



Présentation P4 - Projet fil rouge

Toxic Comment Classification & Identification

25/06/2020

Hamza AMRI, Camille COCHENER, Sophie LEVEUGLE, Rodolphe SIMONEAU



Plan de la présentation

1. Contexte et enjeux du projet
2. Méthodes et difficultés rencontrées
3. Solution proposée
4. Avantages et limites
5. Perspectives de développement

Contexte et enjeux du projet

 Featured Prediction Competition

Toxic Comment Classification Challenge

Identify and classify toxic online comments

\$35,000

Prize Money



Jigsaw/Conversation AI · 4,550 teams · 2 years ago

[Overview](#)

[Data](#)

[Notebooks](#)

[Discussion](#)

[Leaderboard](#)

[Rules](#)

[Team](#)

[My Submissions](#)

[Late Submission](#)

Overview

[Description](#)

[Evaluation](#)

Discussing things you care about can be difficult. The threat of abuse and harassment online means that many people stop

Toxic Comment Classification Challenge



"FUCK YOUR FILTHY MOTHER IN THE ASS, DRY!"

"Stupid piece of shit stop deleting my stuff
asshole go die and fall in a hole go to
hell!"

Toxic Comment Classification Challenge



Conversation AI



Jigsaw

Développement d'outils pour améliorer les conversations en lignes

Etude de commentaires toxiques



Beaucoup de modèles disponibles sur l'**API Perspective**

MAIS

**Les modèles font
encore des erreurs**

**Pas de sélection possible
du type de toxicité**

Toxic Comment Classification Challenge



Construire un **modèle multi-labels** qui permet de **détecter plusieurs types de toxicité...**

Menaces

obscénité

Insultes

Haine

à partir de **commentaires des pages de discussion de Wikipedia**



WIKIPÉDIA
L'Encyclopédie libre

Méthodes et difficultés rencontrées

L'équipe



Hamza
AMRI



Camille
COCHENER



Sophie
LEVEUGLE



Rodolphe
SIMONEAU

Support Académique

Geoffroy PEETERS
(Spécialiste des données audio)

Béatrice BIANCARDI
(Spécialiste des sciences cognitives)

Moyens

Sources de l'état de l'art



F. Chollet



Développement du code
pour les analyses



Librairies majoritaires
utilisées

NLTK



 Keras

Seaborn

Plateformes de calcul

colab

kaggle

Déploiement



plotly | Dash

Méthodologie

Etat de l'art

- Analyse de l'état de l'art et recherches des solutions de la classification multilabels en NLP
- Recherche des solutions de déploiement de modèles

Analyses exploratoires

- Analyses exploratoires des commentaires et des labels
- Exploration des méthodes de preprocessing en NLP
- Exploration des modélisations

Sélection de modèle

- Entraînement et validation de modèles sur un train set et un validation set
- Sélection de modèle
- Prédictions sur un test set non labellisé

Déploiement

- Sauvegarde du modèle
- Création d'une application
- Mise en production

Difficultés et ajustements réalisés



**Changement de
sujet au bout de 6
mois de projet**



Prise de décision
rapide sur un
nouveau sujet

**Travail à distance
dû au confinement**



Github,
visioconférence,
drive...

**Limites dans la
puissance de calcul**

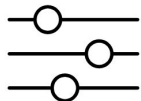


Utilisation
d'outils gratuits
(Colab, ...)

**Limites d'un
projet Kaggle**



Recherche de
complexité (création
d'une API...)



Solution proposée

Rappel de la démarche globale

Première phase

Explorations

Dataset
labellisé
issu de
Kaggle

Analyse exploratoire du
dataset labellisé

Explorations de
modélisations

Choix de construire
un modèle de Deep
Learning

Deuxième phase

Construction du modèle final et Prédictions

Construction
LSTM

Entraînement et
validation

AUC = 0.9804

Construction
CNN

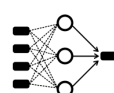
Entraînement et
validation

AUC = 0.9765

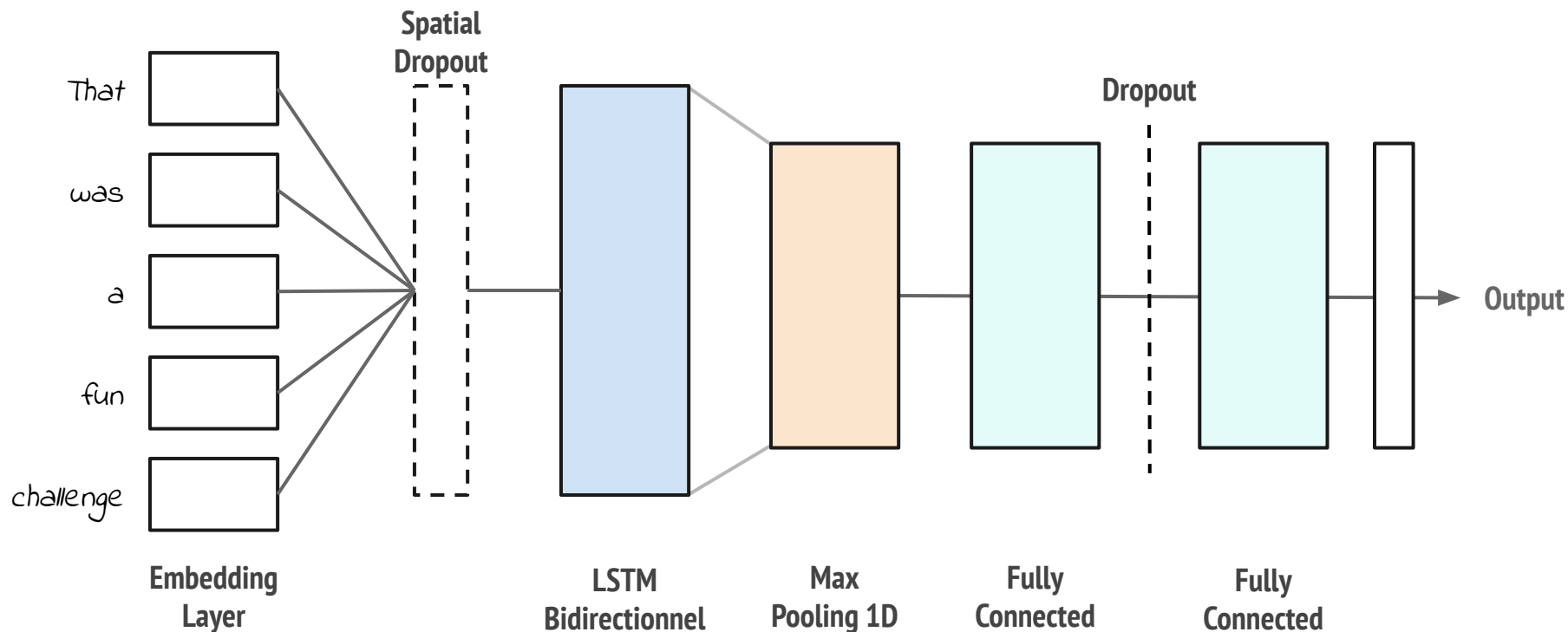
Sélection du modèle
final

Prédictions sur le test
set non-labellisé issu
de Kaggle

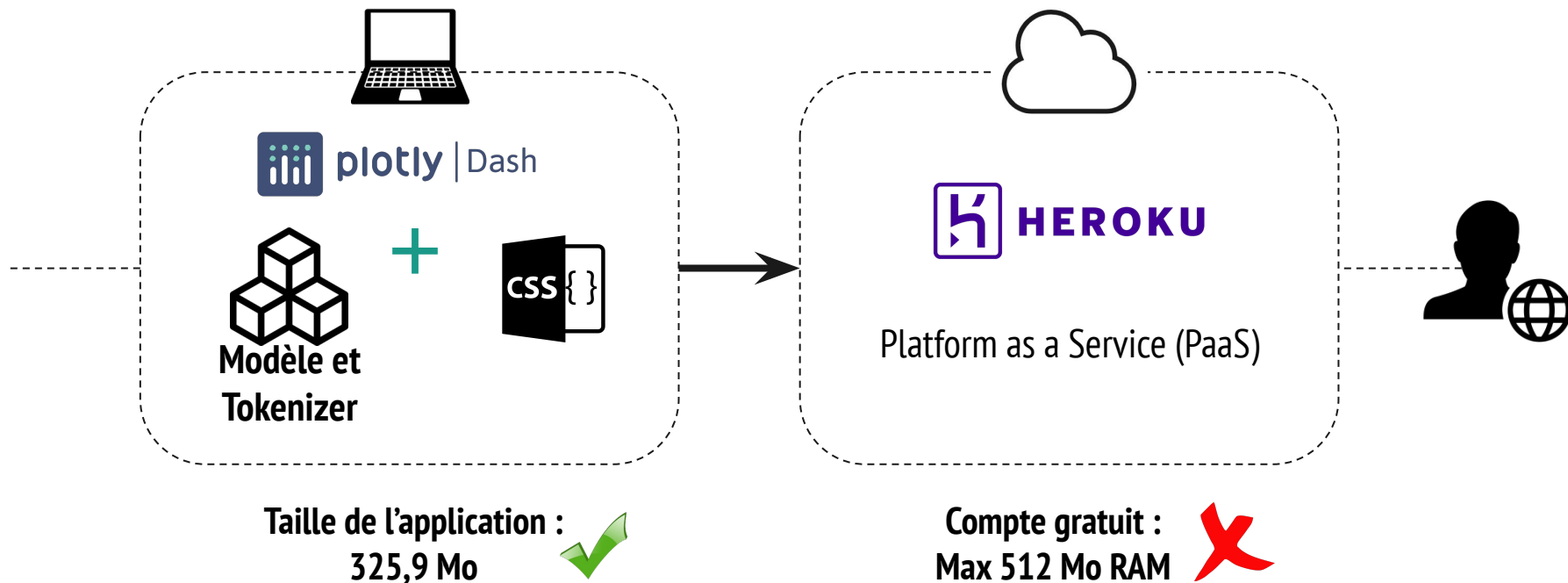
AUC = 0.9807



Architecture du réseau de neurones (LSTM)



Construction d'une application et déploiement



Une interface prédictive



About this project

This predictive tool was built as part of a student project during our Post Master degree in Big Data at Télécom Paris. The data used to build the tool are from the Kaggle "Toxic Comment Classification Challenge" organized by Jigsaw and Conversation AI.



About us



Camille COCHENER



Hamza AMRI



Sophie LEVEUGLE



Rodolphe SIMONEAU

Avantages et limites

Avantages et limites

Avantages

- ❖ Interface d'utilisation **simple d'utilisation** et **interactive**
- ❖ Outil **rapide**
- ❖ Solution alignée avec les enjeux et demandes initiales de Perspective
- ❖ Modèle parfaitement adapté au langage en-ligne

Limites

- ❖ A affiner sur les expressions utilisant des négations
- ❖ Entraîné sur une seule langue (anglais)
- ❖ Modèle volumineux et difficile à déployer en ligne

Sujets abordés au cours du challenge

- ❖ Construction d'architectures de réseaux de neurones (CNN, LSTM)
- ❖ Cas d'application concrète de NLP
- ❖ Gestion d'un projet Data de A à Z
- ❖ Travail d'équipe en conditions

Perspectives de développement

Axes d'améliorations

NLP



BERT

web interface

- ❖ Utilisation de cloud computing pour raccourcir et alléger les temps de calcul
- ❖ Utilisation de la version payante de Heroku pour obtenir un déploiement en ligne plus fluide

Merci pour votre attention !

Architecture du réseau de neurones (LSTM)

