



**HAL**  
open science

## Détection de mésusages de médicaments dans les réseaux sociaux

Elise Bigeard, Natalia Grabar, Frantz Thiessard

► **To cite this version:**

Elise Bigeard, Natalia Grabar, Frantz Thiessard. Détection de mésusages de médicaments dans les réseaux sociaux. TALN, May 2018, Rennes, France. halshs-01968335

**HAL Id: halshs-01968335**

**<https://shs.hal.science/halshs-01968335>**

Submitted on 2 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Détection de mésusages de médicaments dans les réseaux sociaux

Elise Bigeard<sup>1,2</sup> Natalia Grabar<sup>1</sup> Frantz Thiessard<sup>2,3</sup>

(1) CNRS, Univ Lille, UMR 8163 STL - Savoirs Textes Langage, F-59000 Lille, France

(2) Univ. Bordeaux, Inserm, Bordeaux Population Health Research Center, team ERIAS, UMR 1219, F-33000 Bordeaux, France

(3) CHU de Bordeaux, Pole de sante publique, Service d'information medicale, F-33000 Bordeaux, France  
elise.bigeard@u-bordeaux.fr

## RÉSUMÉ

---

Un mésusage apparaît lorsqu'un patient ne respecte pas sa prescription et fait des actions pouvant mener à des effets nocifs. Bien que ces situations soient dangereuses, les patients ne signalent généralement pas les mésusages à leurs médecins. Il est donc nécessaire d'étudier d'autres sources d'information pour découvrir ce qui se passe en réalité. Nous proposons d'étudier les forums de santé en ligne. L'objectif de notre travail consiste à explorer les forums de santé avec des méthodes de classification supervisée afin d'identifier les messages contenant un mésusage de médicament. Notre méthode permet de détecter les mésusages avec une F-mesure allant jusqu'à 0,810. Cette méthode peut aider dans la détection de mésusages et la construction d'un corpus exploitable par les experts pour étudier les types de mésusages commis par les patients.

## ABSTRACT

---

### **Detection of drug misuse in social media.**

Misuses happen when patients do not follow the prescriptions and do actions which lead to potentially harmful situations. Although such situations are dangerous, patients usually do not report the misuse of drugs to their physicians. Hence, it is necessary to study other sources of information to discover what is happening in reality. We propose to study online health fora. The purpose of our work is to explore online health fora with supervised classification methods in order to identify messages that contain drug misuses. Our method permits to detect misuses with up to 0.810 F-measure. This method can help the detection of messages with misuses and building corpus with drug misuses, which can be used by experts to further study the types of misuses committed by patients.

---

**MOTS-CLÉS :** TAL, forums de discussion en santé, classification supervisée, domaine médical, mésusage de médicaments.

**KEYWORDS:** NLP, health discussion forums, supervised classification, medical domain, drug misuse.

---

## 1 Introduction

Un mésusage de médicament apparaît lorsqu'un patient ne respecte pas sa prescription : sous-dosage, sur-dosage, utilisation de médicaments pour des raisons autres de celles de la prescription, consommation de médicaments prescrits pour une autre personne, etc. Ces situations sont dangereuses car elles mettent en danger la santé de la personne. Comme les patients reportent rarement les mésusages à leurs médecins, il est nécessaire de consulter d'autres sources d'informations pour

découvrir ce qui se passe en réalité. Nous proposons d'étudier les réseaux sociaux, où les patients communiquent librement et facilement sur leur processus de santé (Gauducheau, 2008), et sans doute sur l'usage de médicaments. Actuellement, les réseaux sociaux sont largement étudiés par plusieurs disciplines et en poursuivant différents objectifs : identification de géolocalisation, fouille d'opinions, extraction d'événements, traduction et résumé automatique... (Louis, 2016). Dans le domaine médical, qui est au centre de notre travail, les réseaux sociaux peuvent être exploités pour fournir des informations pour la surveillance épidémiologique (Collier, 2011; Lejeune *et al.*, 2013), la qualité de vie des patients (Tapi Nzali, 2017) et les effets indésirables de médicaments (Morlane-Hondère *et al.*, 2016). Cependant, peu de travaux s'intéressent au mésusage de médicaments. Nous pouvons citer par exemple l'étude des tweets concernant l'usage non-médical de médicaments avec des méthodes non supervisées (Kalyanam *et al.*, 2017) et la création d'une plateforme générique pour l'étude de sur-usages (Cameron *et al.*, 2013).

L'objectif général de notre travail consiste également à aider l'étude des mésusages commis par les patients. Notre travail propose d'identifier, au sein des forums de discussion, les messages contenant un mésusage. Dans la suite de ce travail, nous présentons d'abord le matériel (section 2) et les étapes de la méthode (section 3). Nous présentons et discutons les résultats ensuite (section 4).

## 2 Matériel

Nous utilisons plusieurs types de matériel décrit dans cette section : un corpus de messages postés sur des forums de santé, un lexique de médicaments et un lexique de maladies.

**Corpus.** Nous construisons le corpus à partir de deux forums de Doctissimo : médicaments<sup>1</sup> et grossesse<sup>2</sup>. Les messages collectés ont été postés entre 2010 et 2015. Nous conservons uniquement les messages contenant au moins un nom de médicament et excluons les messages de plus de 2 500 caractères, dont le contenu est hétérogène et difficile à analyser. Le corpus obtenu contient 119 562 messages (15,5M mots).

**Noms de médicaments.** Nous utilisons un ensemble de noms commerciaux de médicaments associés à leur code ATC (Skrbo *et al.*, 2004) provenant de différentes sources : la base CNHIM Thériaque<sup>3</sup>, la *base publique du médicament*<sup>4</sup> et la base *Medic'AM* de l'Assurance Maladie<sup>5</sup>. Ce lexique contient 4 133 noms de médicaments répartis en 1 690 codes ATC distincts.

**Noms de maladies.** Nous utilisons un ensemble de 29 maladies traitées par les médicaments anxiolytiques et antidépresseurs. Les maladies sont associées à leurs codes CIM-10 (OMS, 1995) : *panique/F41.0*, *anxiété/F41.9*... Ce codage, employé par les professionnels de santé, fait une distinction fine des maladies, qui reste difficile à faire pour les patients. En effet, les non-spécialistes peuvent confondre les maladies, comme par exemple *agoraphobie/F40.0* et *phobie sociale/F40.1*, *panique/F41.0* et *anxiété/F41.9*. Un groupement simplifié est donc effectué par les experts. Ce lexique source est étendu dans une expérience antérieure grâce à l'utilisation de lexiques existants et à l'analyse du corpus (Bigeard, 2017).

---

1. [http://forum.doctissimo.fr/medicaments/liste\\\_categorie.htm](http://forum.doctissimo.fr/medicaments/liste\_categorie.htm)

2. [http://forum.doctissimo.fr/grossesse-bebe/liste\\\_categorie.htm](http://forum.doctissimo.fr/grossesse-bebe/liste\_categorie.htm)

3. <http://www.theriaque.org>

4. <http://base-donnees-publique.medicaments.gouv.fr>

5. <https://www.ameli.fr/l-assurance-maladie/statistiques-et-publications/donnees-statistiques/medicament/medic-am/medic-am-mensuel-2017.php>

## 3 Méthodes

### 3.1 Prétraitements

Les messages sont segmentés en mots, étiquetés et lemmatisés avec Treetagger (Schmid, 1994). La casse est neutralisée et les diacritiques sont supprimées pour diminuer la variation orthographique, comme {*Anxiété, anxiète*}. Aucune correction orthographique n'est effectuée. Les mots grammaticaux (articles, prépositions, verbes auxiliaires...) sont supprimés.

Nous avons effectué des expériences avec plusieurs formes du texte :

- texte non lemmatisé,
- texte lemmatisé, chiffres remplacés par la séquence *nombre*,
- texte lemmatisé, mots grammaticaux supprimés.

Les messages sont annotés avec les lexiques de médicaments et de maladies, et indexés selon le code correspondant : CIM-10 pour les maladies et ATC pour les médicaments. Nous pouvons ainsi remarquer que certaines classes de médicaments sont très fréquentes, avec 60 % des messages dédiés à la pilule contraceptive et 15 % aux antidépresseurs et anxiolytiques.

### 3.2 Annotation manuelle et Corpus de référence

Afin de constituer un corpus d'entraînement et de test, une annotation manuelle est réalisée. Trois annotateurs ont pour objectif d'associer chaque message à l'une des catégories suivantes :

**usage normal** : contient un usage normal de médicament, comme dans : *Mais la question que je pose est 'est ce que c'est normal que le loxapac que je prends met des heures à agir ? ?*

**pas d'usage** : ne contient pas d'usage de médicament, comme dans : *ouf boo, repose toi surtout, il ne t'a pas prescrit d'aspegic nourisson ? ?*

**mésusage** : contient un mésusage. Lorsque cette catégorie est sélectionnée, l'annotateur explique brièvement en quoi consiste le mésusage (sur-dosage, sevrage brutal...). Dans l'exemple qui suit, le mésusage est dû à un oubli de prise : *bon moi la miss boulette et la tete en l'air je devais commencer mon "utrogestran 200" a j16 bien sur j'ai oublier! donc je l'ai pris ce soir!!!!*

**incertitude** : impossible de décider

Trois annotateurs participent à la tâche : un pharmacologue et deux informaticiens familiers avec les textes médicaux. Deux annotateurs (un pharmacologue et un informaticien) effectuent l'annotation, alors que le troisième vérifie le contenu de la catégorie de mésusage. Les cas de désaccord ou de messages annotés *mésusage* sont discutés entre les annotateurs pour trouver la catégorie consensuelle.

Les données de référence sont constituées à partir de trois corpus :

- *C1* contient 150 messages sélectionnés aléatoirement. Chaque message est annoté indépendamment par deux annotateurs. Nous utilisons cette annotation pour calculer l'accord inter-annotateur selon la mesure du Kappa (Cohen, 1960). En cas de désaccord, les annotateurs se concertent et décident ensemble de la catégorie finale ;
- *C2* contient 1 200 messages sélectionnés aléatoirement. Il est divisé en deux et chaque partie est annotée par l'un des deux annotateurs ;
- *C3* contient 500 messages. Puisque certaines classes de médicaments sont plus fréquentes, nous construisons *C3* pour qu'il contienne des médicaments plus variés : pour chacune des

50 classes les plus fréquentes dans le corpus, 10 messages sont sélectionnés aléatoirement. Ce corpus permet de diversifier le contenu étudié. Il est annoté par l'expert en pharmacologie. Le corpus total contient 1 850 messages (202 726 mots) dont 600 dans la catégorie *pas d'usage*, 1 117 dans la catégorie *usage normal* et 133 dans la catégorie *mésusage*.

### 3.3 Catégorisation supervisée

Nous utilisons l'implémentation Weka (Witten & Frank, 2005) de divers algorithmes de catégorisation supervisée : NaiveBayes (John & Langley, 1995), Bayes Multinomial (McCallum & Nigam, 1998), J48 (Quinlan, 1993), Random Forest (Breiman, 2001) et Simple Logistic (Landwehr *et al.*, 2005). Nous utilisons trois types de descripteurs : texte lemmatisé et vectorisé, classes ATC des médicaments, classes CIM-10 des maladies.

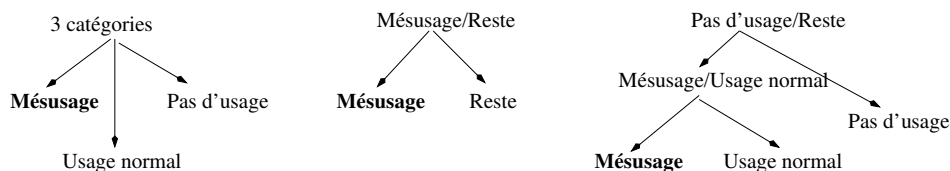


FIGURE 1 – Organisation des expériences pour la détection des mésusages

L'objectif des expériences est d'identifier les messages contenant un mésusage de médicament. Trois chemins sont possibles pour y arriver (figure 1) :

- *Trois catégories*. Nous utilisons directement les trois catégories de l'annotation. Chaque catégorie contient 133 messages. Cette distinction est sans doute la plus difficile car le modèle doit différencier les trois catégories en même temps ;
- *Catégorisation binaire mésusage-reste*. Ce modèle contraste la catégorie *mésusage* avec les deux autres (*usage normal* et *pas d'usage*). Le corpus contient 133 messages dans la catégorie *mésusage* et 133 messages dans les deux autres catégories. Ce modèle est la méthode la plus directe pour détecter les mésusages ;
- *Catégorisation binaire pas d'usage-reste* suivie de *catégorisation binaire usage normal-mésusage*. Cette expérience est effectuée en deux temps : il s'agit d'abord d'isoler les cas de *pas d'usage* (le corpus contient alors 2\*300 messages) et ensuite de distinguer entre les cas de *mésusage* et d'*usage normal*. Il s'agit d'un modèle en cascade. En raison du faible nombre de messages de type *mésusage*, lors de l'évaluation de la seconde étape, nous exploitons l'ensemble des messages de types *mésusage*, ce qui nous fournit un corpus de 2\*133 messages.

Nous effectuons quatre expériences principales, où nous utilisons les descripteurs suivants :

- *texte* : le texte lemmatisé et vectorisé uniquement ;
- *médicaments* : le texte lemmatisé et vectorisé, avec les codes de médicaments ;
- *maladies* : le texte lemmatisé et vectorisé, avec les codes de maladies ;
- *médicaments+maladies* : le texte lemmatisé et vectorisé, avec les codes de médicaments et de maladies.

Ces descripteurs permettront d'observer l'impact des types de médicaments et de maladies par rapport au seul texte des messages. Pour mieux comprendre le rôle de médicaments et de maladies,

TABLE 1 – Résultats pour l'expérience mésusage / reste avec le descripteur médicaments

	NaiveBayes			NaiveBayes Multinomial		
	Précision	Rappel	F-mesure	Précision	Rappel	F-mesure
formes	0,735	0,734	0,733	0,669	0,646	0,627
lemmes	0,778	0,772	0,773	0,608	0,582	0,579
lemmes lexicaux	0,821	0,810	0,809	0,812	0,810	<b>0,810</b>

TABLE 2 – Résultats pour l'expérience mésusage / usage avec le descripteur médicaments

	NaiveBayes			NaiveBayes Multinomial		
	Précision	Rappel	F-mesure	Précision	Rappel	F-mesure
formes	0,733	0,734	0,734	0,641	0,646	0,633
lemmes	0,782	0,772	0,772	0,787	0,759	0,758
lemmes lexicaux	0,801	0,797	<b>0,797</b>	0,661	0,595	0,560

nous effectuons deux ensembles d'expériences supplémentaires, pour les médicaments et pour les maladies :

- *normal* : le texte des messages lemmatisé et vectorisé ;
- *code* : les noms de médicaments ou de maladies sont remplacés par leurs codes ATC ou CIM-10 ;
- *normal+code* : le texte des messages contient les noms de médicaments et maladies, et leurs codes sont ajoutés en plus ;
- *substitution* : les noms de médicaments et de maladies sont remplacés par la séquence *drug* ou *disorder* ;
- *supprimé* : les noms de médicaments et de maladies sont supprimés.

Nous ne disposons que de 133 messages dans la catégorie *mésusage*, qui est aussi la catégorie que nous cherchons à isoler. Pour chaque expérience, le nombre de messages de chaque catégorie est équilibré afin que les classifieurs accordent de l'importance à la catégorie *mésusage*. Les messages sont sélectionnés aléatoirement. Un même corpus est utilisé pour chaque séquence d'expériences. 70 % du corpus est utilisé pour l'entraînement et 30 % pour l'évaluation. Nous ne faisons donc pas de cross-validation. Les résultats sont évalués avec des mesures classiques (Sebastiani, 2002) : vrais positifs pour les messages correctement classifiés ; faux négatifs pour les messages non détectés ; faux positifs pour les messages détectés à tort ; de même que la précision  $P$ , le rappel  $R$  et la F-mesure  $F$ .

## 4 Résultats et Discussion

L'accord inter-annotateur est de 0,46, ce qui est un accord modéré (Landis & Koch, 1977) et indique qu'il s'agit d'une tâche potentiellement difficile. En effet, il existe de nombreux cas où il est difficile de décider si le mésusage a lieu ou pas : par exemple, lorsque le patient exprime une intention de commettre un mésusage ou un questionnement qui peut mener à un mésusage selon la réponse obtenue. Dans cet exemple "*mon psy m'a prescrit effexor enfin son generique a 37.5 je le prend ou pas ? j'ai trop peur des effet secondaire*" si la personne décide de ne pas prendre un médicament qui lui a été prescrit elle se retrouve en situation de mésusage, mais l'usage restera normal si la personne suit la prescription.

Les expériences de détection automatique de mésusages montrent que la *catégorisation binaire mésusage-reste* atteint le meilleur résultat avec 0,810 de F-mesure (l’algorithme NaiveBayes Multinomial, le texte lemmatisé, sans les mots grammaticaux et avec les descripteurs *médicaments*), comme présenté dans le tableau 1. La catégorisation en cascade obtient 0,756 de F-mesure pour la tâche *pas d’usage / reste* (l’algorithme NaiveBayes, le texte lemmatisé, sans les mots grammaticaux, avec les descripteurs *médicaments*) et 0,797 de F-mesure pour la tâche *usage / mésusage* (les mêmes paramètres), comme indiqué dans le tableau 2. Le modèle *trois catégories* atteint au mieux 0,607 de F-mesure (l’algorithme NaiveBayes Multinomial, le texte lemmatisé, les descripteurs *médicaments*).

TABLE 3 – Résultats de l’expérience sur les médicaments, avec texte lemmatisé, algorithme de type NaiveBayes, exprimés en F-mesure

descripteurs	3 classes	mésusage/reste	pas d’usage/reste	usage/mésusage
normal	0,544	0,734	<b>0,713</b>	0,741
code	0,546	0,728	0,700	0,739
normal + code	<b>0,563</b>	<b>0,740</b>	0,684	<b>0,758</b>
substitution	0,540	0,731	0,702	0,721
supprimé	0,554	0,731	0,694	0,721

TABLE 4 – Résultats de l’expérience sur les maladies, avec texte lemmatisé, algorithme de type NaiveBayes, exprimés en F-mesure

descripteurs	3 classes	mésusage/reste	pas d’usage/reste	usage/mésusage
normal	0,577	0,763	<b>0,768</b>	<b>0,763</b>
code	0,542	0,720	0,750	0,651
normal + code	<b>0,579</b>	<b>0,793</b>	0,749	<b>0,763</b>
substitution	0,554	0,734	0,755	0,661
supprimé	0,544	0,730	0,751	0,676

Les résultats des expériences sur les médicaments et les maladies sont présentés dans les tableaux 3 et 4. Pour chaque expérience, la F-mesure obtenue avec le meilleur algorithme (NaiveBayes ou NaiveBayes Multinomial) est présentée. Le meilleur résultat pour chaque tâche est mis en valeur en gras, ce qui nous permet de constater que le descripteur *normal* ou *normal + code* dépasse toujours les autres ensembles de descripteurs. Il semble donc que le nom même du médicament ou de la maladie soient exploités par les classifieurs, plutôt que leur catégorie (descripteur *code*) ou leur simple présence (descripteur *substitution*). Cependant, on note que les résultats obtenus par les différents descripteurs sont proches les uns des autres, ce qui suggère que les noms des médicaments et des maladies ne sont que peu exploités par les classifieurs.

Notons de plus que (1) dans la majorité des expériences, les algorithmes bayésiens se montrent les plus efficaces, aussi avons-nous détaillé uniquement les résultats obtenus avec ces algorithmes ; (2) les noms de médicaments et de maladies ont un rôle positif sur la détection de mésusages ; (3) la lemmatisation et la suppression de mots grammaticaux ont un effet positif ; (4) la précision est en général plus élevée que le rappel, mais l’écart entre les deux reste faible.

Une analyse des faux positifs et des faux négatifs permet de faire les observations suivantes :

- Pour l’expérience *pas d’usage/reste*, 27 messages sont incorrectement classifiés dans *reste* et 33 dans *pas d’usage*. Parmi ces 33 messages, 11 ne contiennent pas d’information explicite sur la prise du médicament, comme dans *elina a quoi pour sa toux ? Ici antibio rebelotte*.

- Dans 5 messages, le médicament n'est pas cité mais reste sous-entendu à travers la maladie, comme dans *j'ai pris mon traitement et les allergies ça va mieux et aussi un spray nasal* ;
- Pour l'expérience *mésusage/reste*, 12 messages sont incorrectement classifiés dans *mésusage* et 9 dans *reste*. Parmi ces 12 messages, 4 contiennent des termes associés à l'excès et à des effets nocifs, comme dans *Je n'imaginais pas que c'était si grave* ou *s'il vous plait ne faites pas n'importe quoi*. Ils sont donc très similaires aux messages comportant des mésusages ;
  - Pour l'expérience *3 catégories*, 14 messages sont incorrectement classifiés dans *pas d'usage*, 11 dans *usage normal* et 20 dans *mésusage*. Hormis le fait que ce modèle est plus complexe à résoudre, y compris pour les cas de mésusage, il ne permet pas de dégager d'explications sur les erreurs commises.

## 5 Conclusion

Ce travail propose un ensemble d'expériences pour détecter les mésusages de médicaments dans les messages de forums de discussion. Des forums francophones de *Doctissimo* sont exploités. Les messages sont d'abord prétraités et indexés avec un lexique adapté. Nous effectuons ensuite plusieurs séries d'expériences de catégorisation supervisée avec pour objectif de catégoriser chaque message dans une des trois catégories : *pas d'usage*, *usage normal* et *mésusage*. Les descripteurs exploités sont : le texte lemmatisé et vectorisé sans les mots grammaticaux, le code CIM-10 pour les maladies et le code ATC pour les médicaments, et leurs différentes variations et combinaisons.

Les meilleurs résultats sont obtenus avec l'expérience qui différencie entre les messages contenant un mésusage et le reste de messages. La F-mesure atteint alors jusqu'à 0,810 points. Les algorithmes bayésiens sont les plus efficaces face à cette tâche. Nous effectuons également une série d'expériences complémentaires pour déterminer l'impact des médicaments et des maladies présents dans le texte sur la qualité de la catégorisation. Il apparaît que les noms de médicaments et de maladies ont une influence positive sur les résultats, bien qu'elle soit faible.

La limitation principale de ce travail est que le nombre de messages décrivant des mésusages n'est pas élevé, ce qui réduit le potentiel de la catégorisation supervisée. Malgré ceci, la méthode proposée peut être utilisée pour détecter les messages contenant un mésusage de médicament. Dans l'avenir, nous utiliserons cette méthode pour enrichir la catégorie des mésusages et améliorer les modèles de catégorisation. Ensuite, une catégorisation plus fine pourra être effectuée pour différencier les types de mésusages, comme ceux proposés dans un travail existant (Bigeard *et al.*, 2018).

## Remerciements

La présente publication s'inscrit dans le programme *Drugs Systematized Assessment in real-life Environment (DRUGS-SAFE)* financé par l'Agence Nationale de Sécurité du Médicament et des Produits de Santé. Cette publication ne représente pas nécessairement l'opinion de l'ANSM.

Nous remercions Pierre Simonetti pour son aide avec l'annotation de messages et l'équipe ERIAS pour les conseils et le soutien.



# Références

- BIGEARD E. (2017). Construction de lexiques pour l'extraction de maladies dans les forums santé. In *RECITAL 2017*, p. 1–12.
- BIGEARD E., GRABAR N. & THIESSARD F. (2018). Typology of drug misuse created from information available in health fora. In *MIE 2018*, p. 1–5.
- BREIMAN L. (2001). Random forests. *Machine Learning*, **45**(1), 5–32.
- CAMERON D., SMITH G. A., DANIULAITYTE R., SHETH A. P., DAVE D., CHEN L., ANAND G., CARLSON R., WATKINS K. Z. & FALCK R. (2013). PREDOSE : a semantic web platform for drug abuse epidemiology using social media. **46**(6), 985–997.
- COHEN J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**(1), 37–46.
- COLLIER N. (2011). Towards cross-lingual alerting for bursty epidemic events. *J Biomed Semantics*, **2**(5), S10.
- GAUDUCHEAU N. (2008). La communication des émotions dans les échanges médiatisés par ordinateur : bilan et perspectives. *Bulletin de Psychologie*, **61**(4), 389–404.
- JOHN G. H. & LANGLEY P. (1995). Estimating continuous distributions in bayesian classifiers. In M. KAUFMANN, Ed., *Eleventh Conference on Uncertainty in Artificial Intelligence*, p. 338–345, San Mateo.
- KALYANAM J., KATSUKI T., LANCKRIET G. R. G. & MACKEY T. K. (2017). Exploring trends of nonmedical use of prescription drugs and polydrug abuse in the twittersphere using unsupervised machine learning. **65**, 289–295.
- LANDIS J. & KOCH G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, **33**, 159–174.
- LANDWEHR N., HALL M. & FRANK E. (2005). Logistic model trees. *Machine Learning*, **95**(1-2), 161–205.
- LEJEUNE G., BRIXTTEL R., LECLUZE C., DOUCET A. & LUCAS N. (2013). Added-value of automatic multilingual text analysis for epidemic surveillance. In *Artificial Intelligence in Medicine (AIME)*.
- LOUIS A. (2016). Natural language processing for social media. *Computational Linguistics*, **42**(4), 833–836.
- MCCALLUM A. & NIGAM K. (1998). A comparison of event models for naive bayes text classification. In *AAAI workshop on Learning for Text Categorization*.
- MORLANE-HONDÈRE F., GROUIN C. & ZWEIGENBAUM P. (2016). Identification of drug-related medical conditions in social media. In *LREC*, p. 1–7.
- OMS (1995). *Classification statistique internationale des maladies et des problèmes de santé connexes — Dixième révision*. Organisation mondiale de la Santé, Genève.
- QUINLAN J. (1993). *C4.5 Programs for Machine Learning*. San Mateo, CA : Morgan Kaufmann.
- SCHMID H. (1994). Probabilistic part-of-speech tagging using decision trees. In *ICNMLP*, p. 44–49, Manchester, UK. treetagger.
- SEBASTIANI F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, **34**(1), 1–47.

SKRBO A., BEGOVIĆ B. & SKRBO S. (2004). Classification of drugs using the atc system (anatomic, therapeutic, chemical classification) and the latest changes. *Med Arh*, **58**(2), 138–41.

TAPI NZALI M. (2017). *Analyse des médias sociaux de santé pour évaluer la qualité de vie des patientes atteintes d'un cancer du sein*. Thèse de doctorat, Université de Montpellier, Montpellier, France.

WITTEN I. & FRANK E. (2005). *Data mining : Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco.