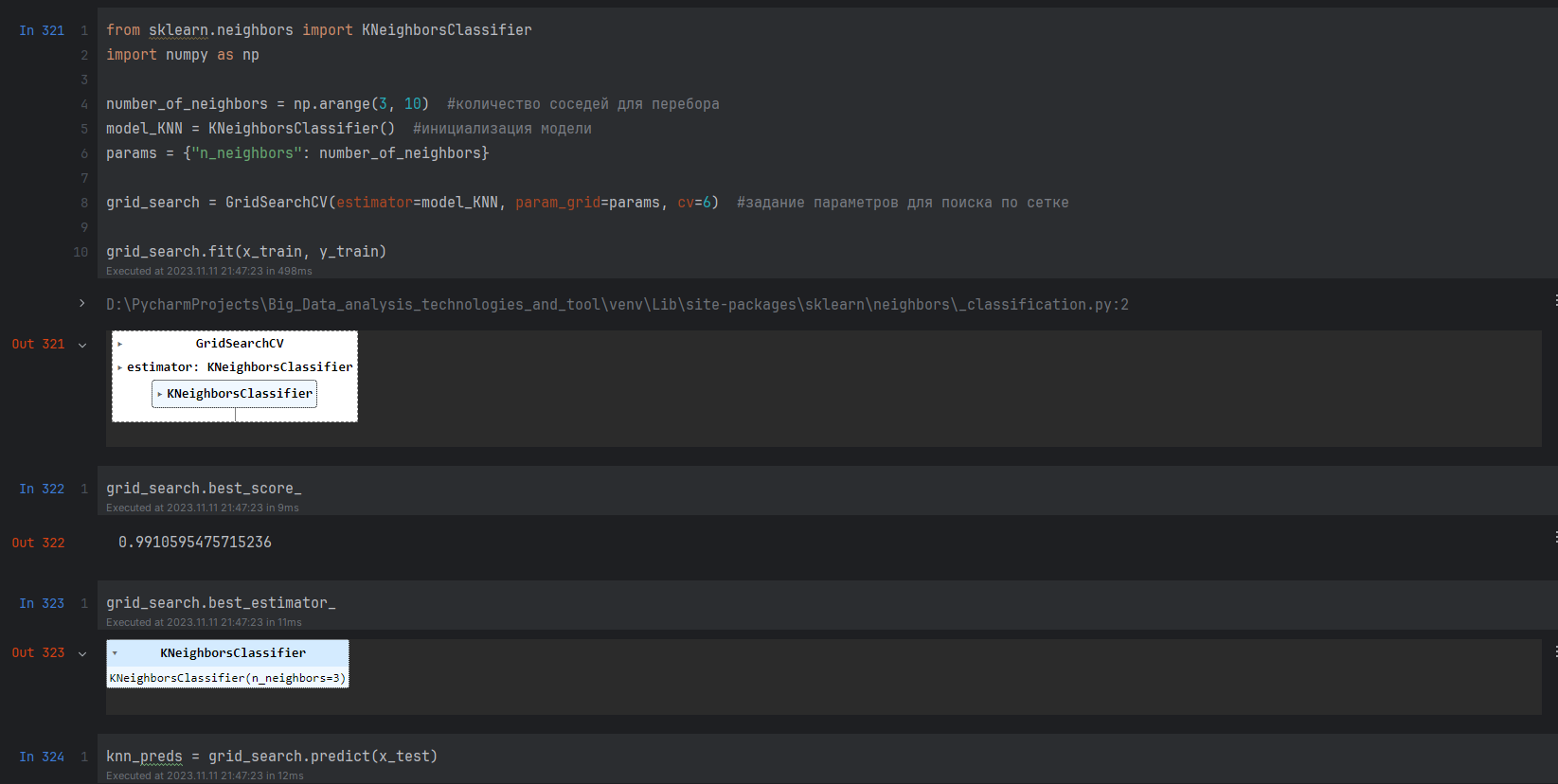


Рисунок 10 – SVC



Рисунок 11 – SVC