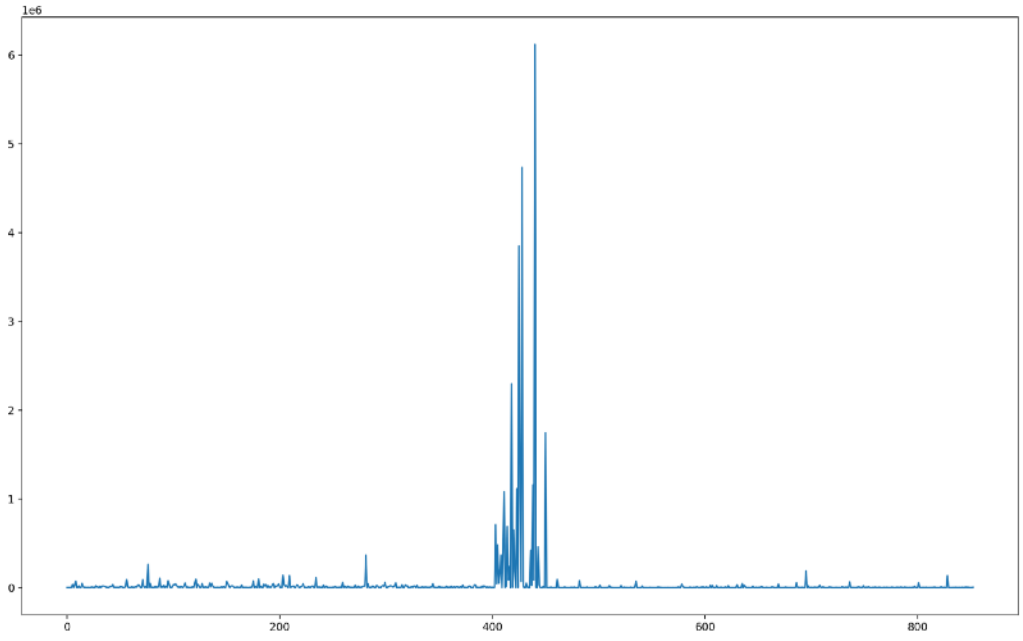
Классификация вредоносного поведения.

1) Оценка важности различных признаков

После обработки данных файла behavior.csv создается 2 файла Harmful.txt и Unharmful.txt в этих файлах через запятую указывается количество i-ых действий (i соответствует названию действия в i-м столбце 1 строки (первый столбец не в счет так как в нем id пользователя или тп))

Далее по данным этих файлов строятся графики (графики строил при помощи matplotlib на Питоне)

График вредоносного поведения:



Из графика видно, что самые частые действия во вредоносном поведении имеют индексы от 402 до 450 (за исключением пары сильных проседов)

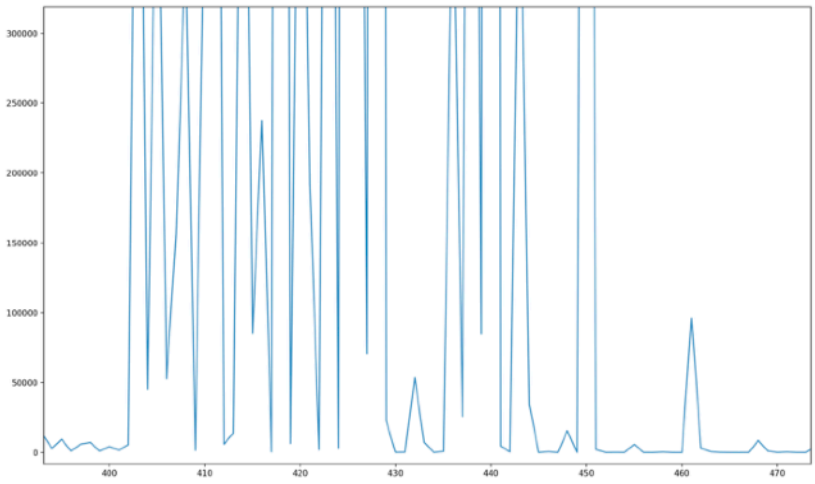
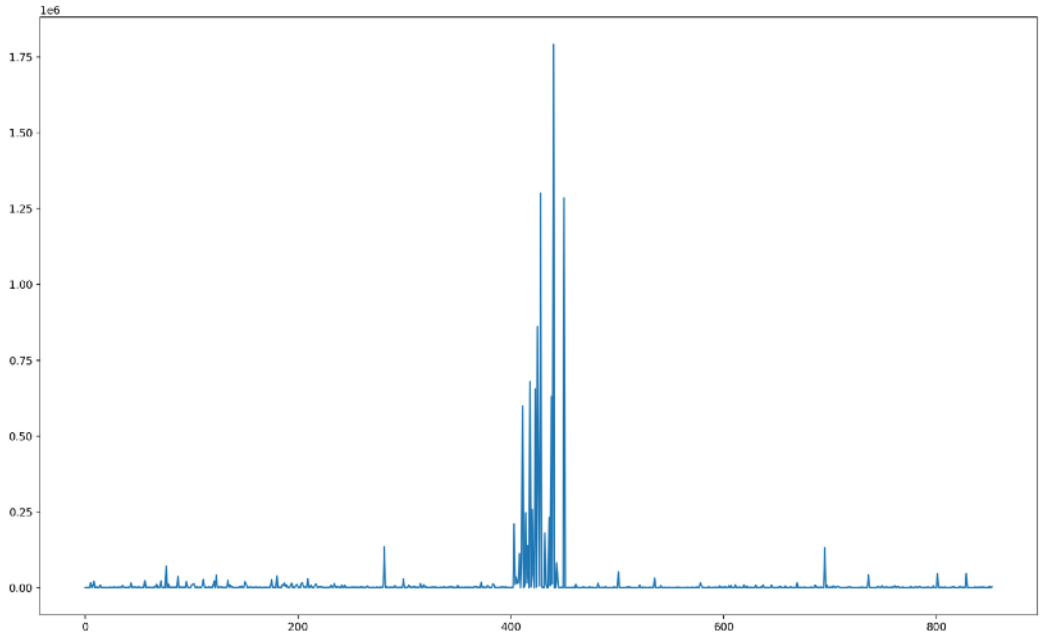
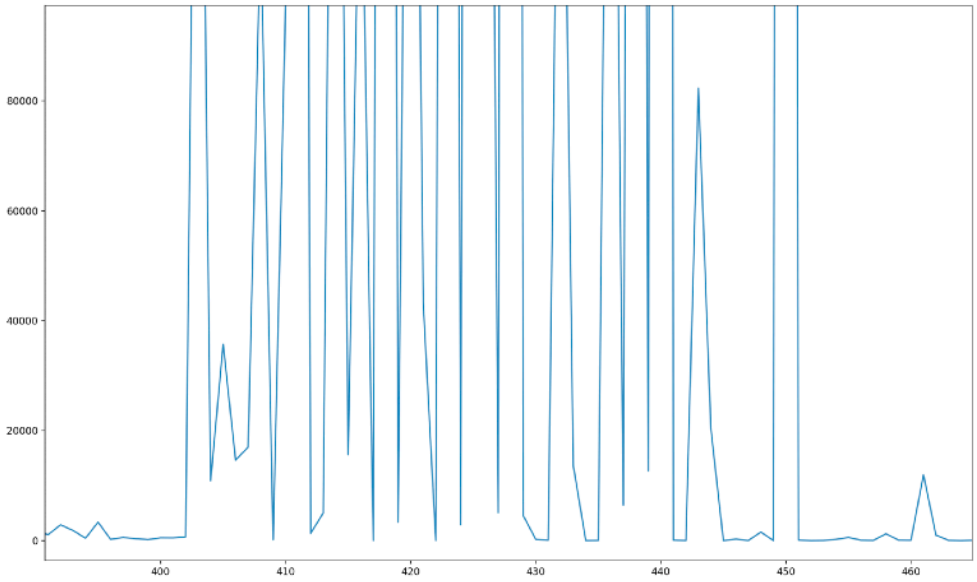


График не вредоносного поведения:



Аналогичный диапазон наиболее говорящих признаков и у не вредоносного поведения



2) Реализация модели и оценка её точности.

Я решил осуществлять классификацию при помощи наивного Байесовского классификатора с классами «вредоносное поведение» и «не вредоносное поведение». Т.е. классификация происходит по формуле:

$$arg\max[P(Q_k)\prod_{i=1}^{n}P(x_i|Q_k)]$$

где x_i - действие пользователя соответствующее наименованию i-го столбца
Q_k - класс («вредоносное поведение» или «не вредоносное поведение»)

На обучающей выборке точность модели составила: **56,4 %**

Такая малая доля правильных предсказаний скорее всего связана с наивным предположением, что действия пользователя независимы друг от друга :(