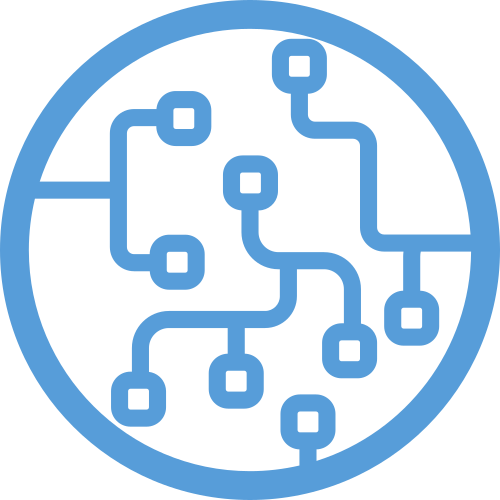
**Financial Sentiments**

****

***Projet P3***

***Capocasale Romain***

***INF3dlm-A***

***He-Arc, Suisse***

***2019***

Table des matières

[1) Cahier des charges 2](#_Toc21031002)

[2) Résumé 2](#_Toc21031003)

# 

# Résumé

# Introduction

Dans le cadre du 5ème semestre de Bachelor en développement logiciel et multimédia, nous avons pour objectif de réaliser un projet individuel qui permet de nous permet de nous préparer le mieux possible au travail de Bachelor. J’ai alors sélectionné un projet dans le domaine du machine learning.

Dans le cadre du machine learning, le domaine de l'analyse de sentiments dans les textes a pris beaucoup d'intérêt ces dernières années grâce au succès et aux bonnes performances des algorithmes de classification de texte. Dans un premier temps cette technologie à été utilisé pour analyser les sentiments dans des domaines tels que les commentaires de films, les hôtels, les restaurants, etc

Le but du projet est d’appliquer cette technologie d’analyse de sentiment au domaine du marché financier dans le but de dire si un texte à plutôt un sentiment positif ou négatif. Ce sentiment pourrait alors être utilisé dans le domaine du FOREX (Foreign Exchange Market). Une information perçue comme positive pourrait indiquer une montée d’une devise par rapport à une autre tandis qu’une information négative pourrait en indiquer la chute.

L’objectif du projet est de créer un analyseur de sentiment pour le domaine de la finance. Le modèle sera entrainé avec des textes et vocabulaires propres au thème de la finance. Dans un premier temps, les solutions d’analyse de sentiment présente sur le web seront analysé pour résoudre ce problème. Dans un second temps, le but est de créé un modèle de machine learning dans le but de prédire un sentiment par rapport à un texte financier.

# Cahier des charges

Le but du projet est de créer un modèle d'analyse de sentiment pour le domaine de la finance. Un analyseur de sentiment est un programme informatique qui en entrée demande une phrase ou un texte et en sortie indique un sentiment pour cette phrase sur une échelle de -1 à 1. -1 pour un sentiment négatif et 1 pour un sentiment positif.

Étant donné que le domaine de la finance dispose de son propre langage, il est intéressant de développer un modèle spécialisé dans ce secteur dans le but d'obtenir de meilleur résultat qu'un analyseur de sentiment global qui aurait été entraîné à faire des prédictions sur tout le vocabulaire de la langue anglaise.

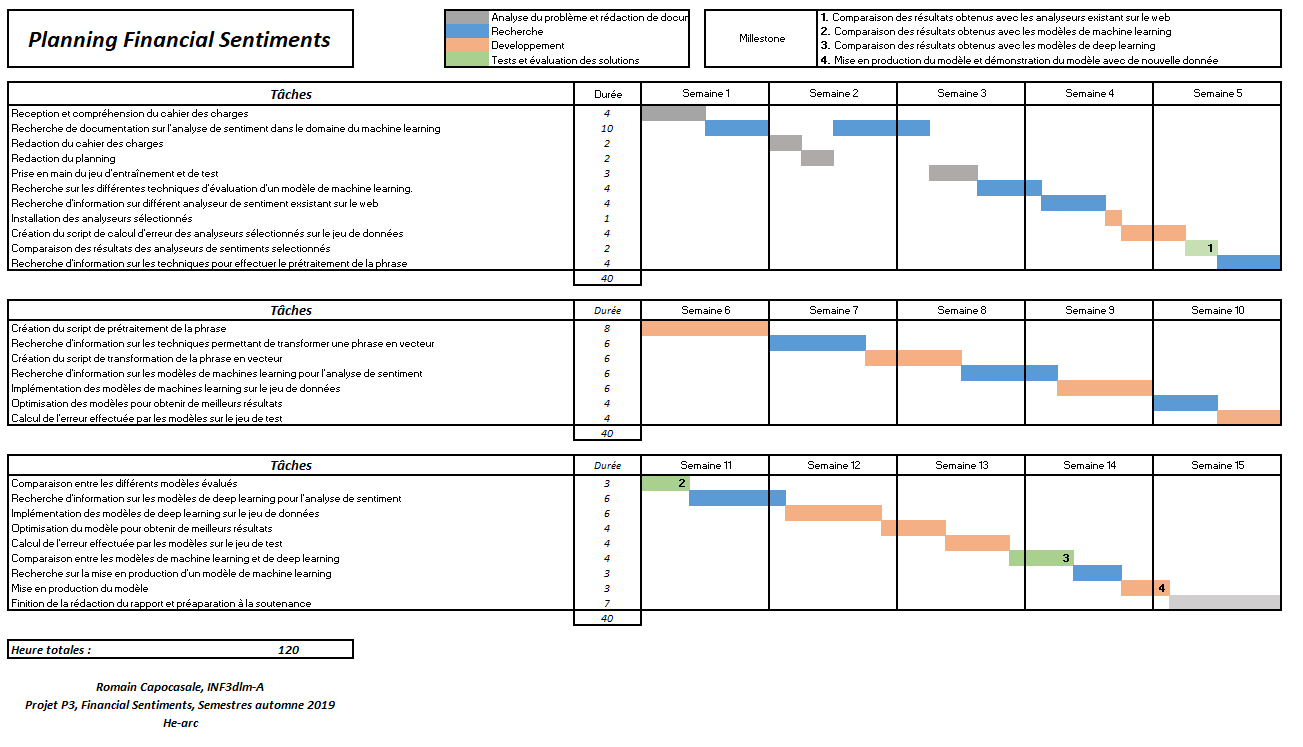
Pour l'entraînement et le test du modèle, un ensemble de titres d'article de finance avec la valeur du sentiment correspondant a été fourni. Le modèle sera réalisé en Python.

Pour ce projet, il faudra dans un premier temps, comparer le score des analyseurs de sentiment existant sur le web avec le jeu de donnée fourni. Par la suite, le but est de développer un modèle spécialisé dans l'analyse de sentiment pour la finance. La première étape consistera à prétraiter la phrase pour enlever la ponctuation, les déterminants, etc. Il faudra par la suite, transformer la phrase en vecteur dans le but d'être utilisé dans un modèle de machine learning. Différents modèles de machine learning devront être implémentés et testés à l'aide du jeu de donnée fourni. Par la suite, un modèle plus complexe utilisant des techniques de deep learning pourra être implémenté. Il faudra également comparer ces différents modèles entre eux pour sélectionner celui effectuant le moins d'erreurs de prédiction.

Le but final(hors projet), serait d'utiliser ce modèle d'analyse de sentiment dans le but de suivre la tendance haussière et baissière dans les pairs de devises (par exemple USD/CHF et EUR/USD). Des sources textuelles seraient récoltées depuis le site de Reuters et d'autres canaux financiers spécialisés dans le but d'en prédire un sentiment (entre -1 et 1) d'une période de trading (une heure, une demi-journée, journée, semaine ...) et relatif à un sujet (par exemple Brexit, échanges commerciaux avec la Chine, ....).

Ce classificateur avec d'autres en cours de développement, sera intégré dans un outil de prédiction du taux change.

# Planning



# Etat de l’art

Il existe de nombreux analyseur de sentiment sur le web comme par exemple : Textblob, VADER (nltk), etc. Ces analyseurs sont très simples à utiliser, il suffit d’indiquer la phrase et la librairie va indiquer la polarité de la phrase entre -1 et 1. Ces analyseurs ont cependant été entrainé sur tout le vocabulaire de langue anglaise et non sur le vocabulaire de la finance spécifiquement. Ils auront alors tendance à commettre plus d’erreur qu’un classificateur entrainé sur un vocabulaire spécifique. Un modèle entrainé avec un vocabulaire spécifique donnera une pondération plus importante aux mots qu’il considère plus important dans le contexte donné.

Si on regarde dans le domaine de la finance spécifiquement, d’après mes recherches ils n’existent pas de libraire ou d’analyseur de sentiment spécifique à la finance. Ils existent cependant de nombreux tutoriels qui expliquent comme créé un modèle à partir de texte contenant du vocabulaire financier. On trouve également pas mal d’article qui explique l’utilité du NLP (Natural Language Processing) appliqué au domaine de la finance.

Avec ces différentes recherches, on constate donc qu’il serait intéressant de créer un modèle de Machine Learning entrainé avec un vocabulaire spécifique à la finance.