

# analyse du sentiment

## Analyse du sentiment.

On utilisera la librairie Pattern pour le français, sur Python 2.7. :

<http://www.clips.ua.ac.be/pages/pattern-fr>

## Installation

Elle peut être installée (au labo) avec **easy\_install pattern** ou **pip install pattern**, depuis la ligne de commande (**cmd**)

## Exécution

Pour exécuter les fichiers, depuis le shell Python (IDLE ou Pythonwin):

```
>>> execfile("c:\\chemin\\vers\\le\\fichier\\sentiment_reviews.py")
```

Sortie:

```
- Lire textes du fichier '/home/pablo/projects/scr/otc/sentiment/books-fr.test-aligne
d.csv'
  Lu 1000 textes
- Analyser 1000 textes
  - Done 100 lines
  - Done 200 lines
  - Done 300 lines
  - Done 400 lines
  - Done 500 lines
  - Done 600 lines
  - Done 700 lines
  - Done 800 lines
  - Done 900 lines
- Écriture vers /home/pablo/projects/scr/otc/sentiment/books-fr.test-aligned.csv.out.
txt
```

## Exécution "interactive"

Une fois que le script a fini et créé le fichier de sortie pour les reviews (qui finit par "out.txt"), on a aussi accès aux fonctions contenues dans le script depuis le shell Python.

Pour analyser des tweets en les entrant sur le shell, on appelle la fonction `analyser_textes`, qui demande comme premier argument une liste de textes. Chaque texte dans la liste contient un champ "texte" et un champ "polarité"

```
{ "texte": u"Ceci est un livre très mauvais, c'est nul 😞", "polarite": "neg" }
```

Le champ "polarité" est la polarité attendue, pour comparer avec la polarité calculée par l'outil (mais évidemment l'outil fait ses propres calculs sans connaître cette polarité attendue, qui est là juste pour l'évaluation automatique disponible depuis le script (voir la variable `WRITE_METRICS`).

```
>>> analyser_textes( [ { "texte": u"Ceci est un livre très mauvais, c'est nul 😞", "polarite": "neg" } ] )
```

Ça donne comme sortie

```
>>> - Analyser 1 textes  
Pol: -0.553333333333 | Sub: 0.666666666667 | Texte: Ceci est un livre très mauvais,  
c'est nul 😞 | Indices: mauvais, -0.71, 0.8, None~~😞, -0.75, 1.0, mood~~
```

## Sentiment en anglais

Pour faire l'analyse en anglais, on change lors de l'appel de la fonction la valeur de l'argument "lan", qui est optionnel, et "fr" par défaut:

```
>>> analyser_textes( [ { "texte": u"This is a horrible book.", "polarite": "neg" } ]  
, lan="en")  
- Analyser 1 textes  
Pol: -1.0 | Sub: 1.0 | Texte: This is a horrible book. | Indices: horrible, -1.0, 1.  
0, None
```



exercices\_sentiment.zip



sentiment\_pres.pdf

Modifier