



---

# Twitter et les sondages automatiques

## RAPPORT DU PROJET DAC

---

Kim-Anh Laura NGUYEN  
Arij RIABI  
Promo DAC 2018-2019

*Encadrants :* Vincent GUIGUE  
Nicolas BASKIOTIS

# Table des matières

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>   | <b>2</b>  |
| <b>2</b> | <b>Analyse de sentiments et transfert</b>   | <b>3</b>  |
| 2.1      | Apprentissage sur des critiques de films . . . . .  | 3         |
| 2.2      | Apprentissage sur des tweets traduits . . . . .   | 4         |
| 2.3      | Application des modèles sur d'autres données étiquetées . . . . .                         | 4         |
| 2.4      | Application des modèles sur les tweets des élections présidentielles . . . . .            | 5         |
| <b>3</b> | <b>Analyse de communautés</b>   | <b>7</b>  |
| <b>4</b> | <b>Fusion des deux approches</b>  | <b>13</b> |
| 4.1      | Classification de sentiments . . . . .  | 13        |
| 4.2      | Prédiction d'opinions . . . . .   | 14        |
| 4.3      | Bilan . . . . .   | 16        |
| <b>5</b> | <b>Conclusion et perspectives</b>   | <b>17</b> |
| <b>A</b> | <b>Sélection de tweets de test pour l'évaluation du classifieur de sentiments</b>         | <b>18</b> |
| <b>B</b> | <b>Sélection de tweets de test pour l'évaluation du classifieur associé à J-L. Mélen-</b> |           |
|          | <b>chon</b>   | <b>18</b> |
| <b>C</b> | <b>Sélection de tweets de test pour l'évaluation du classifieur associé à E. Macron</b>   | <b>19</b> |

# 1 Introduction

Avec l'avènement des réseaux sociaux et la rapidité du relais de l'information, l'analyse de sentiments, ou *Sentiment Analysis*, est devenu l'un des nouveaux défis en traitement automatique des langues, ou *Natural Language Processing*. L'objectif est d'analyser automatiquement les opinions exprimées dans les textes : pour un texte donné, nous cherchons à identifier sa polarité comme étant soit positive, soit négative.

L'analyse de sentiments a connu un récent essor dans les secteurs de la recherche scientifique et de l'industrie. Les grandes masses de données issues des réseaux sociaux permettent d'extraire l'opinion générale des utilisateurs pour un sujet donné. Parmi les nombreux cas applicatifs, nous pouvons citer la recommandation, l'explication de sondages, la détection de spams, la consultation d'avis et l'étude de marché.

Les travaux de Pang et al. (2002) marquent un tournant dans le traitement automatique des langues : en entraînant des classifieurs standards sur des critiques de films, ils montrent que les techniques de *machine learning*, ou apprentissage automatique, produisent de meilleurs résultats en analyse de sentiments que les modèles construits par des linguistes. Pang & Lee (2008) dressent quelques années plus tard un panorama de méthodes d'apprentissage automatique permettant de produire des systèmes de recherche d'information basés sur les opinions. Pour pallier à l'absence de corpus annotés, Blitzer et al. (2007) se confrontent à la problématique d'apprentissage par transfert, ou *transfer learning*, et offrent une méthode pour l'adaptation au domaine qui permet de détecter des correspondances de comportement entre des traits spécifiques au domaine source et des traits spécifiques au domaine cible. Pan & Yang (2009) montrent que, dans les bonnes conditions, le *transfer learning* peut faciliter considérablement l'apprentissage en contournant la phase d'annotation des données. En comparant des données sur des produits de consommation issues de différents réseaux sociaux, Mejova (2012) montre qu'il peut être avantageux d'entraîner des modèles de classification de sentiments sur des textes sources différents des textes cibles, ce qui n'est en revanche pas le cas lorsqu'il s'agit d'opinions politiques, démontrant au passage la difficulté de la classification de sentiments politiques. Enfin, V. D. Blondel et al. (2008) proposent une méthode heuristique basée sur l'optimisation modulaire pour extraire les communautés de graphes volumineux, qui est donc applicable aux données issues de réseaux sociaux.

Nous disposons d'une base de données non supervisée constituée de plus de cinq millions de tweets collectés durant le premier tour des élections présidentielles françaises de 2017 et s'étalant sur une période de deux jours. En plus du contenu textuel, nous avons également accès à des métadonnées : nous savons, par exemple, quels utilisateurs sont concernés par un tweet et quels candidats sont mentionnés dans un tweet. De plus, d'autres jeux de données étiquetées en fonction des sentiments sont à notre disposition : des critiques de films, des tweets traduits de l'anglais au français ainsi que des avis sur des produits Décathlon.

À partir de la base de tweets collectés durant les élections présidentielles, nous souhaitons déterminer les sentiments ainsi que les opinions politiques des utilisateurs.

Pour répondre à notre problématique, nous avons suivi la démarche suivante. Nous considérons d'abord une approche classique en traitement automatique des langues avec transfert d'un classifieur de sentiments appris sur des sources étiquetées vers les données textuelles Twitter. Nous étudierons ensuite ces questions en exploitant, non pas le contenu textuel, mais les interactions entre utilisateurs en construisant un graphe de communautés. Enfin, nous fusionnerons les deux approches dans l'objectif d'apprendre un modèle de classification de sentiments sur des données faiblement supervisées.

## 2 Analyse de sentiments et transfert

Nous cherchons dans un premier temps à déterminer la polarité des tweets. La difficulté à laquelle nous nous heurtons est l'absence de supervision dans cette base de données. Cependant, le web participatif nous donnant accès à un nombre important de données étiquetées, nous pouvons exploiter ces bases pour apprendre un classifieur que nous appliquerons sur les tweets des élections présidentielles. Il s'agit d'une problématique de *transfer learning*, ou apprentissage par transfert, qui consiste à entraîner un modèle sur une tâche source, puis appliquer les connaissances apprises vers une tâche cible. Tous les problèmes étudiés dans cette section sont des tâches de classification de sentiments. Les domaines, en revanche, diffèrent d'un problème à un autre.

Nous nous servons de trois jeux de données étiquetées en fonction des sentiments : une base de critiques de films, une base de tweets traduits de l'anglais au français et une base d'avis sur des produits Décathlon. Les textes sont rédigés en français. Leurs caractéristiques sont résumées dans le tableau 1. Il est important de noter que chacun de ces jeux de données présente des anomalies : la base de critiques de films contient des avis rédigés en anglais, les tweets sont traduits approximativement de l'anglais au français, et certains avis Décathlon ne sont composés que d'un seul mot.

| Données            | Taille | Taux de sentiments positifs |
|--------------------|--------|-----------------------------|
| Critiques de films | 50000+ | 16%                         |
| Tweets traduits    | 1.5M+  | 50%                         |
| Avis Décathlon     | 15000+ | 87%                         |

TABLEAU 1 – Jeux de données étiquetées

Nous optimiserons un modèle sur les critiques de films et un autre sur les tweets traduits. Nous évaluerons quantitativement leurs performances sur les jeux de données sus-mentionnées, puis nous analyserons qualitativement les résultats obtenus sur la base de tweets des élections présidentielles.

Pour chaque classifieur, nous mettrons en place une chaîne de traitements standard en traitement automatique des langues :

- pré-traitements du texte
- construction d'un vocabulaire
- mise sous forme vectorielle (*Bag Of Words*) des tweets
- optimisation d'un modèle par validation croisée et évaluation par le score F1 et le taux de reconnaissance, aussi appelé taux de bonne classification.

Le SVM linéaire étant un modèle efficace pour la classification de sentiments (Rafrafi et al., 2013), nous utiliserons dans toute la suite le modèle **LinearSVC** de la librairie Scikit-learn (Pedregosa et al., 2011).

### 2.1 Apprentissage sur des critiques de films

Nous disposons d'une base de critiques de films, décrite dans le tableau 1, chaque critique étant constituée d'un commentaire ainsi que d'une note « en étoiles », c'est-à-dire comprise entre 0 et 5. Nous binarisons ces notes afin de produire une tâche de classification de sentiments. Nous essayons différentes combinaisons de pré-traitements du texte (stemmatisation, suppression des stopwords, mise en minuscules) afin de construire un dictionnaire pertinent, puis entraînons un SVM sur les revues traitées. Dans le tableau 2 figurent les mots du dictionnaire les plus polarisés, récupérés après entraînement du modèle final. Nous remarquons que ces résultats sont peu interprétables, notamment à cause de la présence d'entités nommées.

| Avis    | Termes   |
|---------|--|
| Négatif | dowhatyoucan, empapaout, mâtur, évit, natur, tribun, match, méchancon, ignorent, doutent         |
| Positif | parfait, coraliecarm, sénior, longtemp, cpasbourdinleboss, medefdt, equivalent, fumister, 4m, be |

TABLEAU 2 – Termes les plus polarisés

Cependant, il y a une différence non-négligeable entre ces critiques de films et les tweets des élections présidentielles : la manière d’écrire un tweet – ou un contenu provenant d’un réseau social de manière générale – n’est pas la même que celle de rédiger une critique. Il y a donc un risque d’avoir trop peu de termes en commun entre les données sources et les données cibles. De plus, Bessalov et al. (2011) montrent que les performances en reconnaissance de sentiments et la taille des ensembles d’apprentissage sont liées ; la quantité de données est donc un facteur important.

## 2.2 Apprentissage sur des tweets traduits

Nous souhaitons donc entraîner un modèle sur un domaine où les données sont plus nombreuses et proches, en termes de format et de contenu, des tweets des élections présidentielles. Nous nous servons donc de la base de tweets traduits, décrite dans le tableau 1. Contrairement aux critiques de films, les classes sont ici équilibrées. Comme précédemment, nous cherchons la combinaison de pré-traitements la plus pertinente et optimisons notre modèle afin d’obtenir les meilleurs scores sur ce même domaine. Les termes les plus polarisés figurent dans le tableau 3. Nous remarquons qu’ils sont moins spécifiques au domaine que les termes du tableau 2, et la plupart expriment explicitement un sentiment positif ou négatif.

| Sentiment | Termes  |
|-----------|---|
| Négatif   | triste, neige, déteste, manque, malade, froid, perdu, Pourquoi, désolé, mal |
| Positif   | adore, merci, Salut, heureux, haha, génial, cool, bonne, Bonjour, Merci     |

TABLEAU 3 – Termes les plus polarisés

## 2.3 Application des modèles sur d’autres données étiquetées

Nous conservons les classifieurs issus des deux domaines précédents puisqu’ils possèdent chacune leur propre avantage, l’un étant que le texte est originalement en français, l’autre que le contenu est plus proche de nos données.

En utilisant le même vocabulaire et les mêmes pré-traitements qu’en apprentissage, nous évaluons ces deux classifieurs sur des domaines de classification de sentiments différents, ainsi que sur leur domaine source respectif. De plus, nous entraînons un classifieur aléatoire sur chaque domaine source et l’évaluons sur les différents domaines cibles afin d’obtenir des scores de référence. Les tableaux 4 et 5 recensent ces résultats.

|                                 | Taux de reconnaissance |           | Score F1     |           |
|---------------------------------|------------------------|-----------|--------------|-----------|
|                                 | SVM linéaire           | Aléatoire | SVM linéaire | Aléatoire |
| Critiques de films              | 91%                    | 73%       | 67%          | 17%       |
| Tweets traduits                 | 50%                    | 50%       | 66%          | 62%       |
| Avis sur des produits Décathlon | 13%                    | 25%       | 22%          | 22%       |

TABLERAU 4 – Résultats du transfert des connaissances issues de critiques de films sur d’autres bases de données étiquetées

|                                 | Taux de reconnaissance |           | Score F1     |           |
|---------------------------------|------------------------|-----------|--------------|-----------|
|                                 | SVM linéaire           | Aléatoire | SVM linéaire | Aléatoire |
| Critiques de films              | 37%                    | 50%       | 27%          | 24%       |
| Tweets traduits                 | 72%                    | 50%       | 72%          | 49%       |
| Avis sur des produits Décathlon | 74%                    | 50%       | 21%          | 13%       |

TABLERAU 5 – Résultats du transfert des connaissances issues de tweets traduits sur d’autres bases de données étiquetées

**Performances du classifieur entraîné sur les critiques de films** Le taux de bonne classification obtenu sur les avis Décathlon est très largement en-dessous de 50%, le score moyen d’un classifieur aléatoire sur un problème où les deux classes sont équilibrées. Or, les domaines source et cible sont asymétriques : la base Décathlon contient 87% d’avis positifs alors que la base de critiques de films n’en comprend que 16%. La mesure F1, qui correspond au taux de reconnaissance des données de la classe minoritaire, est donc plus adaptée à cette tâche. Elle fournit un score légèrement meilleur que le taux de bonne classification. Sur les tweets traduits, notre modèle offre le même taux de reconnaissance qu’un classifieur aléatoire mais un meilleur score F1 par rapport à l’aléatoire et par rapport au score obtenu sur la base Décathlon, ce qui montre qu’il généralise mieux aux tweets traduits qu’aux avis Décathlon.

**Performances du classifieur entraîné sur les tweets traduits** Ce modèle génère, sur les données Décathlon, un taux de reconnaissance ainsi qu’un score F1 meilleurs que ceux obtenus par le classifieur aléatoire. De plus, comme la base de tweets traduits contient autant d’exemples positifs que négatifs, le classifieur correspondant obtient un taux de reconnaissance bien meilleur sur les données Décathlon que le classifieur basé sur les critiques de films. Or, ce score est même meilleur que le taux de bonne classification obtenu sur le domaine source : encore une fois, nous ne pouvons pas évaluer ce classifieur avec cette métrique.

Comme les classes sont déséquilibrées dans les deux jeux de données cibles, nous préférons à nouveau évaluer les performances avec le score F1. Nous constatons que notre modèle produit de meilleurs résultats que le classifieur aléatoire sur chaque domaine, et qu’il classe mieux les critiques de films que les avis Décathlon.

**Remarque** La base Décathlon est spécifique au sport et est donc très différente des autres jeux de données. Il est ainsi plus difficile d’y transférer les connaissances apprises sur des critiques de films ou des tweets, ce qui explique aussi pourquoi le modèle entraîné sur les critiques de films est plus performant sur les tweets traduits que sur les données Décathlon, tout comme le modèle appris sur les tweets traduits est meilleur sur les critiques de films que sur les avis Décathlon.

## 2.4 Application des modèles sur les tweets des élections présidentielles

Comme les tweets des élections présidentielles ne sont pas annotés, il est impossible d’évaluer quantitativement les modèles sur ces données. Nous procédons à une rapide analyse qualitative basée sur la prédiction de chaque modèle sur quelques tweets. Nous utiliserons :

- un tweet négatif : @FrancoisFillon La seule question qu’on doit vous poser c’est : QUAND EST CE QUE VOUS NE MENTEZ PAS ??!!
- un tweet positif : La révolution citoyenne est train de se produire ! #Melenchon2017

- un tweet sarcastique : `Best-of du grand poète Macron`
- un tweet neutre : `#Rediff François Fillon refuse de se rendre chez Jean-Jacques Bourdin`

Les prédictions des deux classifieurs sur les différents tweets figurent dans le tableau 6.

| Nature du tweet | Classifieur basé critiques de films | Classifieur basé tweets traduits |
|-----------------|-------------------------------------|----------------------------------|
| Positif         | —                                   | +                                |
| Négatif         | —                                   | —                                |
| Sarcastique     | —                                   | +                                |
| Neutre          | —                                   | —                                |

TABLEAU 6 – Prédiction des classifieurs pour chacun des tweets représentant un sentiment particulier

Le classifieur entraîné sur les critiques de films prédit essentiellement un sentiment négatif. Nous interprétons ce phénomène par un problème de transfert couplé à un déséquilibre entre les deux classes au sein des données d'apprentissage, la base de critiques de films étant constituée d'un nombre bien plus important d'avis négatifs que d'avis positifs.

Quant au classifieur construit à partir des tweets traduits, il classifie la neutralité, qu'on ne peut pas catégoriser, comme étant négative. Le tweet positif ainsi que le tweet négatif sont correctement classifiés. En revanche, le tweet sarcastique, qui correspond à un sentiment négatif plus difficile à détecter, est considéré comme étant positif; en effet les termes « best » et « grand » sont généralement associé à des propos positifs.

Le domaine cible est très spécifique : les textes sont en français, relatifs aux élections présidentielles et la manière d'écrire sur les réseaux sociaux est particulière (en termes de syntaxe, comme les fautes d'orthographe, et de contenu, comme le sarcasme et l'ironie). Il y a donc très peu de jeux de données à disposition qui puissent fournir des résultats convenables par transfert. De plus, nous ne pouvons pas utiliser de métrique d'évaluation car nous ne disposons pas de supervision : il est donc impossible de conclure.

L'absence de données étiquetées étant un problème majeur, l'idée est maintenant d'obtenir une *faible supervision* en exploitant les autres informations contenues dans la base de tweets.

### 3 Analyse de communautés

Dans l'objectif d'extraire des communautés d'opinions, en particulier favorables à un seul candidat, nous considérons désormais une approche axée non pas sur le contenu des tweets, mais sur les liens entre utilisateurs. Nous souhaitons donc construire un graphe de communautés basé sur les interactions entre utilisateurs, i.e., un graphe où chaque nœud correspond à un utilisateur et chaque arc désigne un retweet, une citation ou une réponse. Nous utilisons la bibliothèque Python `networkx` (Hagberg et al., 2008), spécialisée dans l'étude des graphes et des réseaux.

L'algorithme de partitionnement que nous utilisons est la méthode de Louvain (V. D. Blondel et al., 2008), qui permet d'extraire les communautés de réseaux en optimisant la modularité, i.e., la densité d'arêtes à l'intérieur des communautés comparée à celle des arêtes reliant les communautés entre elles.

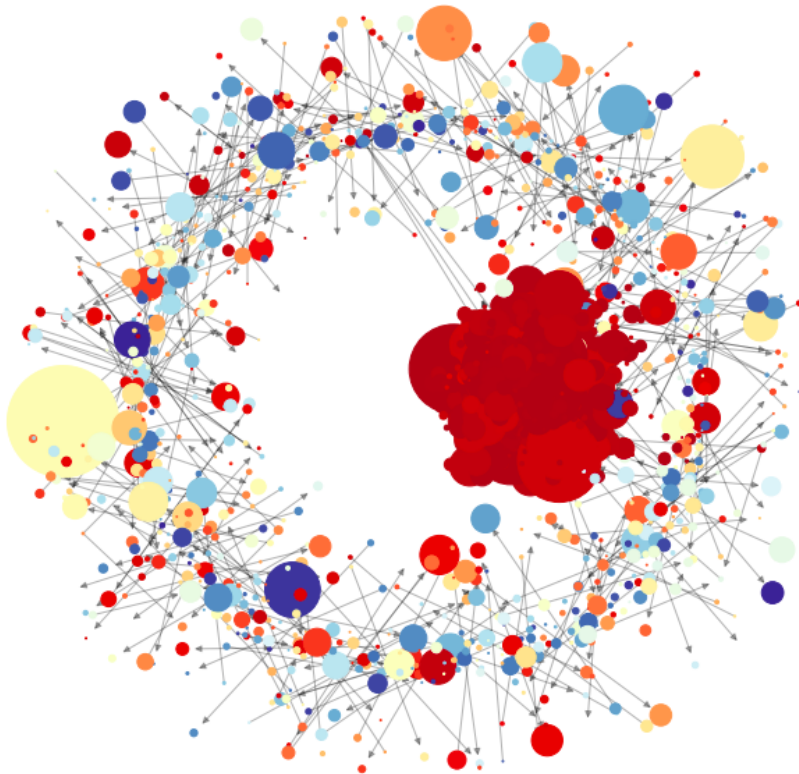


FIGURE 1 – Graphe de communautés construit sur 10000 tweets

La figure 1 contient le graphe de communautés construit à partir des utilisateurs impliqués dans un échantillon de 10000 tweets. Cette partition contient plus de 700 communautés constituées, en majorité, de très peu d'utilisateurs.

Nous décidons donc de garder uniquement les communautés de plus de 100 nœuds. La figure 2 contient les douze plus grandes communautés de la partition.



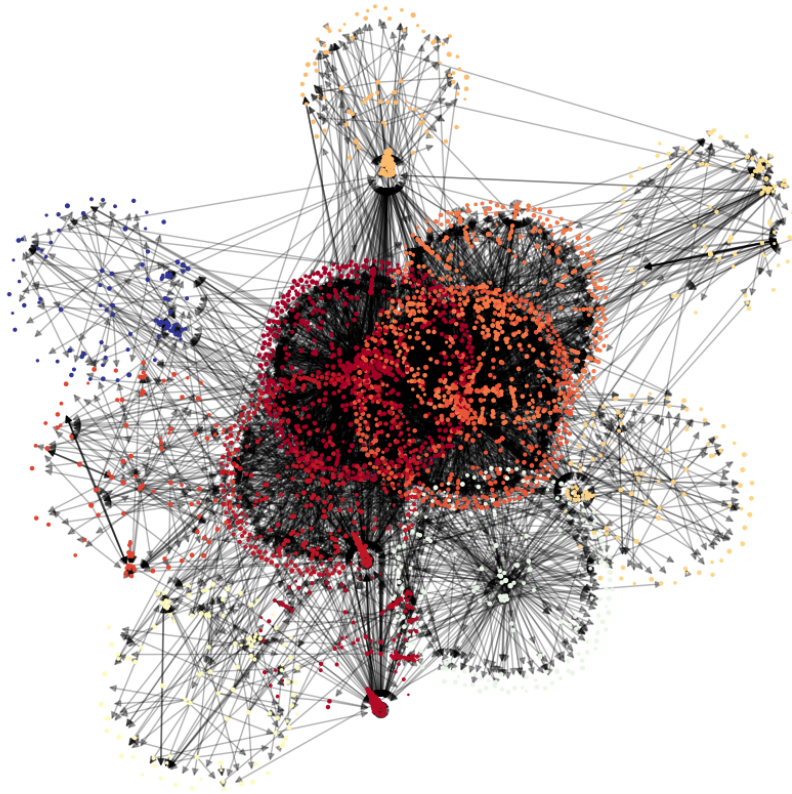


FIGURE 2 – Graphe contenant uniquement les 12 plus grandes communautés

De plus, nous enlevons les utilisateurs ne faisant que retweeter, et qui sont donc possiblement des bots. Nous retirons également ceux qui ne font pas suffisamment partie de leur communauté en choisissant comme critère de sélection le degré de centralité d'un nœud (Opsahl et al., 2010), qui est défini par le nombre de liens incidents à ce nœud. Cet indicateur capture la notion d'importance dans un graphe en identifiant les sommets les plus significatifs (plus un nœud est central, plus il est influent).

Enfin, nous affectons à chaque communauté le candidat le plus mentionné par les utilisateurs de la communauté en question. Nous faisons l'hypothèse que *ces utilisateurs parlent positivement de ce candidat et négativement des autres* : il s'agit donc de relier à chacune des communautés le candidat qu'elle soutient pour ainsi créer des communautés de partisans politiques. Cependant, comme les utilisateurs ont tendance à parler uniquement des candidats les plus populaires, seuls F. Fillon, B. Hamon, M. Le Pen, E. Macron et J-L. Mélenchon sont affectés à une communauté. Par la suite, nous analyserons uniquement les communautés de ces cinq candidats.

| Candidat  | Nombre de partisans |
|-----------|---------------------|
| Fillon    | 495                 |
| Hamon     | 196                 |
| Le Pen    | 562                 |
| Macron    | 587                 |
| Mélenchon | 722                 |

TABEAU 7 – Nombre de partisans politiques pour chaque candidat

Sur le tableau 7 figure le nombre d'utilisateurs favorables à chaque candidat. Ces résultats re-

flètent la réalité : selon la figure 3, tirée d’une étude de 01net <sup>1</sup>, les cinq candidats auxquels nous nous intéressons sont les plus suivis sur Twitter.

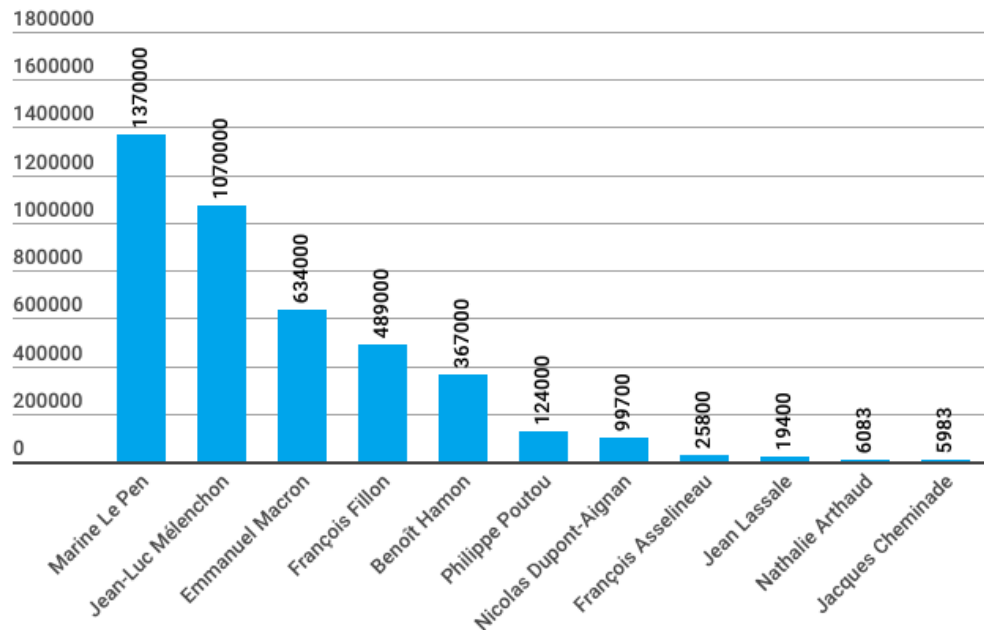


FIGURE 3 – Nombre d’abonnés par candidat

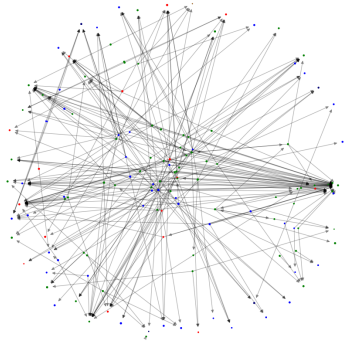
Le tableau 8 contient, pour chaque candidat, le nombre moyen de tweets postés par un utilisateur de sa communauté. Nous remarquons que la communauté de J-L. Mélenchon est la plus active.

| Candidat  | Nombre moyen de tweets par utilisateur |
|-----------|--|
| Fillon    | 182                                    |
| Hamon     | 205                                    |
| Macron    | 238                                    |
| Mélenchon | 302                                    |
| Le Pen    | 156                                    |

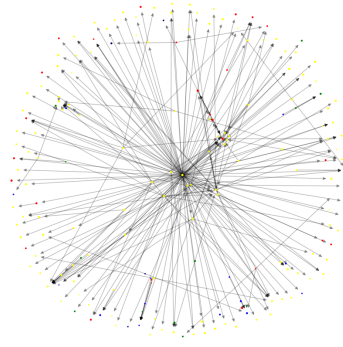
TABLEAU 8 – Nombre moyen de tweets par utilisateur pour chaque communauté

La figure 4 contient les communautés associées à chaque candidat. La couleur de chaque nœud correspond au candidat dont l’utilisateur parle le plus : vert pour F. Fillon, jaune pour B. Hamon, bleu pour E. Macron, rouge pour J-L. Mélenchon et bleu marine pour M. Le Pen. Nous remarquons que les graphes associés à E. Macron (figure 4d) et J-L. Mélenchon (figure 4e) sont très similaires : ils contiennent un nombre important de nœuds centraux, et les utilisateurs parlent essentiellement de leur candidat. Les communautés de F. Fillon (figure 4a) et M. Le Pen (figure 4c) possèdent une structure similaire mais contiennent moins de nœuds significatifs - en particulier F. Fillon. En revanche, le graphe associée à B. Hamon (figure 4b) se démarque du reste par sa structure en étoile : contrairement aux autres candidats, il y a très peu de nœuds qui sont très significatifs. De plus, les utilisateurs de cette communauté ne parlent quasiment que de leur candidat, contrairement à celle de Le Pen où les utilisateurs mentionnent beaucoup E. Macron, et celle de F. Fillon, dans laquelle E. Macron et J-L. Mélenchon sont beaucoup évoqués. Cette différence de structure de communauté peut donc refléter des stratégies différentes chez chaque candidat.

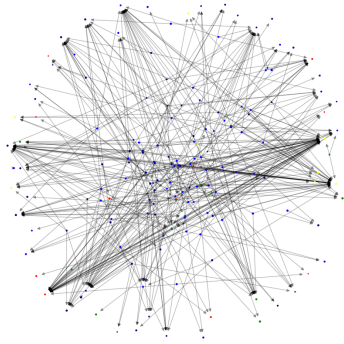
1. <https://www.01net.com/actualites/presidentielle-2017-qui-est-le-candidat-le-plus-populaire-sur-les-reseaux-sociaux-1141990.html>



(a) Fillon



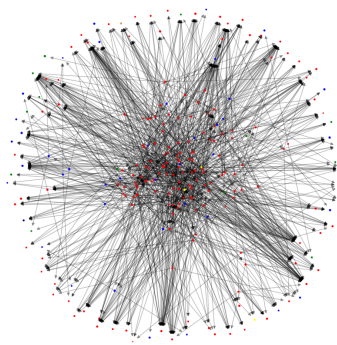
(b) Hamon



(c) Le Pen



(d) Macron



(e) Mélenchon

FIGURE 4 – Communauté de chaque candidat

Sur la figure 5 se trouve, pour chaque candidat, l'histogramme des degrés de centralité des nœuds de sa communauté. D'après la figure 5a, la majorité des sommets du graphe associé à F. Fillon

diffusent très peu (plus de 400 sur les 495 ont un degré de centralité compris entre 0 et 0.05) et très peu sont centraux. En revanche, les histogrammes correspondant aux autres candidats sont plus équilibrés : leurs communautés sont composées de davantage d'utilisateurs qui diffusent. Ces résultats sont cohérents avec ce que nous avons pu déduire de la figure 4.

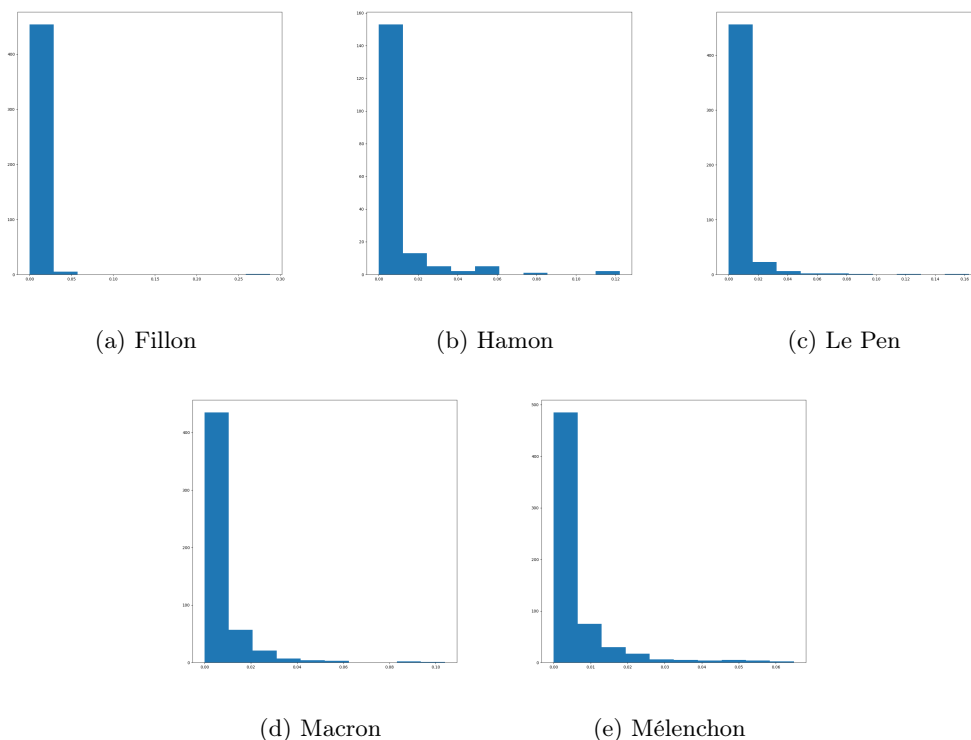


FIGURE 5 – Histogramme des degrés de centralité des utilisateurs de chaque communauté

Enfin, nous récupérons cinq tweets de chaque communauté afin de vérifier nos résultats à vue d'œil.

#### Tweets de la communauté Fillon

- Je sens que #Fillon va devenir désagréable.#RendsLargent
- Il montre enfin son vrai visage ce pourri #fillonlahonte
- Avec l'élection de M. Fillon, les journalistes, comme les Juges, seront remis au pas ! #JeVoteFillon
- Sens commun pourrait être présent dans un gouvernement Fillon
- @FrancoisFillon sera le prochain président car il est courageux, solide et déterminé face à la violence des attaques

#### Tweets de la communauté Hamon

- Merci aux ami-e-s de la #Creuse et aux habitant-e-s Aubusson pour l'accueil et pour @benoithamon #Hamon2017
- Raison de plus pour voter pour lui #Hamon2017
- Avec @benoithamon, priorité à l'éducation pour la réussite des élèves! #Hamin2017
- Soyons nombreux ce soir à la République avec @benoithamon
- Ceux qui attaquent #HAMON feraient mieux d'attaquer Macron, l'espion de la droite chargé de semer la zizanie à gauche. Ne soyez pas collabos

#### Tweets de la communauté Le Pen

- L'islam engendre l'islamisme. L'islamisme engendré le terrorisme. Le terrorisme engendre la guerre. Dimanche #marine2017
- Bonne chance !! Moi #JeVoteMarine #Marine2017 ???? #LaFranceVoteMarine #AuNom-DuPeuple ????
- Mélenchon n'a prononcé le mot "islamiste" dans aucun de ses meetings, pas une fois. Vu la période 2012-2017, c'est presque de la complicité.
- Va falloir l'éduquer cette jeunesse... #JeVoteMarine #AuNomDuPeuple
- Hors antenne sur #BeurFM, #Macron reconnaît que #MohamedSaou est un #islamiste radical, mais un type "très bien"

#### Tweets de la communauté Macron

- On dirait que #EmmanuelHollande commence à horripiler les Français... #Macron
- Je propose de voter #Fillon pour des mesures fortes et de bons sens pour l'#education #jevotefillon
- "MACRON, PRÉSIDENT, MACRON, PRÉSIDENT, MACRON, PRÉSIDENT" #MacronPrésident
- Une raison supplémentaire de ne pas voter Macron
- Macron , je ne le sens pas ce type là, il se regarde parlé, il ne doit pas être net

#### Tweets de la communauté Mélenchon

- Le vote @JLMelenchon pour donner plus de forces à nos luttes pour le progrès social !
- #Mélenchon: Il faut un "protectionnisme solidaire" pour défendre les producteurs français. #DemainPresident #JLM2017
- L'#AvenirEnCommun s'offre dans cette élection. Ensemble, il est à portée de #vote. A dimanche, les gens! #JLM2017
- @SImLyon @BobLeCentriste Être de gauche, c'est avoir du respect pour tous, et pas le mépris social affiché par macron
- L'avenir de la France ENTIÈRE C'est le vote Mélenchon ! #JLM2017 #LaForceDuPeuple

Une rapide analyse qualitative de ces tweets permet de mettre en avant les forces et les faiblesses de l'hypothèse qu'un utilisateur faisant partie d'une communauté soutient le candidat associé. En effet, la grande majorité des tweets provenant des communautés de B. Hamon, M. Le Pen et J-L. Mélenchon émettent sans aucun doute des opinions favorables à l'égard de leur candidat, et défavorables à l'égard des autres; notre hypothèse est donc justifiée pour ces communautés. En revanche, une partie non-négligeable des tweets extraits de la communauté de F. Fillon expriment un sentiment négatif envers ce candidat. Celle d'E. Macron ne comprend quasiment que des opinions défavorables à son égard, et semble même être une communauté anti-Macron. Ce dernier cas montre que l'hypothèse que nous avons émise peut être trop forte : l'algorithme de Louvain permet d'extraire des communautés de partisans politiques, tout aussi bien que des communautés de détracteurs.

Cette approche par étude de graphe de communautés nous permet donc d'extraire des groupes de partisans politiques sur lesquels nous avons pu mener quelques analyses. Cependant, extraire les communautés du graphe des utilisateurs ne suffit pas pour pouvoir répondre à notre problématique de classification de sentiments. Néanmoins, ces communautés ainsi que notre hypothèse pour l'affectation d'une communauté à son candidat induisent une pseudo-supervision : nous savons quels utilisateurs sont favorables à quel candidat, et pouvons étiqueter ses tweets en conséquence.

## 4 Fusion des deux approches

La faible supervision induite par l’analyse de graphes de communautés nous permet de mettre en place une chaîne de traitements standard. Nous considérerons d’abord un modèle « global » permettant de prédire la polarité d’un tweet, puis plusieurs classifieurs « par candidat », qui déterminent l’opinion d’un utilisateur vis-à-vis d’un candidat en particulier.

### 4.1 Classification de sentiments

L’objectif est ici de déterminer la polarité d’un tweet. L’hypothèse émise dans la section 3 nous fournit un moyen d’étiqueter nos données : étant donné un ensemble de tweets issus d’une communauté regroupée autour d’un candidat, ceux qui mentionnent uniquement ce candidat sont considérés positifs, et les autres, négatifs. Le tableau 9 recense le pourcentage de tweets positifs au sein de chaque communauté et pour toutes communautés confondues. Comme les classes ne sont pas équilibrées, nous évaluerons notre modèle avec le score F1.

| Candidat               | Pourcentage de tweets positifs |
|------------------------|--------------------------------|
| Fillon                 | 46.7%                          |
| Hamon                  | 53.5%                          |
| Le Pen                 | 31.5%                          |
| Macron                 | 41.1%                          |
| Mélenchon              | 53.5%                          |
| Tout candidat confondu | 45.8 %                         |

TABLEAU 9 – Pourcentage de tweets positifs dans chaque communauté

Pareillement à ce qui est fait dans la section 2, nous mettons en place une chaîne de traitements standard. Après optimisation de notre modèle, nous obtenons un score F1 de 65.6% en test. Nous ne pouvons néanmoins pas garantir totalement la fiabilité de ce résultat car les étiquettes affectées aux tweets n’étant que des *pseudo-labels*, certains tweets peuvent être mal annotés.

Pour estimer à vue d’œil la qualité de notre modèle, nous évaluons ses prédictions sur des tweets, fournis en annexe A, que nous jugeons positifs, négatifs, neutres et sarcastiques. Les prédictions figurent dans le tableau 10.

| Nature du tweet | +   | −   |
|-----------------|-----|-----|
| Positif         | 2/5 | 3/5 |
| Négatif         | 2/5 | 3/5 |
| Neutre          | 0/2 | 2/2 |
| Sarcastique     | 0/2 | 2/2 |

TABLEAU 10 – Prédications du modèle pour chaque nature de tweet

Le modèle considère les tweets neutres et sarcastiques comme étant négatifs. Sur les tweets positifs et négatifs, nous obtenons les mêmes résultats : deux tweets sur cinq jugés positifs, les trois autres jugés négatifs.

Nous souhaitons comprendre quels termes du dictionnaire définissent la polarité d’un tweet selon notre classifieur. Dans le tableau 11 figurent les mots les plus polarisés. Comme les tweets dont nous disposons correspondent en majorité à des débats entre utilisateurs et des comparaisons entre candidats, les termes extraits sont essentiellement des entités nommées liées aux candidats. Ces derniers déterminent donc la polarité d’un tweet : par exemple, si « dupontaigan » apparaît, le tweet en question est probablement négatif, l’inverse s’il s’agit de « francinsoumis ».

| Avis    | Termes  |
|---------|---|
| Négatif | dupontaignan, clairock, feedly,<br>fillonpresident, scolaire,<br>jevotefillonpresident, erichacquemand,<br>jacqu, stopmelenchon, benoît |
| Positif | jlmtoulouse, franceinsoumis, bisounour,<br>corrompu, marineàpar, manson_fx,<br>ptit, emmanuelholland, hamon2017, stopmacron             |

TABLEAU 11 – Termes les plus polarisés

Ces termes étant liés aux candidats et non à des émotions particulières, il nous semble intéressant de construire un modèle de classification de sentiments par candidat, afin de déterminer l’opinion politique d’un utilisateur.

## 4.2 Prédiction d’opinions

Nous souhaitons affiner la classification : l’objectif est désormais de prédire, étant donné un tweet, l’opinion d’un utilisateur vis-à-vis d’un candidat donné. Nous construisons donc un classifieur par candidat. Pour chaque classifieur, les données sont étiquetées de sorte à ce que les tweets issus de la communauté du candidat et ne mentionnant que ce dernier sont considérés positifs, et tous les autres, négatifs. Comme le montre le tableau 13, cet étiquetage induit un fort déséquilibre entre les deux classes et il est donc nécessaire d’évaluer les modèles construits avec la mesure F1.

| Candidat spécifique au classifieur | Pourcentage de tweets positifs |
|------------------------------------|--------------------------------|
| Fillon                             | 8.2%                           |
| Hamon                              | 8.3%                           |
| Le Pen                             | 5.6%                           |
| Macron                             | 11.2%                          |
| Mélenchon                          | 17.9%                          |

TABLEAU 12 – Pourcentage de tweets positifs pour chaque étiquetage des tweets

Les performances de chaque classifieur sont présentées dans le tableau 13. Nous remarquons que, plus la communauté soutient son candidat, meilleur est le résultat : les scores F1 correspondant à B. Hamon, M. Le Pen et J-L. Mélenchon sont les plus élevés et nous avons déduit dans la section 3 que les communautés liées à ces candidats sont essentiellement constituées de partisans politiques. Au contraire, l’évaluation sur les communautés de F. Fillon et E. Macron, plus hétérogènes, fournit des scores plus bas.

| Candidat spécifique au classifieur | Score F1 |
|------------------------------------|----------|
| Fillon                             | 50.3%    |
| Hamon                              | 80.1%    |
| Le Pen                             | 63.1%    |
| Macron                             | 53.0%    |
| Mélenchon                          | 76.5%    |

TABLEAU 13 – Score F1 de chaque modèle de classification

Dans la suite de cette section, nous n’étudierons que classifieurs liés à J-L. Mélenchon, dont la communauté comprend essentiellement des partisans politiques, et E. Macron, dont la communauté contient vraisemblablement beaucoup de détracteurs.

| Sentiment | Termes   |
|-----------|--|
| Négatif   | macron, fillon, hamon, hamon2017, clairock, marin, envi, francoisfillon, mlp, mlp_officiel   |
| Positif   | jlm_2017, penicheinsoumis, avenirencommun, révolutionfiscal, jlmtoulous, insoumis, manson_fx, jlm2017, franceinsoumis, jlmdesintox |

TABLEAU 14 – Termes les plus polarisés selon le classifieur spécifique à la communauté Mélenchon

Les termes les plus polarisés selon le classifieur lié à la communauté de J-L. Mélenchon figurent dans le tableau 14. Les entités nommées correspondant aux autres candidats définissent le caractère négatif d'un tweet, alors que les entités associées à J-L. Mélenchon sont des marqueurs de sentiments positifs pour ce modèle d'apprentissage. Ces résultats confirment l'hypothèse émise, dans la section 3 : la communauté associée à J-L. Mélenchon est constituée d'utilisateurs qui le soutiennent.

| Sentiment | Termes  |
|-----------|---|
| Négatif   | fillon, hamon, francoisfillon, jlm2017, aucoeurdudébat, marin, jevotefillon, pen, hamon2017, édito              |
| Positif   | consid, ridiculetv, impostur, incohérent, étouff, bariza_khiar, saougat, emmanuelholland, stopmacron, bisounour |

TABLEAU 15 – Termes les plus polarisés selon le classifieur spécifique à la communauté Macron

En revanche, nous ne retrouvons pas le phénomène précédent dans la communauté liée à E. Macron. D'après le tableau 15, les termes négatifs correspondent bien aux autres candidats, mais les termes positifs sont des marqueurs d'opposition à E. Macron. Ces résultats sont cohérents : les tweets annotés positifs ne mentionnent que ce candidat et sont en réalité des commentaires négatifs envers ce dernier, les tweets annotés négatifs mentionnent tout autre candidat. Nous pouvons confirmer ce qui est dit dans la section 3 : la communauté associée à E. Macron est constituée en grande partie de détracteurs.

Nous récupérons ensuite, pour chaque classifieur, quelques tweets, fournis en annexes B et C, jugés positifs, négatifs et neutres envers le candidat en question. Les prédictions de chaque modèle sur leur ensemble de tweets figurent dans les tableaux 16 et 17.

| Nature du tweet | +   | -   |
|-----------------|-----|-----|
| Positif         | 1/2 | 1/2 |
| Négatif         | 0/2 | 2/2 |
| Neutre          | 0/2 | 2/2 |

TABLEAU 16 – Prédiction du modèle associé à la communauté de J-L. Mélenchon

D'après le tableau 16, le modèle d'apprentissage propre à J-L. Mélenchon classe correctement les tweets négatifs, annote négativement les tweets neutres et considère un seul des deux tweets positifs comme étant positif.



| Nature du tweet | +   | −   |
|-----------------|-----|-----|
| Positif         | 0/2 | 2/2 |
| Négatif         | 1/2 | 1/2 |
| Neutre          | 0/2 | 2/2 |

TABLEAU 17 – Prédications du modèle associé à la communauté d’E. Macron

Sur les quelques tweets qui lui sont associés, les résultats contenus dans le tableau 17 montrent que le modèle associé à E. Macron est moins performant : aucun des tweets positifs n’est correctement classifié et un tweet négatif sur deux est mal classé. Cette performance est en accord avec les analyses faites précédemment sur cette communauté. Quant aux tweets neutres, ils sont aussi considérés comme étant négatifs.

### 4.3 Bilan

Si les communautés sont extraites de sorte à ce qu’un utilisateur d’une communauté donnée soutient en effet le candidat associé, alors la méthode fournie dans cette section montre qu’il est possible de classifier les données texte issues d’un corpus de tweets politiques sans qu’il soit nécessaire de les annoter manuellement.

## 5 Conclusion et perspectives

Ce projet nous a permis de travailler sur des données textuelles issues de Twitter et de mettre en œuvre différentes stratégies de classification de sentiments. Nous avons considéré une approche basée sur la problématique de *transfer learning*, exploité les interactions entre utilisateurs pour extraire des communautés regroupées autour des candidats, puis tiré profit de la faible supervision induite par ces communautés pour classer les tweets selon le sentiment de l'utilisateur puis selon son opinion politique.

Il a été difficile d'appliquer des connaissances apprises sur un domaine à un autre et d'évaluer les modèles obtenus, d'autant plus lorsque les domaines sont asymétriques et qu'aucune supervision n'est disponible. Nous avons cependant montré qu'il est possible, à partir d'un graphe d'utilisateurs, d'annoter nos tweets. Les expériences menées sur ces données faiblement supervisées attestent de l'efficacité de cette méthode mais aussi de ses limites.

Nous aimerions raffiner l'extraction des communautés, par exemple en fusionnant les communautés soutenant le même candidat, afin d'obtenir une pseudo-supervision plus fiable. Nous souhaiterions aussi étudier plus en détails les techniques d'apprentissage par transfert. Dans l'objectif d'améliorer la pertinence du vocabulaire appris, nous pourrions aussi retravailler le pré-traitement du texte, en particulier en affinant la reconnaissance d'entités nommées.

## A Sélection de tweets de test pour l'évaluation du classifieur de sentiments

- Tweets positifs :
  - La #jeunesse : vous aurez un avenir avec @FrancoisFillon. #Fillon2017 #JeVoteFillon
  - #JeVotePour #Hamon2017 #BenoitHamon2017 le candidat de la jeunesse
  - Pour l'avenir de mon pays dans 10 jours je vote #Marine2017
  - #JeVoteMacron 10 raisons de voter Macron ! #EnMarche #Macron2017 #Presidentielle2017
  - La révolution citoyenne est train de se produire ! #Melenchon2017
- Tweets négatifs :
  - @FrancoisFillon La seule question qu'on doit vous poser c'est : QUAND EST CE QUE VOUS NE MENTEZ PAS ???!
  - #Hamon2017 Désigner un frondeur pour rassembler la gauche ? Une erreur de casting #macron2017 #Presidentielle2017
  - Quand Marine rêve son Monde #Marine2017 ! il y a des rires glaçants, des coups sanglants et un peuple réduit à néant
  - @BFMTV macron ferme la ! tu tacles le Fn mais tu es 1salopard corrompu Les français s'en rendront compte vermine
  - REPUGNANT ! MELENCHON EMBRASSE CHAVEZ ! Nous allons cracher du sang si MELENCHON est élu !
- Tweets sarcastiques :
  - Best-of du grand poète Macron
  - @FrancoisFillon Achetez-vous un parachute M. Fillon
- Tweets neutres :
  - 'La vie d'un entrepreneur est souvent plus difficile que celle d'un salarié'- Macron
  - #Rediff François Fillon refuse de se rendre chez Jean-Jacques Bourdin

## B Sélection de tweets de test pour l'évaluation du classifieur associé à J-L. Mélenchon

- Tweets positifs :
  - Ne vous trompez pas de bulletin de vote #JLM2017 pour la paix les autres c'est la guerre
  - La révolution citoyenne est train de se produire ! #Melenchon2017
- Tweets négatifs :
  - REPUGNANT ! MELENCHON EMBRASSE CHAVEZ ! Nous allons cracher du sang si MELENCHON est élu !
  - Pratiques indignes! Voilà le vrai visage de cette gauche communiste extrémiste #melanchon #JLM2017 #stopmelenchon
- Tweets neutres :
  - #Rediff François Fillon refuse de se rendre chez Jean-Jacques Bourdin
  - mdr le papillon dans fantasy life j'ai cru il s'appelait fillon

## C Sélection de tweets de test pour l'évaluation du classifieur associé à E. Macron

- Tweets positifs :
  - #JeVoteMacron 10 raisons de voter Macron ! #EnMarche #Macron2017 #Presidentielle2017
  - Justement si #macron2017 a annoncé beaucoup de mesures pour les handicapés... #JeVoteMacron
- Tweets négatifs :
  - Emmanuel #Macron l'escroquerie en marche
  - Je lis le programme de #Macron... Ça ratisse super large...Mais, ça se finance comment ?? #Marine2017
- Tweets neutres :
  - #Rediff François Fillon refuse de se rendre chez Jean-Jacques Bourdin
  - mdr le papillon dans fantasy life j'ai cru il s'appelait fillon

## Références

- Bespalov, D., Bai, B., Qi, Y. & Shokoufandeh, A. (2011). Sentiment classification based on supervised latent n-gram analysis. In *Proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, p. 375–382.
- Blitzer, J., Dredze, M. & Pereira, F. (2007). Biographies, bollywood, boom-boxes and blenders : Domain adaptation for sentiment classification. In *Proceedings of the 45th annual meeting of the association of computational linguistics*. Association for Computational Linguistics, p. 440–447.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. (2008). Fast unfolding of communities in large networks. In *Journal of statistical mechanics : theory and experiment* 2008.10, P10008.
- Hagberg, A., Swart, P. & S Chult, D. (2008). *Exploring network structure, dynamics, and function using NetworkX*. Rapp. tech. Los Alamos National Lab.(LANL), Los Alamos, NM (United States).
- Mejova, Y. A. (2012). Sentiment analysis within and across social media streams. In.
- Opsahl, T., Agneessens, F. & Skvoretz, J. (2010). Node centrality in weighted networks : Generalizing degree and shortest paths. In *Social networks* 32.3, p. 245–251.
- Pan, S. J. & Yang, Q. (2009). A survey on transfer learning. In *IEEE Transactions on knowledge and data engineering* 22.10, p. 1345–1359.
- Pang, B. & Lee, L. (2008). Opinion mining and sentiment analysis. In *Foundations and Trends® in Information Retrieval* 2.1–2, p. 1–135.
- Pang, B., Lee, L. & Vaithyanathan, S. (2002). Thumbs up ? : sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*. T. 10. Association for Computational Linguistics, p. 79–86.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. et al. (2011). Scikit-learn : Machine learning in Python. In *Journal of machine learning research* 12.Oct, p. 2825–2830.
- Rafrafi, A., Guigue, V. & Gallinari, P. (2013). Classification de Sentiments Multi-Domains en Contexte Hétérogène et Passage à l’Echelle. In *CORIA*.