

Московский авиационный институт
(Национальный исследовательский университет)

Институт: «Информационные технологии и прикладная математика»

Кафедра: 806 «Вычислительная математика и программирование»

Лабораторная работа № 1
по курсу «Искусственный интеллект»

Студент:	Егорова А. В.
Группа:	М8О-307Б-18
Оценка:	
Дата:	

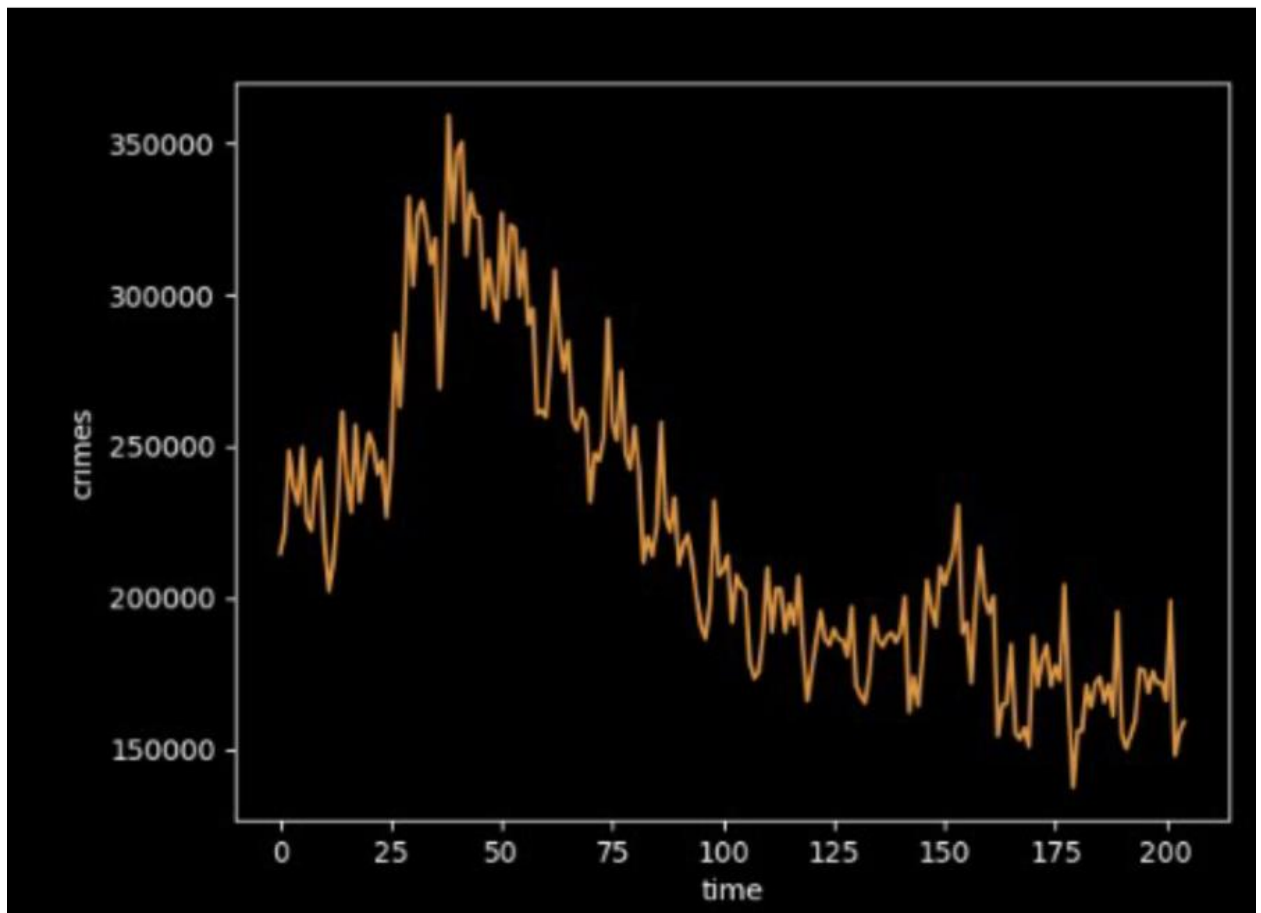
Москва, 2021

Постановка задачи:

Найти себе набор данных (датасет), для следующей лабораторной работы, и проанализировать его. Выявить проблемы набора данных, устранить их. Визуализировать зависимости, показать распределения некоторых признаков. Реализовать алгоритмы К ближайших соседа с использованием весов и Наивный Байесовский классификатор и сравнить с реализацией библиотеки sklearn.

Датасет: статистика преступлений в РФ за последнее время.

График зависимости числа преступлений за годы



KNN:

Вычисляем расстояние до каждого из объектов обучающей выборки, отбираем k объектов обучающей выборки, расстояние до которых минимально

Во внимание принимаем удаленность классов от нового значения. Для каждого класса j определяется оценка близости. У какого класса значение близости выше, тот класс и присваивается новому объекту.

Наивный Байесовский классификатор:

Он предполагает независимость параметров

Сравнение с sklearn

```
my knn:      [0, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]
my bayes:    [0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1]
true val:    [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]
knn accuracy = 0.8260869565217391 bayes accuracy = 0.8260869565217391

sklearnKnn:  [0 0 0 0 1 0 1 0 1 1 0 1 1 1 1 1 1 1 1 1 1]
accuracy = 0.8260869565217391

sklearnNB:   [0 0 0 0 0 0 1 0 1 1 0 1 0 1 1 1 1 1 1 1 1]
accuracy = 0.8260869565217391
```

Точности алгоритмов и sklearn совпадают