

Projet : filtre anti-spam

Ladeveze Quentin
Lallemand François

5 mai 2018

1 Utilisation du programme

Les sources de ce projet se trouvent dans le fichier `filtreAntiSpam.java`, à compiler par la commande :

```
javac filtreAntiSpam.java
```

La classe `filtreAntiSpam` se lance par les commande :

- `java filtreAntiSpam <test_folder> <number of spam in test folder> <number of ham in test folder>` Effectue un test avec les fichiers situés dans `<test_folder>`, sur le nombre de SPAM et de HAM indiqués. Le programme demande ensuite un nombre de SPAM et de HAM dans la base d'apprentissage et lance le test.
- `java filtreAntiSpam save <file where to store classifier > <number of spam> <number of ham>` Cette option effectue un apprentissage sur le nombre de SPAM et de HAM indiqués et enregistre le classifieur dans le fichier dont le nom est donné en paramètre.
- `java filtreAntiSpam filtre_ligne <source file for the classifier > <message> <HAM or SPAM>` Cette option effectue un apprentissage en ligne en se basant sur un classifieur existant.
- `java filtreAntiSpam <source file for the classifier > <message to test>` Cette option permet de tester un mail rapidement à partir d'un classifieur existant.

2 Répartition du travail

Nous nous sommes répartis le travail de la manière suivante : Nous avons réfléchi ensemble aux questions avant d'implémenter les réponses sur une seule machine.

3 Choix effectués

Afin de décider si un message est ou non un spam, plutôt que de calculer :

$$\prod_{j=1}^d (b^j SPAM)^{x^j} (1 - b^j SPAM)^{1-x^j}$$

Nous avons utilisé les logarithmes afin de calculer :

$$\sum_{j=1}^d \ln((b^j SPAM)^{x^j}) \ln((1 - b^j SPAM)^{1-x^j})$$

Les

$$x^j$$

et

$$1 - x^j$$

sont représentés sous forme de booléens, afin de simplifier les calculs du classifieurs en utilisant des conditions “if” plutôt que des calculs de puissance.

La fonction `lire_message` sépare le texte selon l’expression régulière “W+”. W correspond aux caractères qui ne sont ni des chiffres, ni des lettres. Cela permet de reconnaître les mots suivis d’un signe de ponctuation, comme “House,”.

4 Exemple d’exécution

```
java filtreAntiSpam basetest 100 200
Nombre de spams dans la base d’apprentissage : 200
```

```
Nombre de hams dans la base d’apprentissage : 200
```

```
Apprentissage sur: 200 spams et 200 hams...
=== TEST ===
```

```
P(Y=SPAM | X=x) = -291,529731, P(Y=HAM | X=x) = -319,106067.
Spam number 0 classified as SPAM
P(Y=SPAM | X=x) = -295,396522, P(Y=HAM | X=x) = -338,647222.
Spam number 1 classified as SPAM
P(Y=SPAM | X=x) = -296,500981, P(Y=HAM | X=x) = -350,242648.
Spam number 2 classified as SPAM
P(Y=SPAM | X=x) = -296,500981, P(Y=HAM | X=x) = -350,242648.
Spam number 3 classified as SPAM
P(Y=SPAM | X=x) = -205,249638, P(Y=HAM | X=x) = -253,033832.
...
```

$P(Y=\text{SPAM} \mid X=x) = -99,320038$, $P(Y=\text{HAM} \mid X=x) = -87,989156$.
 Spam number 29 classified as HAM *** error ***
 ...
 $P(Y=\text{SPAM} \mid X=x) = -103,779099$, $P(Y=\text{HAM} \mid X=x) = -100,698392$.
 Spam number 97 classified as HAM *** error ***
 $P(Y=\text{SPAM} \mid X=x) = -186,837117$, $P(Y=\text{HAM} \mid X=x) = -196,483656$.
 Spam number 98 classified as SPAM
 $P(Y=\text{SPAM} \mid X=x) = -220,585287$, $P(Y=\text{HAM} \mid X=x) = -248,344009$.
 Spam number 99 classified as SPAM
 ...
 $P(Y=\text{SPAM} \mid X=x) = -386,925411$, $P(Y=\text{HAM} \mid X=x) = -369,281157$.
 Ham number 0 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = -168,757847$, $P(Y=\text{HAM} \mid X=x) = -143,911920$.
 Ham number 1 classified as HAM
 ...
 $P(Y=\text{SPAM} \mid X=x) = -417,327559$, $P(Y=\text{HAM} \mid X=x) = -431,827431$.
 HAM number 54 classified as SPAM *** error ***
 $P(Y=\text{SPAM} \mid X=x) = -466,173949$, $P(Y=\text{HAM} \mid X=x) = -474,356725$.
 HAM number 55 classified as SPAM *** error ***
 ...
 $P(Y=\text{SPAM} \mid X=x) = -174,426532$, $P(Y=\text{HAM} \mid X=x) = -155,199404$.
 Ham number 196 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = -255,184287$, $P(Y=\text{HAM} \mid X=x) = -236,380796$.
 Ham number 197 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = -130,871171$, $P(Y=\text{HAM} \mid X=x) = -118,576840$.
 Ham number 198 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = -368,254599$, $P(Y=\text{HAM} \mid X=x) = -340,982368$.
 Ham number 199 classified as HAM

 Test error on 100 spams : 12.0%
 Test error on 200 hams : 2.0%
 Global **test** error : 5.333333333333333%