

# **Analyse de sentiments à base d'apprentissage profond dans les forums de discussions MOOC**

Ahlem Mejri\*, Dr. Mariem Mahfoudh\*, \*\*, Dr. Nihel Kooli\*\*\*

\*Université de Kairouan, Tunis, mejriahlem25@gmail.com

\*\*Laboratoire MIRACL, Université de Sfax, Tunis, mariem.mahfoudh@gmail.com

\*\*\* Direction générale de l'armement de France, France, Nihel.Kooli@gmail.com

## **1 Introduction**

Selon (Liu, 2012), l'analyse des sentiments (AS) est "le domaine d'étude qui analyse les opinions, les sentiments, les évaluations, les attitudes et les émotions des gens à partir du langage écrit". Elle vise à analyser les avis pour suivre l'humeur du public à propos d'un produit ou d'un sujet particulier. Notre travail s'intéresse au domaine d'apprentissage en ligne, plus particulièrement à l'AS des avis publiés sur les cours en lignes ouvertes et massives (MOOC). Le MOOC est un type de formation en ligne qui offre aux étudiants/apprenants la possibilité de suivre une formation à distance. Au cours de la formation, les apprenants ont tendance à discuter et publier fréquemment leurs avis sur les forums de discussions, rendant ainsi le travail des instructeurs très difficile en matière d'analyse des avis. Des travaux de la littérature ont cherché à analyser automatiquement les polarités des avis (Nikolić et al., 2020). Cependant, ils se sont concentré sur l'identification des sentiments selon trois classes : négatifs, neutres et positifs.

Nous utilisons des techniques de TAL avec une approche basée sur l'apprentissage profond (CNN) pour déterminer leurs polarités émotionnelles. Les avis sont classés en fortement positif, positif, fortement négatif, négatif, ou neutre.

## **2 Méthode proposée**

Notre approche d'AS sur les MOOC comprend cinq étapes principales :

1. Collection des données : consiste à collecter les commentaires des apprenants de site MOOC Coursera. Ils sont 140320 avis.
2. Pré-traitement des données : consiste à appliquer certaines techniques de TAL, comme la tokenisation, le stemming, la suppression des mots vides sur les avis afin de bien les préparer à la phase de l'analyse.
3. Extraction des caractéristiques : utilise le TF/IDF et le word embedding pour transformer les avis en représentations numériques.