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

**Институт информационных технологий (ИИТ)**

**Кафедра прикладной математики (ПМ)**

**ОТЧЕТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ**

по дисциплине «Технологии и инструментарий анализа больших данных»

**Практическое занятие №5**

|  |  |  |  |
| --- | --- | --- | --- |
| Студент группы ИНБО-01-17 | *ИНБО-07-21, Веригин Никита Алексеевич* | (подпись) | |
| Преподаватель | *Царёв Роман Юрьевич* | (подпись) | |
| Отчет представлен | «\_\_\_»\_\_\_\_\_\_\_\_2024 г. | |  | |

Москва 2024 г.

Введение

Цели работы:

1. . Найти данные для классификации. Данные в группе повторяться не должны. Предобработать данные, если это необходимо.
2. Изобразить гистограмму, которая показывает баланс классов. Сделать выводы.
3. Разбить выборку на тренировочную и тестовую. Тренировочная для обучения модели, тестовая для проверки ее качества.
4. Применить алгоритмы классификации: логистическая регрессия, SVM, KNN. Построить матрицу ошибок по результатам работы моделей (использовать confusion\_matrix из sklearn.metrics).
5. Сравнить результаты классификации, используя accuracy, precision, recall и f1-меру (можно использовать classification\_report из sklearn.metrics). Сделать выводы.
6. Оформить отчет о проделанной работе.

ХОД РАБОТЫ

Код программ и результаты их выполнения представлены на Рисунках 1-7.

Изображение выглядит как текст, снимок экрана, программное обеспечение, Мультимедийное программное обеспечение

Автоматически созданное описание

**Рисунок 1 *—* Код решения задачи (ч.1)**

**Изображение выглядит как текст, снимок экрана, программное обеспечение, диаграмма

Автоматически созданное описание**

**Рисунок 2 *—* Код решения задачи (ч.2)**

**Изображение выглядит как текст, снимок экрана, программное обеспечение, Шрифт

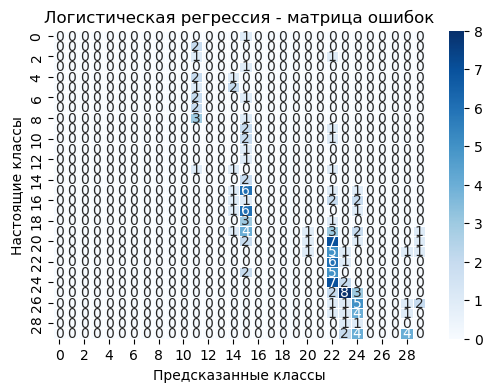
Автоматически созданное описание**

**Рисунок 3 *—* Код решения задачи (ч.3)**

**Изображение выглядит как текст, снимок экрана, программное обеспечение

Автоматически созданное описание**

**Рисунок 4 *—* Код решения задачи (ч.4)**

****

**Изображение выглядит как текст, снимок экрана, линия, Параллельный

Автоматически созданное описание**

**Рисунок 5 *—* Код решения задачи (ч.5)**

**Изображение выглядит как текст, снимок экрана, линия, число

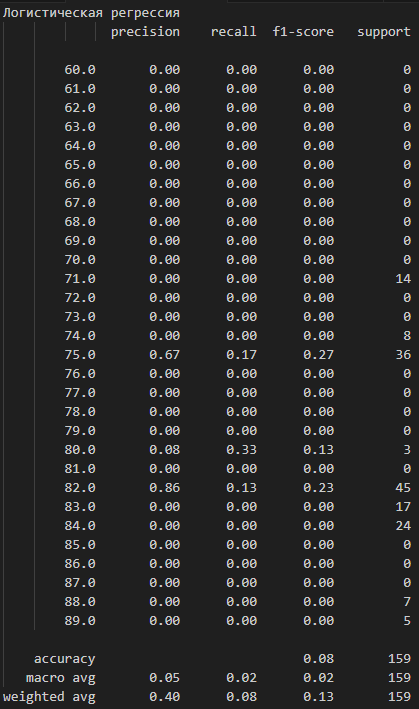
Автоматически созданное описание**

**Рисунок 6 *—* Код решения задачи (ч.6)**

**Изображение выглядит как текст, снимок экрана, Шрифт, программное обеспечение

Автоматически созданное описание**

**Рисунок 7 *—* Код решения задачи (ч.7)**

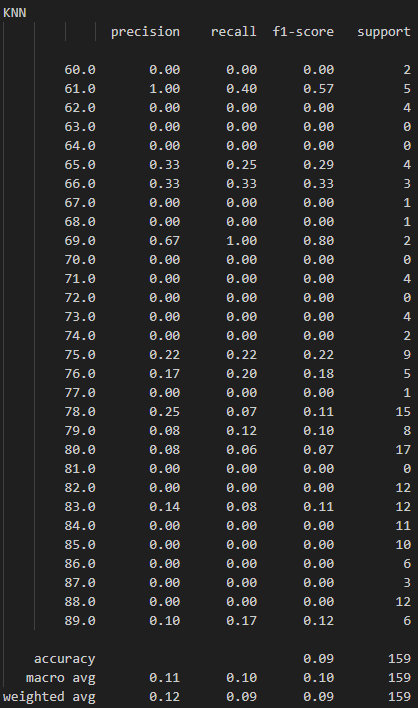
****

**Рисунок 8 — Сравнение результатов классификации**

**Изображение выглядит как текст, снимок экрана, меню

Автоматически созданное описание**

**Рисунок 9 — Сравнение результатов классификации**

****

**Рисунок 10 — Сравнение результатов классификации**

ЗАКЛЮЧЕНИЕ

В ходе выполнения практической работы была изучена задача классификации, метрики оценки качества модели для задачи классификации, алгоритмы классификации (логистическая регрессия, SVM, KNN) и матрица ошибок.