Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования

«Южно-Уральский государственный университет

(национальный исследовательский университет)»

Высшая школа электроники и компьютерных наук

Кафедра «Системное программирование»

ОТЧЁТ

по лабораторной работе №3 на тему «Байесовская классификация»

Вы	полнил	
Сту	дент гр	уппы КЭ-120
		Д.А. Снегирева
«		2020 г.
Ema	ail: dash	asneg@mail.ru

ЗАДАНИЕ

Выполните классификацию набора данных с помощью Байесовской классификации, варьируя соотношение мощностей обучающей и тестовой выборок от 60%:40% до 90%:10% с шагом 5%. Вычислите показатели качества классификации: аккуратность (accuracy), точность (precision), полнота (recall), F-мера. Выполните визуализацию полученных результатов в виде диаграмм.

СОДЕРЖАНИЕ

ВАДАНИЕ	2
СОДЕРЖАНИЕ	3
І КРАТКИЕ СВЕДЕНИЯ О НАБОРАХ ДАННЫХ И СРЕДСТВАХ РЕАЛИЗАЦИІ	И 4
2 ВИЗУАЛИЗАЦИЯ РЕЗУЛЬТАТОВ	5

1 КРАТКИЕ СВЕДЕНИЯ О НАБОРАХ ДАННЫХ И СРЕДСТВАХ РЕАЛИЗАЦИИ

В данной работе был использован набор данных Wine Data Set (https://archive.ics.uci.edu/ml/datasets/wine), содержащий результаты химического анализа вин, выращенных в одном и том же регионе Италии, и состоящий из 13 атрибутов.

В качестве одного из средств реализации была использована библиотека scikit-learn, простое средство для анализа данных.

Репозиторий задания: https://github.com/DasHaSneg/BigDataMiningCourse

Каталог задания: 3 bayesian_classification

2 ВИЗУАЛИЗАЦИЯ РЕЗУЛЬТАТОВ

Для визуализации были вычислены показатели качества классификации: аккуратность (accuracy), точность (precision), полнота (recall) и F-мера для разных значений соотношений мощностей обучающей и тестовой выборок (от 60%:40% до 90%:10% с шагом 5%). Для каждого показателя были построены диаграммы, которые представлены на рис. 1.

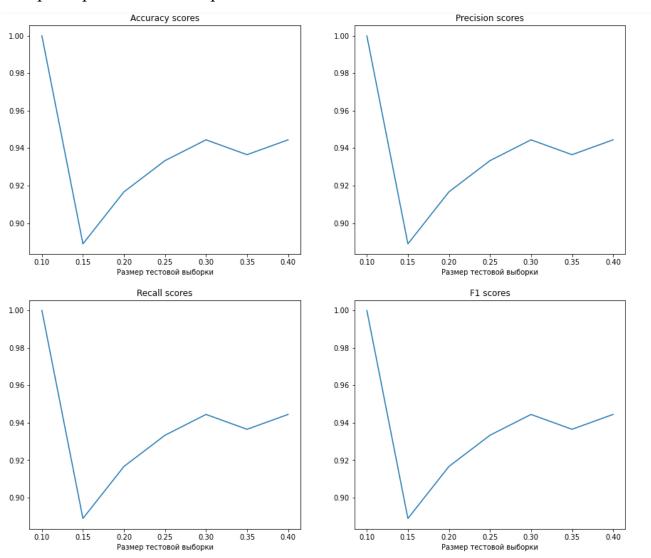


Рис. 1. Результаты

Результаты для каждой метрики идентичны. Исходя из значений, можно сделать вывод, что лучшее достигается при соотношении мощностей обучающей и тестовой выборок 90%:10%.