



Université Paul Valéry - Montpellier 3

UFR 6 - Éducation & Sciences pour les LLASHS

Master MIASHS

Prédiction des comportements suicidaires

RENDU 1 - QUESTION DE RECHERCHE



Équipe :

Lisa BÉTEILLE
Matéo CALSACY
Jean CHABANOL
Anamé ROUMY
Laura SÉNÉCAILLE
Célia TEYSSIER

2021/2022

Encadrants :

Sandra BRINGAY
Pierre LAFAYE de MICHEAUX
Florian LOMBARDO

CONTEXTE

Dans le cadre de notre Master MIASHS, les travaux d'études et de recherche (TER) nous permettent de nous initier à la recherche scientifique par une première expérience tuteurée, donnant lieu à un mémoire et une soutenance devant un jury composé de chercheurs et d'enseignants-chercheurs.

Pour cette composante de notre formation, nous avons choisi le sujet « Prédiction des comportements suicidaires ». Ce dernier est encadré par deux commanditaires :

Sandra Bringay, directrice de l'UFR 6 à l'Université Paul Valéry Montpellier III et également enseignante des modules en Licence et Master MIASHS. Elle effectue en parallèle des recherches au LIRMM portant sur les méthodes de sciences des données appliquées au domaine de la santé.

Florian Lombardo, enseignant contractuel à l'Université Paul Valéry Montpellier III, effectue ses recherches au LEIRIS (Laboratoire d'Études Interdisciplinaires sur le Réel et les Imaginaires Sociaux) dans le cadre d'une thèse en sociologie.

Dans un premier temps, nos commanditaires nous ont partagé quelques chiffres sur le suicide, un problème majeur de santé publique. Selon l'OMS (Organisation Mondiale de la Santé), une personne met fin à ses jours toutes les 40 secondes. Cela représente plus de 800 000 décès par an et au moins 20 fois plus de tentatives de suicides (1). Les taux de suicide ont augmenté globalement de 60% dans le monde sur les 45 dernières années. Prédominant chez les personnes âgées jusqu'à maintenant, il l'est aujourd'hui chez les personnes plus jeunes. C'est la deuxième cause de décès chez les 15-29 ans. De plus, un simple contact humain, réel ou virtuel, peut suffire pour éviter un suicide.

Malheureusement, de nos jours, il n'y a pas assez de médiatisation ou de démarche publique portant sur le suicide. Rien ne nous permet de prévoir à l'avance si quelqu'un va se suicider, alors qu'un simple contact humain, réel ou virtuel, peut suffire pour éviter un suicide. L'importance de la prédiction des comportements suicidaires nous a alors semblé évidente. Autant pour la société, quand l'OMS souhaite réduire d'un tiers les décès par suicide d'ici 2030, que pour nos commanditaires, quand cette prédiction peut s'appuyer sur les méthodes de sciences des données.

PERSPECTIVES

L'intérêt de ce sujet est donc de pouvoir identifier, par une analyse textuelle, les personnes étant les plus à même de faire une tentative de suicide. Cela permettra de sensibiliser concrètement et rapidement les personnes concernées. Le suicide étant lié à de nombreux facteurs sociaux (exclusion, repli sur soi, problème personnel ou professionnel, dépression ...), il est donc intéressant de comprendre les états des personnes qui font des tentatives de suicide (TS).

Il nous est demandé d'élaborer un algorithme de prédiction des risques suicidaires à partir de différents corpus de textes :

- nature (biographie Wikipédia, tweet, publication Reddit)
- taille (un tweet est limité à 254 caractères)
- narration (contrainte ou libre)

L'objectif est d'apprendre des modèles supervisés où les étiquettes des données sont connues (Wikipédia, nous savons que les personnes se sont effectivement suicidées) et d'étudier la possibilité de transférer ces modèles vers des données Twitter ou Reddit où la nature des textes est différente.

Dans un premier temps, il nous faut comprendre les comportements suicidaires. Cela passe par des recherches sur le sujet et des lectures d'articles scientifiques. Des documents nous ont été remis sur les travaux déjà réalisés afin que nous puissions nous imprégner au mieux de notre travail.

Les données utilisées sont des données textuelles issues de biographies Wikipédia de personnes qui se sont suicidées, de publications sur le forum Reddit de personnes qui ont survécu à une tentative de suicide mais également de tweets qui sont associés à la notion de suicide.

Une analyse textuelle de ses données nous permet de connaître les éventuels points de convergences (sexe, âge, secteur d'activité, enfance, relation familiale, situation professionnelle, ...) entre les personnes qui se suicident ou tentent de se suicider.

La finalité du projet est donc d'élaborer un outil qui nous permettra de repérer les textes en lien avec le suicide, à partir de données textuelles. Les messages ou publications pourront ainsi être classés comme à risque.

Dans une perspective à très long termes, les auteurs de messages considérés à risque pourront être contactés par des groupes de soutien dans le but d'agir avant la tentative de suicide. Les groupes de soutien pourront être composés de professionnels de santé ou de modérateur selon la stratégie des médias sociaux.

Nos compétences acquises au cours de nos cursus respectifs, mais également les notions que nous allons aborder au cours de l'année nous servent pour mener ce projet à bien.

BIBLIOGRAPHIE

(1) Communiqué de presse de l'OMS, le 09/09/2019 :

<https://www.who.int/fr/news/item/09-09-2019-suicide-one-person-dies-every-40-seconds>