

Projet : filtre anti-spam

Ladeveze Quentin
Lallemand François

7 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 : 500
```

```
Nombre de hams dans la base d’apprentissage : 500
```

```
Apprentissage sur: 500 spams et 500 hams...
=== TEST ===
```

```
P(Y=SPAM | X=x) = 1,000000, P(Y=HAM | X=x) = 0,000000.
Spam number 0 classified as SPAM
P(Y=SPAM | X=x) = 1,000000, P(Y=HAM | X=x) = 0,000000.
Spam number 1 classified as SPAM
P(Y=SPAM | X=x) = 1,000000, P(Y=HAM | X=x) = 0,000000.
Spam number 2 classified as SPAM
P(Y=SPAM | X=x) = 1,000000, P(Y=HAM | X=x) = 0,000000.
Spam number 3 classified as SPAM
P(Y=SPAM | X=x) = 1,000000, P(Y=HAM | X=x) = 0,000000.
Spam number 4 classified as SPAM
```

$P(Y=\text{SPAM} \mid X=x) = 1,000000$, $P(Y=\text{HAM} \mid X=x) = 0,000000$.
 Spam number 5 classified as SPAM
 $P(Y=\text{SPAM} \mid X=x) = 0,989475$, $P(Y=\text{HAM} \mid X=x) = 0,010525$.
 Spam number 6 classified as SPAM
 ...
 Ham number 47 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 48 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 49 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 50 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 51 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000003$, $P(Y=\text{HAM} \mid X=x) = 0,999997$.
 Ham number 52 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 53 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 1,000000$, $P(Y=\text{HAM} \mid X=x) = 0,000000$.
 HAM number 54 classified as SPAM *** error ***
 ...
 Ham number 94 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 95 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 96 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 97 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 98 classified as HAM
 $P(Y=\text{SPAM} \mid X=x) = 0,000000$, $P(Y=\text{HAM} \mid X=x) = 1,000000$.
 Ham number 99 classified as HAM
 Test error on 100 spams : 9.0%
 Test error on 100 hams : 4.0%
 Global **test** error : 6.5%