

FAKE NEWS

Recherche & Développement

AUGER Nathan / DURAND Pierre
LOPEZ Julio / NOUVELIERE Benjamin

MAY Madeth

Qu'est-ce qu'une fake new ?



Politique



Santé



BBC NEWS

Partout ??

Le cyber-harcèlement très répandu

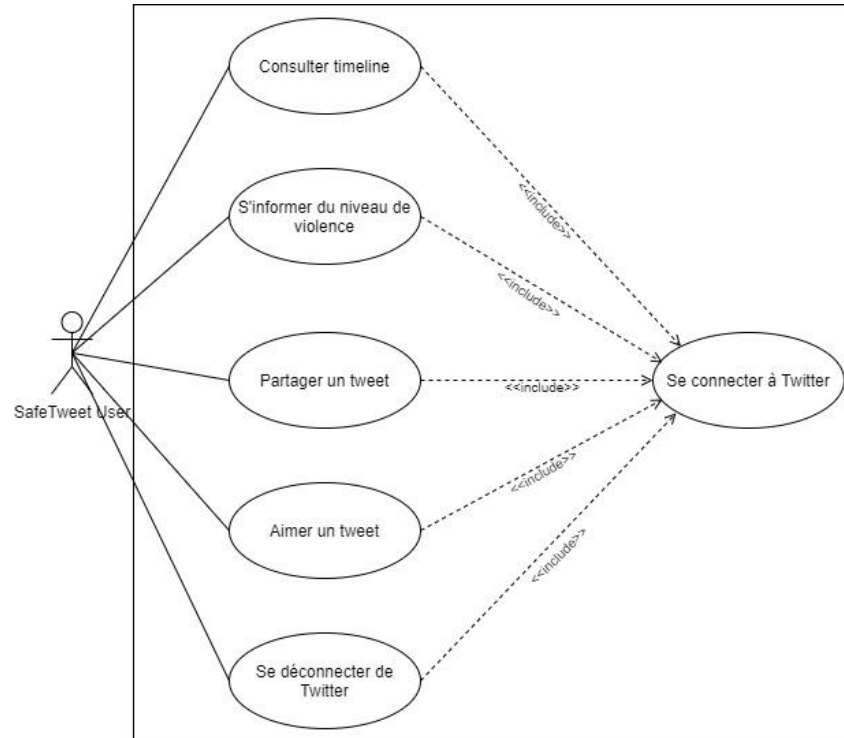
En France en **2019**, **22% des jeunes majeurs** disent avoir déjà été de cyber-harcèlement.
Ce chiffre monte à **27% pour les 18-24 ans**.



Suis-je en **danger** face au cyber-harcèlement ?
Suis-je **protégé** contre la diffamation en groupe ?



Les usages de l'utilisateur



Dans une optique de R&D (Benchmarks)

Benchmark des Frameworks

	<u>Angular</u>	<u>React</u>	Vue
Taille	65.5 <u>KB</u>	36.3 <u>KB</u>	30.8 <u>KB</u>
Temps de performance	1.5 s	1 s	1 s
Temps de changement	1484.6 ms	796.5 ms	806.5 ms
Apprentissage	Facile	Facile	Très facile

Technologies	Spécificités
Régression linéaire	- Facile et rapide à calculer - Réajustement manuel de la courbe
Réseau de neurones	- Réajustement automatique de la courbe

Benchmark des bibliothèques d'apprentissage automatique

<u>TensorFlow</u>	<ul style="list-style-type: none"> - Bien documenté - Grande communauté - <u>Graph</u> statique - <u>Debugging</u> non intuitif - Bonne visualisation de <u>graph</u> - Déploiement très facile - Date de 2011 développé par Google
<u>Pytorch</u>	<ul style="list-style-type: none"> - Documenté - Tutoriels officiels - <u>Graph</u> dynamique - <u>Debugging</u> intuitif - Bonne visualisation de <u>graph</u> - Facile à utiliser (moins d'efforts pour autant de résultats) - Date de 2016 développé par Facebook

Langages	Spécificités
Python	<ul style="list-style-type: none"> - Large panel de bibliothèques (<u>Numpy</u>, <u>Panda</u>, <u>NLTK</u>, <u>SpaCy</u>, <u>SciKit-learn</u>) - Intègre facilement des <u>APIs</u> - Performant en temps - Communauté développée et active
<u>JVM</u>	<ul style="list-style-type: none"> - Bon panel de bibliothèques (<u>CoreNPL</u>, <u>ND4J</u>, <u>DL4J</u>) - Calcule sur CPU - Machine virtuelle - Communauté développée
C/C++	<ul style="list-style-type: none"> - Quelques bibliothèques (<u>CUDA</u>, <u>TensorFlow</u>, <u>Caffe</u>) - Utilisation de pointeurs - Calcule sur <u>GPU</u> - Accède à des <u>APIs</u> de haut niveau et flexibles
JavaScript	<ul style="list-style-type: none"> - Très peu de bibliothèques - Intègre mal les <u>APIs</u>

Technologies utilisées pour les modèles



Simple d'utilisation et possède beaucoup de bibliothèques (NLP, ML, DL)



Bibliothèque qui permet de faire de l'apprentissage automatique (Naive Bayes, SVM)



Bibliothèque de traitement automatique des langues (Lemmatisation)



Bibliothèque qui permet de faire du deep learning (modèle LSTM)

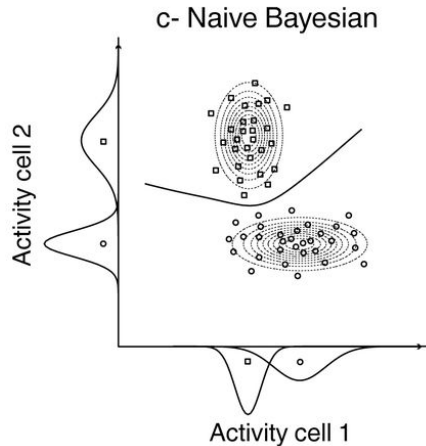


Flask

Framework pour application web sur python (utilisé pour l'API)

Les différents modèles

- Naive Bayes



Avantages :

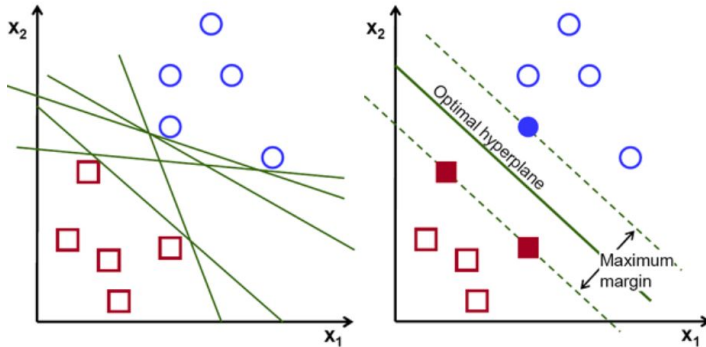
- Simple à comprendre et à mettre en place
- Rapide pour l'entraînement et les prédictions

Inconvénients :

- Fait une prédiction binaire
- Ne fonctionne pas très bien sur des données complexes

Les différents modèles

- SVM



Avantages :

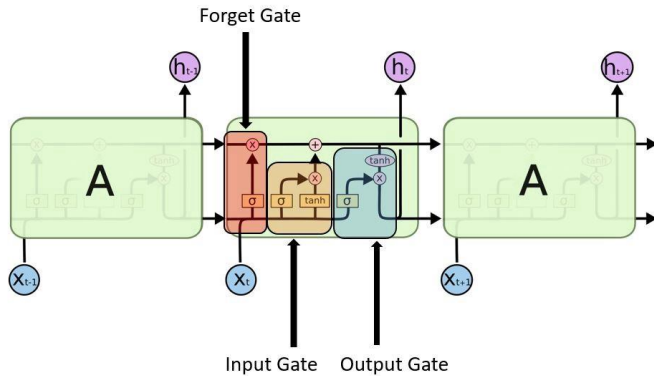
- Résistant au sur-apprentissage
- Performant

Inconvénients :

- Fait une prédiction binaire
- Ne convient pas à des jeux de données volumineux

Les différents modèles

- LSTM



Avantages :

- La prédiction est une probabilité
- Un réseau de neurones récurrent amélioré
- Très utilisé dans le domaine du traitement de la langue

Inconvénient :

- La phase d'apprentissage est plutôt lente

Étapes clés du programme

1. **Nettoyer** les données (mettre en minuscule, enlever la ponctuation, enlever StopWord, Lemmatiser...)
2. **Trier** les mots par fréquence
3. Enlever les **'outliers'**
4. **Remplir / Tronquer** les données
5. Séparer les données en **3 groupes** (Train : 80%, Validation : 10%, Test : 10%)
6. Définir la structure du **réseau LSTM** (taille des batch, nombre d'époque...)
7. **Entraînement** du modèle
8. **Tester** (données de Test, phrase utilisateur)

Démonstration du modèle

- Quelques exemples de tweets classifiés



Bonjour, je suis heureux aujourd'hui : 0.0025246990844607353

$x < 0.7$



Merci beaucoup a tous pour votre soutien : 0.2498244047164917



T'etonne pas si je t'attend a la sortie du college : 0.7055900692939758

$0.7 \geq x > 0.9$



Les violeurs comme toi merite d'aller en taule fils de pute : 0.9998519420623779

$x \geq 0.9$

Le choix d'une **application mobile**



Twitter

- Base de données conséquente
- Peu de filtrage de contenu

+



Application android

- Plateforme mobile plus adaptée aux réseaux sociaux

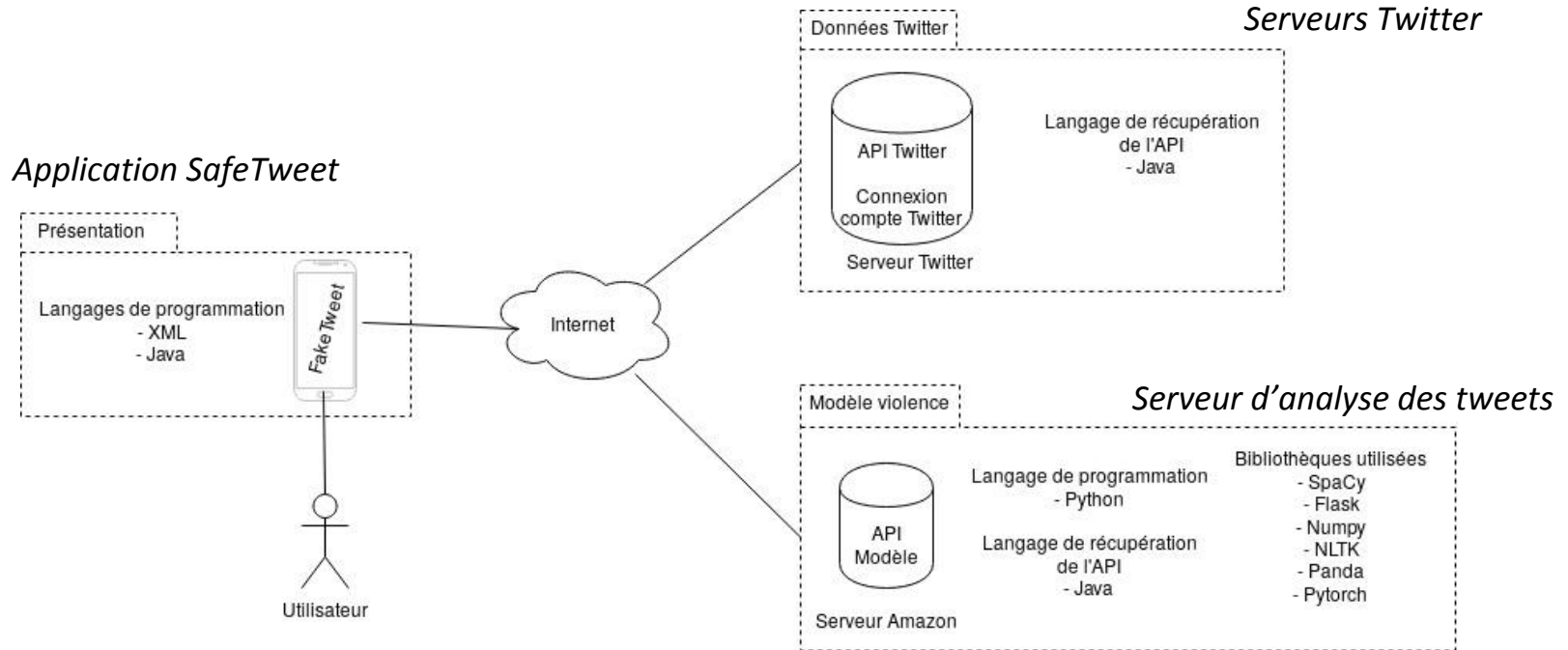
+

Filtrage

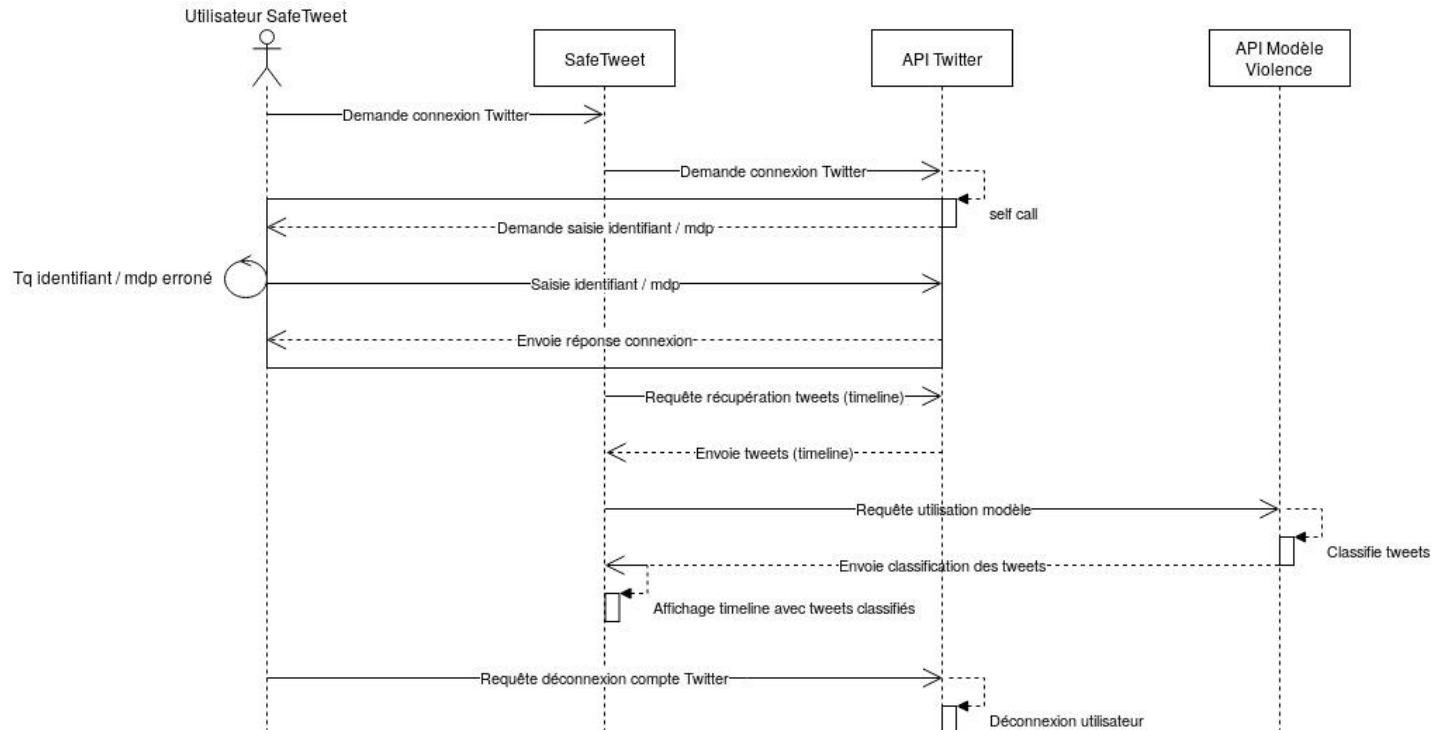
Modification apportée

- Notre plus-value
- Protéger l'utilisateur de contenus potentiellement dangereux

Structure de l'application en 3 axes



Cas d'une première connexion à S



Notre façon de travailler

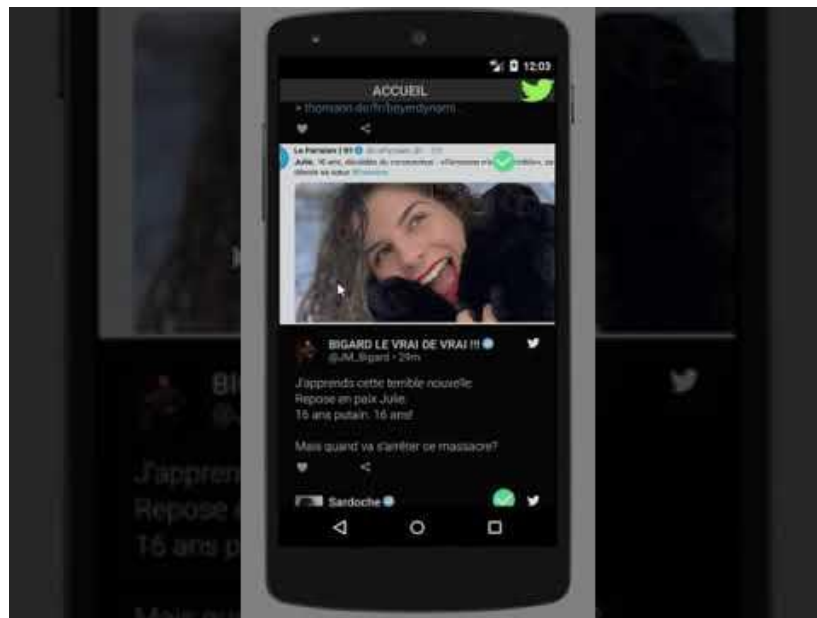


- Versionning



- Gestion de projet
- Méthode agile
- Rush 1 semaine
- 1 séance en fin de rush de mise en commun
- Rédaction du rapport de séance
- Réunion de projet toutes les 3 à 4 séances

Démonstration



Les différentes pistes d'amélioration

- Signaler un compte
- Démentir le résultat de l'analyse d'un tweet
- Informateurs de violence adaptés aux tweets avec images
- Interactions standards de Twitter (Commenter et retweeter)
- Améliorer le modèle (création de notre propre dataset)
- Améliorer le visuel de l'application

