

Fusion anticipée de descripteurs bas niveau pour la détection d'émotions dans les textes

F. Dzogang, M.-J. Lesot, M. Rifqi
WACAI 2012

16/11/2012



Classification de textes bas niveau

| | | |
|---|---|----------------|
| d_1 : "I love you my darling wife." | → | <i>love</i> |
| d_2 : "I am bitterly broken hearted." | → | <i>anger</i> |
| d_3 : "I drink soda." | → | <i>neutral</i> |
| ... | | |

- multi-classes : [modélisation catégorielle des émotions](#)

Caractéristiques

- peu de données étiquetées
- déséquilibre des classes
- complexité des expressions linguistiques émotionnelles
 - ambiguïté, imprécisions, contexte d'apparition

Représentation des textes

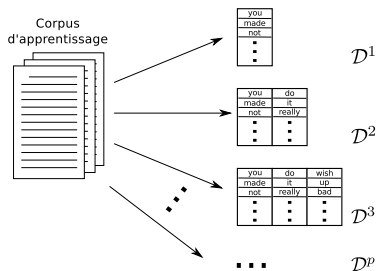
- descripteurs : *n*-grammes de différents ordres
 - *n*-gramme : suite consécutive de *n* termes
- spécialisation des dictionnaires selon les émotions
 - dictionnaire : ensemble de descripteurs
- combinaison par fusion anticipée

Classifieurs

- stratégie “un contre tous”
- 2 étapes de décision
- discrimination linéaire : interprétabilité des décisions

n -grammes de différents ordres

- descripteurs : n -grammes uniques dont occurrences ≥ 2 textes

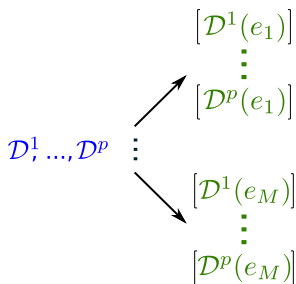


- couverture des descripteurs : n faibles
 - ex : *sad* $\{n = 1\}$ et *sad to hear of* $\{n = 4\}$
- contexte d'apparition des termes : n élevés
 - ex : *bad* $\{n = 1\}$, **not bad** $\{n = 2\}$, **not really bad** $\{n = 3\}$

$\Rightarrow p$ dictionnaires : $\mathcal{D}^1, \dots, \mathcal{D}^p$

Spécialisation des dictionnaires selon les émotions (1/2)

- M classes émotionnelles : e_1, \dots, e_M



- objectif : éliminer les descripteurs les moins pertinents
- hypothèse : vocabulaire spécifique à chaque émotion

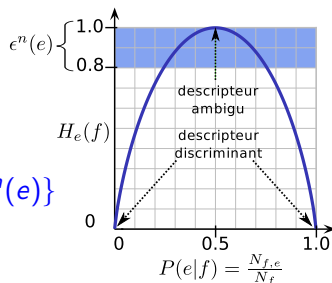
$\Rightarrow p \times M$ dictionnaires : $\mathcal{D}^n(e_i)$, $n \in [1..p], i \in [1..M]$

objectif : éliminer les descripteurs les moins pertinents

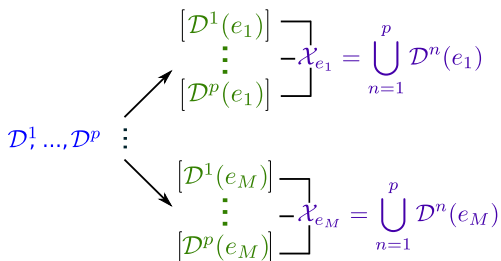
- descripteur $f \in \mathcal{D}^n$ et émotion e : $P(e|f) = \frac{N_{f,e}}{N_f}$
 - vaut 1 lorsque f est spécifique à e
 - vaut 0 lorsque f est spécifique à $\neg e$
 - vaut $1/2$ lorsque f n'est pas spécifique
- pertinence : entropie de Shannon
 - $H_e(f)$ proche de 1 si f ambigu, de 0 si f discriminant

filtrage : $\mathcal{D}^n(e) = \{f \in \mathcal{D}^n / H_e(f) < \epsilon^n(e)\}$

\Rightarrow seuil spécifique à chaque émotion et à chaque ordre



Fusion anticipée des descripteurs spécifiques aux émotions



- espace de représentation final :

$$\mathcal{X}_e = \bigcup_{n=1}^p \mathcal{D}^n(e)$$

- vecteur de représentation d'un texte pour l'émotion e :

$$\mathbf{x}_e = \mathbf{x}_e^1 \oplus \dots \oplus \mathbf{x}_e^p \in \mathcal{X}_e$$

Stratégie un contre tous : M problèmes binaires

$$f_e : \mathcal{X}_e \rightarrow \{e, \neg e\}$$

$$\mathbf{x} \mapsto f_e(\mathbf{x})$$

- corpus d'apprentissage pour f_e
 - exemples associés à l'étiquette e
 - contre-exemples associés à une étiquette $\neq e$
- classification en 2 étapes
 1. f_{neutre} discrimine neutre/émotionnel
 2. f_e émotionnels : exemples *neutre* exclus

Classification linéaire : identification des descripteurs pertinents

- $f_e(\mathbf{x}) = \text{sign}(\langle \mathbf{x}, \boldsymbol{\alpha} \rangle + b)$
 - $\Rightarrow \alpha_j$: poids associés à chacun des $|\mathcal{X}_e|$ descripteurs
- SVM avec noyau linéaire et coût de classification asymétrique

- corpus étiqueté : I2B2 track2 (2011)
 - 4 241 phrases étiquetées
- caractéristiques
 - textes très courts puisque réduits à une phrase
 - parmi les $M = 15$ étiquettes, 2 ne sont pas des émotions
 - étiquettes *instruction* et *information*

| Etiquette | # phrases | Etiquette | # phrases |
|--------------|-----------|------------------------|-----------|
| neutral | 2 460 | anger | 69 |
| instruction | 800 | sorrow | 51 |
| hopelessness | 455 | hopefulness | 47 |
| love | 296 | happiness/peacefulness | 25 |
| information | 295 | fear | 25 |
| guilt | 208 | pride | 15 |
| blame | 107 | abuse | 9 |
| thankfulness | 94 | forgiveness | 6 |

- descripteurs considérés :
 - mots, ponctuation, émoticônes
- présence/absence des descripteurs
 - \mathbf{x}_e : vecteur binaire
- $p = 2$: unigrammes et bigrammes pour toutes les émotions
 - pas de gain pour $p \geq 3$ (résultats expérimentaux)

Performance des classifieurs : évaluation par validation croisée

- évaluation des performances pour étiquette e

$$\text{tp} : f_e(\mathbf{x}) = e \quad \wedge \quad y = e$$

$$\text{fn} : f_e(\mathbf{x}) = \neg e \quad \wedge \quad y = e$$

$$\text{fp} : f_e(\mathbf{x}) = e \quad \wedge \quad y = \neg e$$

- couverture de la classe e
 $\text{rappel} = tp / (tp + fn)$
- prédictions correctes pour e
 $\text{précision} = tp / (tp + fp)$

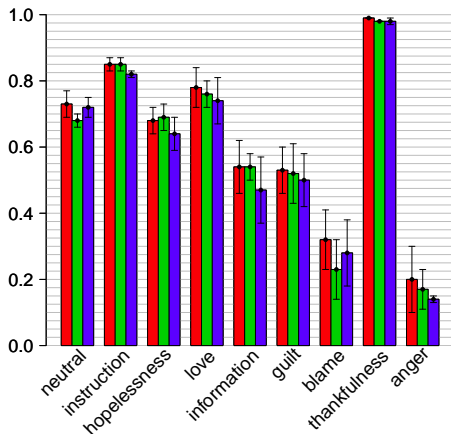
⇒ score F1 : moyenne harmonique entre *rappel* et *précision*

Dictionnaires

- analyse des tailles
- analyse du contenu : descripteurs les plus discriminants
 - au sens de la fonction de prédiction

Résultats (1/3) : scores F1

fusion (rouge), unigrammes (vert), et bigrammes (bleu)



- performances variables selon les étiquettes
- influence du nombre d'exemples d'apprentissage
- fusion anticipée : en moyenne ↗ score F1

- dictionnaires non spécialisés
 - $|\mathcal{D}^1| = 1\,206$, $|\mathcal{D}^2| = 2\,968$
 - soit au total : 4 174 descripteurs
- spécialisation : variable selon les étiquettes
 - filtrage plus faible pour unigrammes que pour bigrammes
 - ex : $|\mathcal{D}^1(\text{neutral})| = 0.48 \times |\mathcal{D}^1|$
 - ex : $|\mathcal{D}^2(\text{neutral})| = 0.45 \times |\mathcal{D}^2|$

⇒ descripteurs plus spécifiques pour $n > 1$
 - filtrage plus faible pour étiquettes rares que fréquentes
 - ex : $|\mathcal{D}^2(\text{neutral})| = 0.45 \times |\mathcal{D}^2|$
 - ex : $|\mathcal{D}^2(\text{abuse})| = 0.85 \times |\mathcal{D}^2|$

⇒ descripteurs discriminants par leur absence

Résultats (3/3) : descripteurs discriminants

7 unigrammes et 7 bigrammes les plus discriminants pour 5 plus fréquentes émotions

| Emotions | Descripteurs |
|--------------|--|
| Thankfulness | thank appreciate than nice effort kindness under be swell than you you dear appreciate it too . have be for your |
| Love | love wonderful bless watch beloved most loving you . do . be wonderful love you god bless your john me on |
| Hopelessness | cancer am suffer die struggle everybody tired without you go on dear jane can not . my be . of all |
| Guilt | sorry forgive excuse fail hurt could burden have be forgive me please forgive have do understand . not to to help |
| Blame | sorry thank love please give wish go to be cause you of it you . you to in the to go |

Résultats (3/3) : descripteurs discriminants

7 unigrammes et 7 bigrammes les plus discriminants pour 5 plus fréquentes émotions

| Emotions | Descripteurs |
|--------------|--|
| Thankfulness | thank appreciate than nice effort kindness under be swell than you you dear appreciate it too . have be for your |
| Love | love wonderful bless watch beloved most loving you . do . be wonderful love you god bless your john me on |
| Hopelessness | cancer am suffer die struggle everybody tired without you go on dear jane can not . my be . of all |
| Guilt | sorry forgive excuse fail hurt could burden have be forgive me please forgive have do understand . not to to help |
| Blame | sorry thank love please give wish go to be cause you of it you . you to in the to go |

- descripteurs non nécessairement corrélés à \neq ordres

Résultats (3/3) : descripteurs discriminants

7 unigrammes et 7 bigrammes les plus discriminants pour 5 plus fréquentes émotions

| Emotions | Descripteurs |
|--------------|---|
| Thankfulness | thank appreciate than nice effort kindness under be swell than you you dear appreciate it too . have be for your |
| Love | love wonderful bless watch beloved most loving you . do . be wonderful love you god bless your john me on |
| Hopelessness | cancer am suffer die struggle everybody tired without you go on dear jane can not . my be . of all |
| Guilt | sorry forgive excuse fail hurt could burden have be forgive me please forgive have do understand . not to to help |
| Blame | sorry thank love please give wish go to be cause you of it you . you to in the to go |

- descripteurs non nécessairement corrélés à \neq ordres
- descripteurs plus spécifiques pour $n = 2$

Résultats (3/3) : descripteurs discriminants

7 unigrammes et 7 bigrammes les plus discriminants pour 5 plus fréquentes émotions

| Emotions | Descripteurs |
|--------------|--|
| Thankfulness | thank appreciate than nice effort kindness under be swell than you you dear appreciate it too . have be for your |
| Love | love wonderful bless watch beloved most loving you . do . be wonderful love you god bless your john me on |
| Hopelessness | cancer am suffer die struggle everybody tired without you go on dear jane can not . my be . of all |
| Guilt | sorry forgive excuse fail hurt could burden have be forgive me please forgive have do understand . not to to help |
| Blame | sorry thank love please give wish go to be cause you of it you . you to in the to go |

- descripteurs non nécessairement corrélés à \neq ordres
- descripteurs plus spécifiques pour $n = 2$
- "." termine une phrase : position des termes semble importer

Conclusions : représentation des textes pour les émotions

- n -grammes d'ordres différents
 - descripteurs **génériques** et descripteurs **spécifiques**
- spécialisation des dictionnaires
 - **élimination** des descripteurs les **moins informatifs**
- **fusion** des dictionnaires
 - **améliore en moyenne les scores F1**
 - unigrammes ↗ *rappel* et bigrammes ↗ *précision*

Perspectives

- ordres plus élevés pertinents selon les émotions
 - $n = 3 \rightarrow$ *sorrow* $F1 = 0.98$ et *hopelessness* $F1 = 0.8$
 - bas niveau limité par la base d'apprentissage
 - ordres plus élevés et émotions rares
- ⇒ combinaison de descripteurs bas niveau et haut niveau

<http://ieee-ssci.org>

deadline : 23 nov. 2012



2013 IEEE Symposium Series on Computational Intelligence

IEEE SSCI 2013

15 Mon - 19 Fri April 2013, Singapore
Grand Copthorne Waterfront Hotel

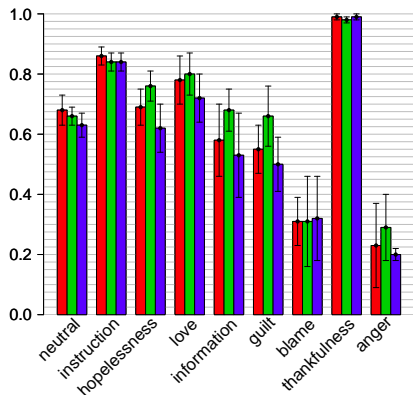
IEEE Computational Intelligence Society

Topics of interest

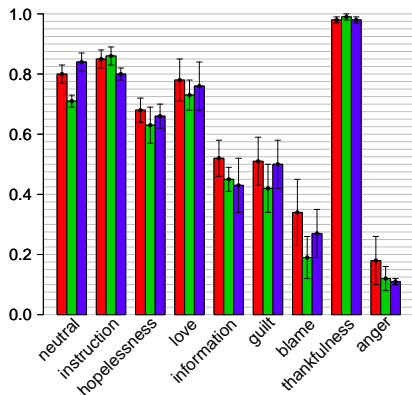
- theories of emotions from psychology and their application to computer sciences
- computational models for processing emotions and other affective states
- multimodal emotional corpora
- automatic emotion recognition from physiological signals, facial expressions, body language, speech
- emotion mining in texts, images, videos, film, multimedia data
- affective interaction with virtual agents and robots

Résultats : rappel et précision

fusion (rouge), unigrammes (vert), et bigrammes (bleu)



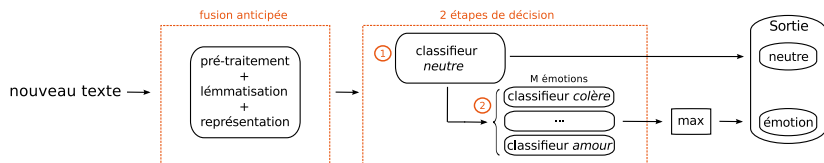
apport des unigrammes
en moyenne ↗ Rappel
(descripteurs couvrants)



apport des bigrammes
en moyenne ↗ Précision
(descripteurs précis)

Vue globale de la chaîne de traitement

Etiquetage d'un nouveau texte



Etape 1 :

- entrée : textes pré-traités
- discrimination entre *neutre* et *émotionnel*

Etape 2 :

- entrée : textes émotionnels
- discrimination entre chaque émotion et toutes les autres
⇒ **émotion détectée : celle de confiance maximale**

Compétition I2B2 track 2 : résultats des participants

| Système | Haut niveau | score F1 |
|-----------------|-------------|----------|
| Open university | oui | 0.61 |
| Msra | oui | 0.59 |
| Mayo | oui | 0.56 |
| Nrciit | oui | 0.55 |
| Oslo | oui | 0.54 |
| Limsi | oui | 0.54 |
| Swatmrc | oui | 0.53 |
| Uman | oui | 0.53 |
| Cardiff | oui | 0.53 |
| Lt3 | oui | 0.53 |
| Utd | oui | 0.52 |
| Wolverine | oui | 0.50 |
| Clips | oui | 0.50 |
| Sri et UcDavis | oui | 0.48 |
| Diego-Acu | oui | 0.48 |
| Notre approche | non | 0.47 |
| Ebi | non | 0.46 |
| Duluth | non | 0.45 |
| Tpavacoe | non | 0.38 |
| Lassa | non | 0.38 |

- les approches bas niveau sont limitées à $F1 \leq 0.47$
 - systèmes exploitant uniquement le corpus d'apprentissage
- les approches haut niveau mettent en œuvre
 - ressources sémantiques
 - ressources affectives
 - données externes