СОФИЙСКИ УНИВЕРСИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“
ФАКУЛТЕТ ПО МАТЕМАТИКА И ИНФОРМАТИКА

КУРСОВ ПРОЕКТ
ПО ИЗКУСТВЕН ИНТЕЛЕКТ

Тема:

Класификатор на спам

Студенти:

Мартин Ивелинов Николаев, 2, 45634
Радослав Руменов Хубенов, 2, 45708

София, юни 2022 г.

1. Формулировка на задачата

Да се напише програма, която при подадено множество от имейли, ги класифицира като “спам” или “хам”.

2. Използвани алгоритми

Използваният алгоритъм е наивен Бейсов класификатор.

Наивните Бейсови класификатори са популярни статистически методи за филтриране на имейли. Типично използват множества от думи за да разпознават даден имейл като “спам” или “хам”, т.е. “нормален”.
Наивните Бейсови класификатори работят като обвързват използваните думи с вече класифицирани имейли, използвайки теоремата на Бейс, за да изчисли вероятността даден имейл да е “спам”.

$$\Pr(\text{spam}|\text{words}) = \frac{\Pr(\text{words}|\text{spam})\Pr(\text{spam})}{\Pr(\text{words})}$$

Източник https://en.wikipedia.org/wiki/Bayes%27_theorem

3. Описание на програмната реализация

Файлове: build_vocabulary.py, gen_frequency.py, naive_bayes.py
Класове: NaiveBayes
Функции: build_vocabulary(curr_email)

Файл: naive_bayes.py

Файлът съдържа имплементацията на класа “NaiveBayes”, който съдържа следните методи:

- init
  запазва броя имейли, броя features, броя класове (в случая 2)
- fit
  За всеки клас (“спам” и “не спам”) намира математическото средно, вариацията и пропорцията от съответния клас.
- predict

Изчислява шанса всеки имейл от подаденият data set да е спам, използвайки density_function.

- density_function

  Изчислява вероятността използвайки функцията за плътност на Гаус.

$$\det(2\pi\boldsymbol{\Sigma})^{-\frac{1}{2}}\ \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right),$$

Файл: gen_frequency.py

Кодът във файла създава два нови файла, първият от които представлява честотата на срещането на всяка дума от речника ни, а вторият - дали даден имейл е в реалност спам или хам.

Файл: build_vocabulary.py

Създава нов файл, който представлява хеш таблица от индексирани думи. Проверява всяка дума от всеки имейл дали е истинска дума от английския език и ако не се среща в таблицата ни, я добавя.

4. Примери, илюстриращи работата на програмната система

   python3 build_vocabulary.py

   Създава vocabulary.txt, който е речника от имейлите

   python3 gen_frequency.py

   Създава X.npy и y.npy, които съдържат честотата на срещане на всяка дума и класификацията за всеки един имейл.

   python3 naive_bayes.py

   Обучава модела и тества точността.

```
0.9149790502793296
```

5. Литература

https://en.wikipedia.org/wiki/Naive_Bayes_spam_filtering - 22.06

https://en.wikipedia.org/wiki/Multivariate_normal_distribution - 22.06