

Национальный исследовательский университет  
компьютерных технологий, механики и оптики

Факультет ПИиКТ

Системы искусственного интеллекта  
Лабораторная работа №3  
Вариант чётный

Работу выполнил: Зубахин Д. С.

Группа: Р33312

Преподаватель: Королёва Ю. А.

Санкт-Петербург  
2022 год

## 1 Описание задания

1. Датасет с данными про оценки студентов инженерного и педагогического факультетов (для данного датасета нужно ввести метрику: студент успешный/неуспешный на основании грейда)
2. Отобрать случайным образом  $\sqrt{n}$  признаков.
3. Реализовать без использования сторонних библиотек построение дерева решений (numpy и pandas использовать можно).
4. Провести оценку реализованного алгоритма с использованием Accuracy, precision и recall
5. Построить AUC-ROC и AUC-PR.

## 2 Реализация.

### 2.1 Подсчет количества признаков

Общее количество признаков в датасете - 22. Тогда требуемое количество:

$$k = \lceil \sqrt{32} \rceil = 5$$

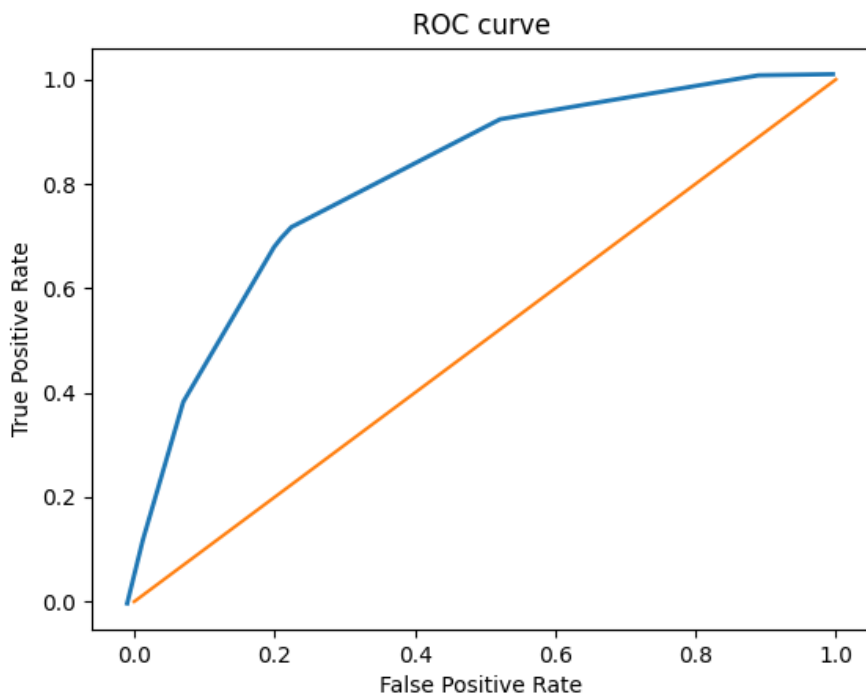
### 2.2 Реализация и вывод формул

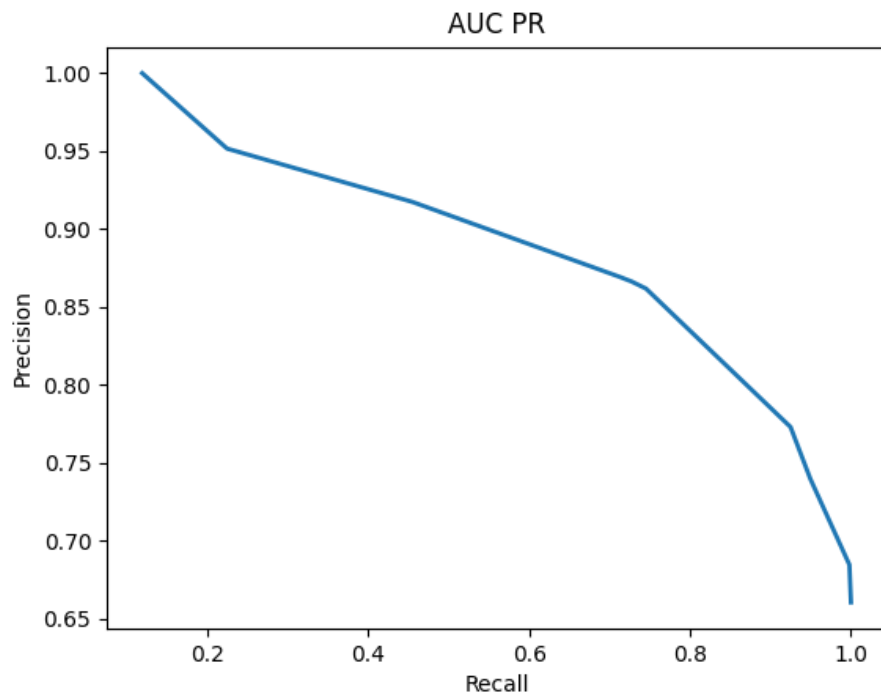
Обратиться к пункту 3.

### 2.3 Оценка реализованного алгоритма с использованием accuracy, precision и recall

Accuracy: 0.8203865566908778  
Precision: 0.7729890764647468  
Recall: 0.9249049429657795

### 2.4 AUC-ROC и AUC-PR





### 3 Исходный код

<https://github.com/dmittrey/Artificial-intelligence-lab3>

### 4 Вывод.

В ходе выполнения лабораторной работы я познакомился с алгоритмом C4.5 и реализовал его на питоне. Также вспомнил, что такое accuracy, precision, recall, а также разобрался, как строить графики AUC-ROC, AUC-PR.