

# Développement d'un algorithme de détection de tweets à connotation suicidaire

## Nice: un homme suicidaire sauvé par une mobilisation d'internautes sur Twitter

Un homme a annoncé sa volonté d'en finir avec la vie sur Twitter. La mobilisation des internautes a permis de le trouver et de le transporter aux urgences.

En 2017, un homme envoyait sur tweeter un message annonçant sa volonté d'en finir avec la vie. Cependant, son message n'est pas resté sans réponse. Les utilisateurs se sont mobilisés pour relayer l'information et ont dans la finalité permis de sauver la vie de cet homme.

Nous nous sommes intéressés à cette problématique et avons décidé de mettre nos connaissances en traitement des données et Machine Learning pour y répondre.

Notre objectif est de mettre en place un algorithme de scraping de Twitter afin de récupérer un ensemble de tweet comportant des messages à connotation suicidaire à l'aide d'un algorithme d'analyse des sentiments. Nous souhaitons mettre en place un Bot qui enverra un message de prévention aux personnes concernées avec un ensemble d'informations et de solutions pouvant leur apporter de l'aide.

Ce document présente l'ensemble des recherches effectuées sur le suicide et les pensées suicidaire afin de comprendre la démarche que nous devront suivre dans la détection des tweet. Nous présentons également l'ensemble des aides disponibles qui nous serviront dans la dernière étape de notre projet: informer les personnes.

## I. Définition du suicide

Dans un premier temps, il faudra s'intéresser au suicide en lui-même, ce que c'est, comment il se remarque.

Se suicider, c'est se donner intentionnellement la mort. Il ne s'agit pas d'une maladie et des signes peuvent être détectés en amont sur un très grand nombre des personnes souhaitant commettre un tel acte.

En France, 1 décès sur 50 est lié au suicide, soit environ un total de 160 000 tentatives entraînant la mort à 10 000 d'entre eux. Il est donc primordial de détecter les potentiels messages d'alertes afin de pouvoir leur proposer une aide adaptée.

La dépression présente un large éventail de symptômes et affecte des millions de personnes chaque année. Les personnes touchées manifestent des signes de tristesse, perdent de l'intérêt pour leurs activités, n'ont plus d'appétit, ont un sommeil irrégulier, ne parviennent plus à réfléchir et ont éventuellement des pensées suicidaires.

Notre objectif est de détecter ces signes via les messages que les personnes ayant de telles pensées peuvent poster sur les réseaux sociaux et les informer sur les aides et solutions disponibles.

Une situation suicidaire est décrite par trois étapes :

- Les idées suicidaires
- La crise suicidaire : envahissement de l'esprit par un projet suicidaire
- Le passage à l'acte

*Nous ne parlerons pas dans ce rapport de l'étape du passage à l'acte et de son déroulement car notre objectif est de prévenir ces situations et donc de comprendre les étapes amenant à cette dernière.*

Sur sa page d'aide, Twitter présente un ensemble de questions qu'il est possible de se poser pour interpréter une publication en ligne grâce à des signes avant-coureurs:

- *Cette personne laisse-t-elle fréquemment transparaître un sentiment de dépression ou de désespoir dans ses publications ?*
- *Publie-t-elle des commentaires morbides ? Évoque-t-elle la mort de manière non équivoque ?*
- *Publie-t-elle des commentaires évoquant des tentatives de suicide passées ?*
- *Décrit-elle un comportement autodestructeur ou publie-t-elle des photos qui en témoignent, ou se présente-t-elle comme suicidaire ?*
- *Avez-vous récemment identifié un changement de son humeur ou du contenu de ses publications ?*

## II. Les situations à risque

Les statistiques montrent, une fréquence beaucoup plus élevée de femmes que d'hommes qui commettent des tentatives de suicide, par contre, les décès par suicide touchent plus d'hommes que de femmes.

Les décès par suicide seraient favorisés par l'âge, un habitat rural (notamment dans le nord et l'ouest de la France), et le faible niveau socio-économique.

Certains événements de vie peuvent générer du mal-être. Dans ces moments-là, il est important de soutenir la personne et d'être attentif aux signaux d'alerte :

- Le décès d'un proche
- Une séparation
- Des relations familiales difficiles
- Un divorce
- Une situation de discrimination, de harcèlement, de violences physiques ou verbales (envers soi ou envers un proche)
- Le chômage et/ou la pauvreté
- La difficulté de reconnaissance ou d'acceptation de l'orientation sexuelle
- Un suicide chez un membre du réseau immédiat de la personne
- La maltraitance dans l'enfance
- Les problèmes de santé mentale ou physique
- Les Jeunes de 15 à 24 ans et les personnes âgées de 65 ans et plus
- Personnes incarcérées
- Personnes ayant vécu un ou plusieurs épisodes de blessures auto-infligées

D'autres facteurs sont également existant mais touchent plutôt certaines catégories de population.

À l'adolescence :

- L'échec scolaire
- Des conflits soudains dans la famille

À l'âge adulte :

- L'échec scolaire
- La perte d'un emploi
- Des difficultés financières
- Un échec professionnel

Chez les personnes âgées

- Le deuil du conjoint
- La perte du permis de conduire
- Le sentiment de ne plus être utile à la société
- La perte de l'autonomie
- La maladie chronique

La plupart des personnes qui se suicident ont envoyé des messages de détresse et de souffrance à leur entourage, et ont manifesté leur intention d'en finir. Il est donc crucial de reconnaître les indices et les comportements suicidaires regroupés en plusieurs catégories.

### **Messages verbaux directs et indirects**

La personne parle de son intention de se tuer clairement ou avec ambiguïté, de mort, d'écoeurement et d'essoufflement : « Je veux me suicider », « J'ai fait mon testament »

### **Indices comportementaux**

La personne commence par effectuer un retrait et isolement par rapport à ses amies, famille, ou encore collègues pour faire une mise en ordre de ses affaires en rédigeant des lettres, régler ses conflits et rédiger son testament. Ses habitudes vont également changer notamment sur son hygiène, son sommeil ou encore ses habitudes alimentaires. Il peut également développer un intérêt soudain pour les armes à feu, le suicide, les choses morbides, la réincarnation ou les cimetières.

### **Indices cognitifs et émotionnels**

Il se peut également que la personne développe des difficultés de concentration, un manque soudain d'intérêt et de motivation et d'indécision que cela soit dans le domaine privé ou professionnel. Elle aura également tendance à changer plus facilement et brusquement d'humeur et montrera des signes d'agressivité et de tristesse

### **Symptômes de dépression**

Il y a dépression si, depuis plus de deux semaines, il y a présence d'au moins cinq des symptômes suivants et qu'ils représentent un changement important par rapport au fonctionnement habituel :

- Humeur dépressive ;
- Diminution de l'intérêt et du plaisir ;
- Trouble de l'alimentation
- Les abus de substances
- Insomnie ou hypersomnie ;
- Fatigue, perte d'énergie ;
- Sentiment de dévalorisation ou de culpabilité ;
- Difficulté à se concentrer ou à prendre des décisions ;
- Pensées sur la mort ou le fait de mourir.

### III. Les Réseaux sociaux et le suicide

Aujourd'hui, les réseaux sociaux sont souvent présents dans le processus dépression, signes précurseurs et suicide en lui-même. Twitter compte pas moins de douze millions d'utilisateurs en France, dont 22% de 18-24 ans, couplé à une addiction des réseaux sociaux par les plus jeunes, cette utilisation intensive peut être néfaste pour les personnes les plus fragiles:

- Une exposition possible à des contenus évoquant le suicide:

De nombreux contenus évoquent le suicide, la dépression et l'automutilation, qui peuvent influencer négativement. De plus, le fonctionnement des algorithmes des réseaux sociaux, qui ont tendance à proposer aux utilisateurs du contenu similaire avec ce qu'ils ont vu précédemment, peut enfermer les utilisateurs dans cette pensée négative.

- L'isolement sur les réseaux sociaux:

Exister et se sentir important, sont des choses importantes pour les adolescents les plus fragiles. La mise en scène du suicide exposée aux yeux de tous sur les réseaux sociaux peut lui paraître héroïque.

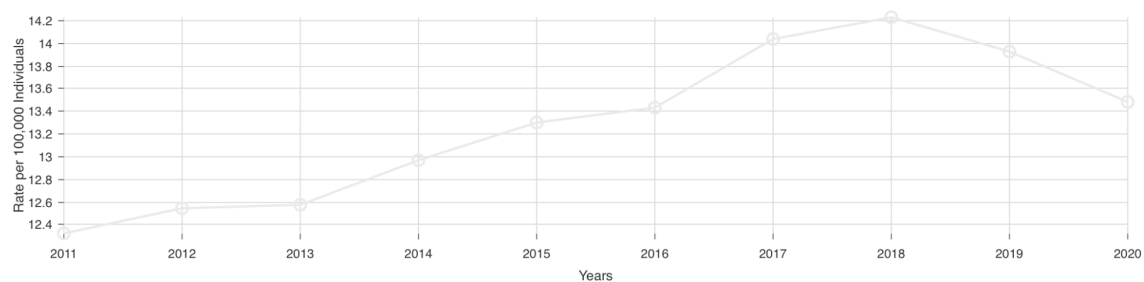
- Les messages ou tweets dans notre cas, indiquant une potentielle dépression:

Comme vu avant, des signes avant coureurs peuvent orienter sur la piste du suicide, tels que des messages sur les réseaux sociaux. En effet, la majorité des victimes de suicide partagent leurs pensées à propos de la mort ou du suicide avant de mourir. C'est sur ce point que devra se concentrer notre travail de détection.

D'après une étude universitaire américaine menée par une équipe de psychologues, sur un échantillon de plus de 500 000 personnes en 2017:

Les adolescents utilisant leurs smartphones plus de cinq heures par jour ont 66% de risque supplémentaire de ressentir des symptômes suicidaires que ceux qui ne consacrent qu'une heure quotidienne aux écrans.

On peut constater que l'explosion des appareils connectés et de l'usage d'internet et en particulier les réseaux sociaux, augmente les risques de pensées suicidaires, en particulier chez les plus jeunes.



ÉVOLUTION DES SUICIDES AUX ÉTATS-UNIS

## IV. Les aides possibles

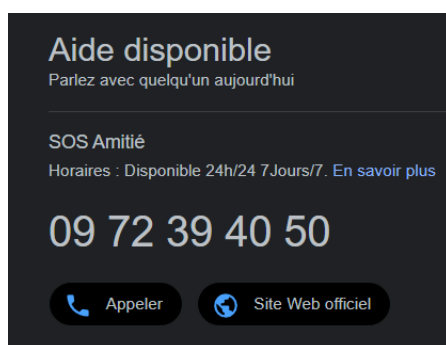
- Sur twitter:

Si nous sommes témoins d'un tweet autodestructeur, il est possible de le signaler comme tel à Twitter.

*“Quand nous recevons un signalement relatif à un risque de suicide ou de comportement autodestructeur, nous contactons la personne concernée et l'informons que quelqu'un qui tient à elle nous a alertés. Nous l'encourageons également à rechercher de l'aide et lui donnons des informations sur des ressources dédiées en ligne et par hotline.” - help.twitter.com*

- Sur Google:

Lorsque une recherche internet présente un lien avec le suicide, un message d'aide s'affiche automatiquement en haut des résultats:



- Le gouvernement français:

Sur le site [solidarite-sante.gouv](https://solidarite-sante.gouv.fr), un ensemble d'aides est présenté:

- Le numéro national de prévention du suicide: le 3114. Un professionnel de soins (infirmier ou psychologue), spécifiquement formé à la prévention du suicide, est à l'écoute afin d'évaluer la situation et proposer des ressources adaptées aux besoins de la personne.
- En cas de risque suicidaire imminent, le SAMU (15) ou le 112 (numéro européen)
- SOS Amitié: 09 72 39 40 50
- Fil Santé Jeunes: 0 800 235 236
- Suicide écoute: 01 45 39 40 00
- SOS Suicide Phénix: 01 40 44 46 45
- Phare Enfants-Parents: 01 43 46 00 62

L'ensemble des aides est disponible [ici](#).

## Sources

- <https://suicideactionmontreal.org/suicide-signes-de-detresse/>
- <https://centredecrisebsl.qc.ca/signes-precursurs-du-suicide/>
- <https://slate.fr/story/153873/reseaux-sociaux-adolescents-suicide>
- <https://preventionsuicide.be/fr/je-cherche-des-infos/signaux-d-alerte.html>
- [https://has-sante.fr/jcms/c\\_271964/fr/la-crise-suicidaire-reconnaitre-et-prendre-en-charge](https://has-sante.fr/jcms/c_271964/fr/la-crise-suicidaire-reconnaitre-et-prendre-en-charge)
- <https://ledauphine.com/magazine-sante/2021/09/09/suicide-les-reseaux-sociaux-facteur-de-ri-sque>
- [https://bfmtv.com/tech/addiction-aux-reseaux-sociaux-instagram-et-snapchat-vises-par-une-plainte-apres-le-suicide-d-un-adolescent\\_AN-202204130429.html](https://bfmtv.com/tech/addiction-aux-reseaux-sociaux-instagram-et-snapchat-vises-par-une-plainte-apres-le-suicide-d-un-adolescent_AN-202204130429.html)
- <https://solidarites-sante.gouv.fr/prevention-en-sante/sante-mentale/la-prevention-du-suicide/article/que-faire-et-a-qui-s-adresser-face-a-une-crise-suicidaire>
- [https://lexpress.fr/actualite/societe/fait-divers/nice-un-homme-suicidaire-sauve-par-une-mobilisation-d-internautes-sur-twitter\\_1901355.html](https://lexpress.fr/actualite/societe/fait-divers/nice-un-homme-suicidaire-sauve-par-une-mobilisation-d-internautes-sur-twitter_1901355.html)
- <https://help.twitter.com/fr/safety-and-security/self-harm-and-suicide>

## La suite du projet

Notre projet suivra 3 étapes principales:

1. Scraper twitter sur une période à l'aide de mots clés
  - a. Déterminer les mots clés qui nous permettront de sélectionner les tweet
  - b. Définir la période d'étude pour l'entraînement de notre modèle
2. Mise en place d'un modèle d'analyse des sentiments pour sélectionner uniquement les tweet montrant réellement des pensées suicidaire
  - a. Etude des modèles d'analyse des sentiments existants
  - b. Sélection et amélioration d'un modèle
3. Création d'un bot sur Twitter afin de contacter la personne et lui donner les aides disponibles



# **Développement d'un algorithme de détection de tweets à connotation suicidaire**

**PROJET PAS - LIVRABLE 02**

16/12/2022

# SOMMAIRE

- I. Revue des objectifs**
- II. Santé du projet**
- III. Réalisations actuelles**
  - A. Questionnaire
  - B. Avancement du code
    - 1. Scraping twitter
    - 2. Nettoyage des données
    - 3. Traduction
    - 4. Modèle de classification des tweets
- IV. Bibliographie**

## I. Revue des objectifs

Notre objectif est d'identifier les personnes évoquant des pensées suicidaires sur Twitter, afin de pouvoir contacter ces dernières avec une liste de l'ensemble des aides disponibles.

Pour ce faire, nous voulons mettre en place un bot twitter, qui se chargerait de tout le processus de manière automatisé, à travers plusieurs étapes:

- Scrapping des tweets, en récupérant le contenu des messages, le pseudo de la personne et la date d'émission. Afin de déjà procéder à un premier filtrage, le scrapping est réalisé à partir de mots clés tels que: suicide, mourir...
- Nettoyage des données en ne gardant que les tweets anglais (pour l'instant) et en enlevant tous les caractères qui ne seront pas utilisés lors de l'analyse des tweets: suppression des tags avec les @ et suppression des émojis, que nous n'utilisons pas pour l'instant.
- Analyse des tweets à l'aide du modèle **suicidal-bert** trouvé sur Hugging face. Le modèle nous donne un label et sa probabilité pour chaque tweet, indiquant si le message correspond à un message de suicide ou non.
- Récupération des résultats par notre bot, et mise en place du processus de contact des personnes concernées via un message privé listant l'ensemble des aides et associations disponibles à ce sujet.

Tous les modèles d'analyse de sentiments disponibles sont aujourd'hui entraînés sur des messages en anglais. On ne traite donc pour l'instant que les tweets anglais, avant d'étudier la possibilité de traduire les tweets d'une langue à l'anglais afin de faire notre analyse.

Le modèle **suicidal-bert** nous indique pour un texte le label suicide ou non avec la probabilité de sa prédiction. On va donc devoir choisir un seuil de probabilité à partir duquel on identifiera la personne comme étant à risques.

Faut-il la jouer large et donner un seuil faible à notre bot afin de s'assurer de ne rater personne au risque de contacter une personne inutilement, ou au contraire faut-il avoir un seuil élevé afin d'aider les personnes qui en ont réellement besoin au risque de manquer des personnes potentiellement fragiles?

La réponse à cette problématique sera essentielle dans la suite de notre projet et dans la crédibilité de notre solution apportée. On pourra envisager une phase de tests sur des tweets témoins pour déterminer ce seuil.

En imaginant qu'on aille au bout de notre idée et qu'on déploie réellement notre bot, l'objectif serait de transmettre les contacts des personnes fragiles à des associations, afin qu'une aide leur soit apportée.

## II. Santé du projet

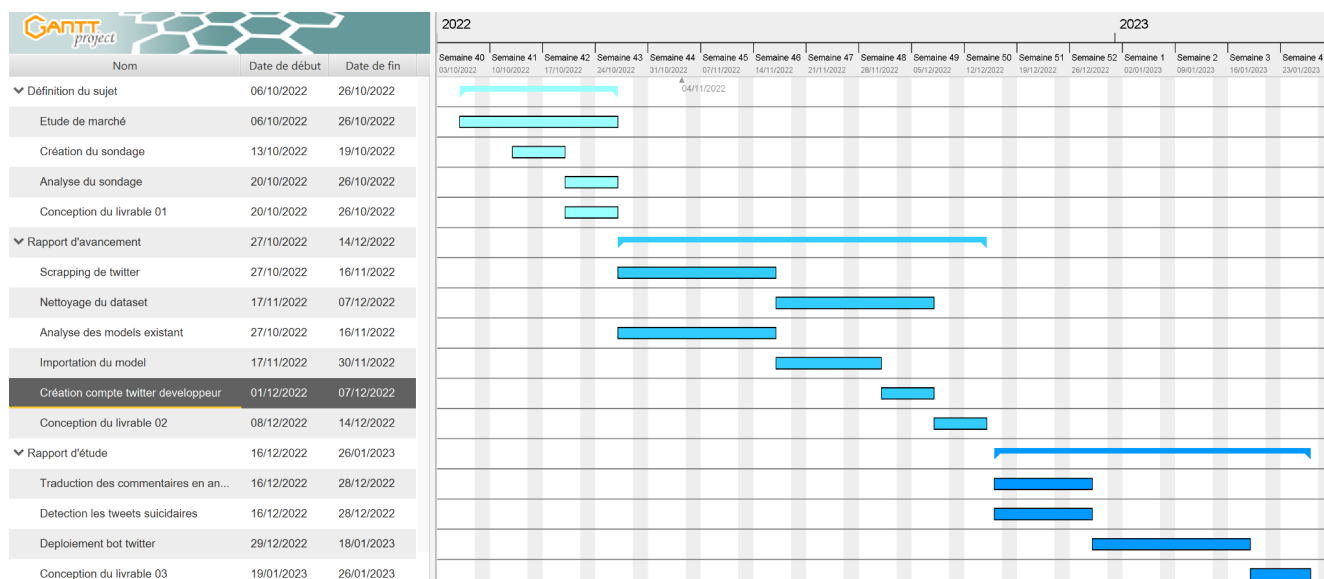


Diagramme de Gantt

Voici le diagramme de Gantt de notre projet. Il est divisé en 3 grandes parties distinctes sur une période de 4 mois. Nous sommes actuellement à la fin de la deuxième grande partie. Entre la première étape et maintenant, nous avons dû réaliser 5 tâches que nous avons réparties au sein de l'équipe.

Si nous suivons notre planning, nous sommes actuellement dans les temps. Chacune des tâches du rapport d'avancement a été finalisée et les résultats obtenus semblent excellents.

Un point de vigilance est cependant existant sur la deuxième partie de ce projet. Notre modèle nommé Suicidal-bert est cependant encore loin d'être parfait et peut commettre des erreurs. En fonction de ses potentielles erreurs et de leur gravité, il faudra probablement coupler celui-ci avec d'autres types de modèles afin d'en augmenter la fiabilité. Un modèle permettant de prédire les émotions ou la subjectivité pourrait être intéressant afin de confirmer la prédiction de notre modèle principal.

Des points de vigilance doivent être également mis en place pour les tâches de la troisième partie. En effet, nous avons déjà commencé à regarder la traduction des tweets en anglais, car seule celle-ci est acceptée par le modèle de classification que nous utilisons. Cependant, toutes nos tentatives de traduction se sont soldées par un échec. En effet, soit les bibliothèques ne fonctionnent plus, soit elles ne supportent pas le nombre élevé de données que nous avons à traduire. Nous avons donc pris, pour le moment, la décision de traiter uniquement les tweets anglais en attendant de trouver une solution. Un autre point de vigilance concerne le déploiement de notre bot twitter, car il s'agit d'une solution technique qu'aucun membre de l'équipe n'a eu à traiter dans sa carrière et qui semble s'être compliqué récemment suite au rachat de la compagnie. Notre solution à ces points de vigilance sera la

communication grâce à la méthode agile et l'utilisation du "pair-programming" pour les parties difficiles.

### **III. Réalisations actuelles**

#### **1) Questionnaire**

Afin d'évaluer approximativement la nécessité de notre projet, nous avons décidé d'élaborer un questionnaire d'une dizaine de questions que nous avons diffusé sur nos réseaux sociaux. Grâce à cela nous avons récolté plus d'une centaine de feedback d'étudiants qui nous ont permis d'en tirer quelques chiffres. Voici la liste des questions transmises :

- Tu es homme/femme/autre ?
- Quel âge as-tu ?
- Utilises-tu Twitter ?
- T'es tu déjà senti dans un état dépressif de par ton utilisation de Twitter? (Contenu photos, vidéos, messages négatifs)
- As-tu déjà fait une dépression ?
- Si oui, voulais-tu de l'aide extérieure ?
- Connais-tu les aides possibles pour les personnes ayant des pensées suicidaires ?
- As-tu déjà eu des pensées suicidaires ?
- As-tu déjà tweeté ou retweeté un contenu laissant penser que tu étais mal au point de mettre fin à tes jours ?
- Si oui, quelqu'un t'a-t-il aidé suite à cela ?
- As-tu déjà vu des tweets te laissant penser que quelqu'un voulait mettre fin à ses jours ?
- Si oui, as-tu essayé de l'aider ?

#### **Analyse du questionnaire (129 répondants) :**

Le questionnaire est à peu près équitable au niveau des sexes, 54% d'hommes 46% de femmes ont répondu. 65% des étudiants utilisent Twitter et 28% l'utilisent tous les jours.

10% de ces étudiants déclarent avoir déjà eu un sentiment dépressif à l'usage et la moitié de ces étudiants disent avoir fait une dépression.

Les personnes ayant fait une dépression sont assez équilibrées sur la question de l'aide extérieure. En effet, 50,6% ont effectivement demandé de l'aide extérieure et on été content de pouvoir être aidé dans ces moments là.

Cependant, seuls 18% des étudiants interrogés connaissent les aides possibles et envisageables auxquelles nous pouvons avoir recours dans ces situations.

Selon l'étude, 36% ont déjà eu des pensées suicidaires et 6% sont déjà allés jusqu'à tweeter cette émotion.

Nous sommes convaincus aujourd'hui qu'à l'aide de notre projet, nous pouvons aider cette population, ces 6% d'étudiants qui sont peut être seuls ou désinformer sur les aides et les moyens qui sont à leur portée pour être secouru.

Pour conclure cette analyse, le chiffre clé de cette étude est 93%. C'est le pourcentage d'étudiants ayant eu recours à ces méthodes n'ayant reçu aucune aide.... 93% des étudiants songeant au suicide ou tweetant des propos d'autodestruction ne sont pas accompagnés.

## 2) Avancement du code

### a) Scraping twitter

La première étape du développement est la récupération des tweets. Nous avons décidé d'utiliser une librairie permettant le scraping de Twitter. Cette dernière est *snsrape*.

*snsrape* est un scraper pour les services de réseaux sociaux (SNS: Social Network Services). Il scrape des éléments comme les profils d'utilisateurs, les hashtags ou les recherches, et renvoie les éléments découverts.

Les services suivants sont actuellement supportés :

- **Facebook** : profils d'utilisateurs, groupes et communautés (alias les publications des visiteurs).
- **Instagram** : profils d'utilisateurs, hashtags et emplacements
- **Mastodon** : profils d'utilisateurs et toots (simple ou fil de discussion)
- **Reddit** : utilisateurs, subreddits et recherches (via Pushshift)
- **Telegram** : canaux
- **Twitter** : utilisateurs, profils d'utilisateurs, hashtags, recherches, tweets (simples ou fils de discussion), listes de messages et tendances.
- **Vkontakte** : profils d'utilisateurs
- **Weibo** (Sina Weibo) : profils d'utilisateurs

Ainsi, il nous est possible grâce à cette librairie de récupérer un ensemble de tweet, mais aussi le nom des utilisateurs ayant écrit ce tweet. Cette information nous sera nécessaire pour la partie Bot dont l'objectif est de contacter l'utilisateur dont le tweet a été détecté comme suicidaire afin de l'aider.

En mai 2022, chaque seconde, en moyenne, environ 10 033 tweets sont postés, soit 867 millions de tweets envoyés par jour. Ainsi, nous ne pouvons pas récupérer l'ensemble des tweets. Nous avons donc décidé de récupérer un dataset par jour avec un filtre sur des mots/phrases clés.

Ainsi, nous avons créé un dictionnaire de mots/phrases nous permettant de filtrer les tweet que nous récupérons.

```
keywords=["suicide", "suicidaire", "mourir", "je veux mourir", "à l'aide", "aidez-moi", "mutiler", "pendre", "corde"]
```

*NB: cette liste n'est pas exhaustive et pourra être complétée au fil du projet.*

Le code complet permettant le scraping des tweets est le suivant:

## Extract from twitter

```
: def extract_tweet(keyword): #format keyword = keyword since:AAAA-MM-JJ until:AAAA-MM-JJ
    #Creating list to append tweet data
    tweet_list=[]

    for i, tweet in enumerate(sntwitter.TwitterSearchScrapper(keyword).get_items()):
        tweet_list.append([str(tweet.date), tweet.id, tweet.content, tweet.user.username])

    #Creating a dataframe from tweets list
    tweet_df=pd.DataFrame(tweet_list, columns=['Datetime', 'Tweet Id', 'Text', 'Username'])

    return tweet_df

: keywords=["suicide", "suicidaire", "mourir", "je veux mourir", "à l'aide", "aidez-moi", "mutiler", "pendre", "corde"]

: debut = "2022-12-14"
  fin = "2022-12-15"

df=pd.DataFrame(columns=['Datetime', 'Tweet Id', 'Text', 'Username'])

for kw in keywords:
    tweet = extract_tweet(f'{kw} since:{debut} until:{fin}')

    frames=[df, tweet]

    df = pd.concat(frames)

df = df.reset_index(drop=True)
```

... et permet d'obtenir le dataset suivant pour les données du 14/12/2022 :

	Datetime	Tweet Id	Text	Username
0	2022-12-14 23:59:54+00:00	1603177994498998272	Canadian Healthcare System Introduces Punch Ca...	DisobbedireOra
1	2022-12-14 23:59:50+00:00	1603177979462287360	You never know what's inside people's mind. It...	RJ_2_SF
2	2022-12-14 23:59:50+00:00	1603177979131334656	@edwardeldridge @hodgetwins Good for you. The...	DjAnnonOmys
3	2022-12-14 23:59:47+00:00	1603177968418795520	@bradgoodson0 @CyrusMMcQueen I knew the nicest...	SchmattCheryl
4	2022-12-14 23:59:46+00:00	1603177963406532609	@gladyscl @LencsikeD @TheEllenShow @official_t...	Lisa93070546
...	...	...	...	...
34773	2022-12-14 00:05:50+00:00	1602817100434784256	@LcorderoOrueta A contar ovejas corde	alvarocarrea
34774	2022-12-14 00:05:05+00:00	1602816911330492416	@ldebacher Faudrait prouver qu'un pouf et une ...	choupete450
34775	2022-12-14 00:04:54+00:00	1602816866195763200	@KCKEPS @vic_T_Wine @imTheRealDSTROY @kidlonni...	LifeGoodMeGood
34776	2022-12-14 00:03:21+00:00	1602816474674073600	tengo sueño y quiero dormirme, POR QUE COJONES...	LcorderoOrueta
34777	2022-12-14 00:02:40+00:00	1602816304469217280	@BrunoLeMaire Mec, vous êtes déjà incapable de...	LeoKar15

34778 rows × 4 columns



## b) Nettoyage des données

Une fois le dataset récupéré, il est nécessaire de préparer ce dernier avant d'appliquer le modèle de classification. En effet, les messages extraits de Twitter peuvent contenir des emoji ainsi que des username. Nous devons donc les enlever des tweets:

### Clean dataset

- remove emojis
- remove usernames

```
def Clean(text):
    regex_pattern = re.compile(pattern = "["
        u"\U0001F600-\U0001F64F" # emoticons
        u"\U0001F300-\U0001F5FF" # symbols & pictographs
        u"\U0001F680-\U0001F6FF" # transport & map symbols
        u"\U0001F1E0-\U0001F1FF" # flags (iOS)
        "]+", flags = re.UNICODE)
    temp = regex_pattern.sub(r'',text) # remove emojis
    temp = re.sub('@[\s]+', '',temp) # remove usernames (@...)
    temp = temp.strip()
    return temp
```

```
pandarallel.initialize(progress_bar=True)
```

```
df.Text = df.Text.parallel_apply(Clean)
```

INFO: Pandarallel will run on 8 workers.

INFO: Pandarallel will use standard multiprocessing data transfer (pipe) to transfer data between the main process and workers.

Cette étape a été réalisée en utilisant les expressions régulières afin de détecter les patterns des emoji et des username.

L'application de la fonction sur le dataset a été faite à l'aide de la méthode *parallel\_apply* de la librairie *pandarallel*, qui permet d'appliquer des fonctions à un dataframe pandas à l'aide d'une parallélisation pour aller plus vite.

	Datetime	Tweet Id	Text	Username
0	2022-12-14 23:59:54+00:00	1603177994498998272	Canadian Healthcare System Introduces Punch Ca...	DisobbedireOra
1	2022-12-14 23:59:50+00:00	1603177979462287360	You never know what's inside people's mind. It...	RJ_2_SF
2	2022-12-14 23:59:50+00:00	1603177979131334656	Good for you. The reason you never got sick w...	DJAnnonOmys
3	2022-12-14 23:59:47+00:00	1603177968418795520	I knew the nicest, funniest, best dad who comm...	SchmattCheryl
4	2022-12-14 23:59:46+00:00	1603177963406532609	He committed suicide	Lisa93070546
...	...	...	...	...
34773	2022-12-14 00:05:50+00:00	1602817100434784256	A contar ovejas corde	alvarocarraa
34774	2022-12-14 00:05:05+00:00	1602816911330492416	Faudrait prouver qu'un pouf et une corde=suici...	choupete450
34775	2022-12-14 00:04:54+00:00	1602816866195763200	Au pire tu lui achètes une corde pour se pendr...	LifeGoodMeGood
34776	2022-12-14 00:03:21+00:00	1602816474674073600	tengo sueño y quiero dormirme, POR QUE COJONES...	LcorderoOrueta
34777	2022-12-14 00:02:40+00:00	1602816304469217280	Mec, vous êtes déjà incapable de produire de l...	LeoKar15

34778 rows x 4 columns

### Dataset après nettoyage des tweets

### c) Traduction

Le modèle de classification que nous souhaitons utiliser a été entraîné sur un jeu de données en anglais. Cependant, notre dataset comporte des tweet dans d'autres langues que l'anglais et il nous est donc nécessaire de traduire les données en anglais.

Après quelques recherches, nous avons souhaité utiliser l'API de Google Translate afin de traduire l'ensemble des tweets.

```
def Translate(text):
    translator = Translator()
    translator.raise_Exception = True
    translated = translator.translate(text, dest = 'en').text
    return translated
```

Cependant, cette méthode n'a pas été concluante, car bien que l'API soit annoncée comme gratuite et illimitée, nous avons été bloqués suite à un trop grand nombre de traductions demandé.

#### Googletrans: Free and Unlimited Google translate API for Python

license MIT build failing docs passing pypi package 3.0.0 coverage 91% maintainability D

Googletrans is a **free** and **unlimited** python library that implemented Google Translate API. This uses the [Google Translate Ajax API](#) to make calls to such methods as detect and translate.

#### API annoncée comme gratuite et illimitée

**Exception:** Unexpected status code "429" from ('translate.google.com',)

#### Message d'erreur 429

Nous nous sommes donc tournés vers une autre librairie de traduction également annoncée comme gratuite et illimitée: *deep-translator*.

Cette dernière fonctionne sans problème. Cependant, le temps de calcul sur l'ensemble du dataset est très long:

```
for i in tqdm(range(len(df1))):  
    try:  
        df1.Text[i] = Translate(df1.Text[i])  
    except:  
        print(i)  
        df1.drop([i], axis=0, inplace=True)
```

100%  10000/10000 [44:45<00:00, 6.77it/s]

### Test sur un sous-dataset

Ainsi, le test nous montre que le modèle traduit environ 222 tweets par minute ( $1000/45$ ). Or, notre dataset contient 34778 tweets, ce qui revient à 2 heures et 36 minutes de temps de traitement.

#### d) Modèle de classification des tweets

Plusieurs solutions se sont présentées à nous pour le choix du modèle. La première était de le réaliser nous-mêmes, mais trouver un dataset contenant des données d'entraînement semblait s'avérer compliqué. De plus, il existe maintenant un grand nombre de modèles très performants et surtout déjà entraînés (ce qui pourrait prendre plusieurs mois pour obtenir le même résultat avec notre propre modèle en fonction de la complexité ou de la taille des données d'entrée). Il nous semblait donc naturel de chercher un modèle déjà prêt et, après quelques recherches, nous sommes tombés sur un modèle nommé **suicidal-bert**. Celui-ci a été conçu pour déterminer si un message semble montrer des envies suicidaires ou non. Le modèle possède en sortie 2 possibilités :

- Label 0: Le modèle pense qu'il ne s'agit pas d'un message suicidaire
- Label 1: Le modèle pense qu'il s'agit d'un modèle suicidaire

```
from transformers import pipeline
classif = pipeline(model="goohjy/suicidal-bert")
classif("Mario kart is the best game ever")

[{'label': 'LABEL_0', 'score': 0.9999561309814453}]
```

#### Exemple de prédiction d'un message non suicidaire

```
from transformers import pipeline
classif = pipeline(model="goohjy/suicidal-bert")
classif("SUICIDAL")

[{'label': 'LABEL_1', 'score': 0.9982531666755676}]
```

#### Exemple de prédiction d'un message suicidaire

## IV. Bibliographie

### 1) Twitter scraping

- <https://github.com/JustAnotherArchivist/snsrape>
- <https://www.renolon.com/number-of-tweets-per-day/>

### 2) Traduction

- <https://py-googletrans.readthedocs.io/en/latest/>
- <https://support.google.com/translate/thread/15027713/when-i-am-calling-google-trans-late-api-multiple-times-in-my-program-it-is-showing-error-why-is-it-so?hl=en>
- <https://pypi.org/project/deep-translator/>

### 3) Models

- <https://huggingface.co/gooohjy/suicidal-bert>
- [https://huggingface.co/docs/transformers/main\\_classes/pipelines](https://huggingface.co/docs/transformers/main_classes/pipelines)
- [https://huggingface.co/transformers/v3.0.2/model\\_doc/auto.html](https://huggingface.co/transformers/v3.0.2/model_doc/auto.html)

### 4) Bot Twitter

- <https://www.sales-hacking.com/post/creer-bot-twitter>