

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
**ECOLE SUPERIEURE EN INFORMATIQUE 08 MAI 1945 SIDI BEL  
ABBES**



## Mémoire

En vue de l'obtention du diplôme d'ingénieur

Filière : Informatique

Spécialité : Système d'information et Web (SIW)

## Thème

---

**Conception et réalisation d'une application Web pour la  
détection des faux profils dans les réseaux sociaux en  
utilisant le machine learning.**

---

**Présenté par :**

AMIRI Fatima Zohra

**Soutenu le :** 06/07/2022

Devant le jury composé de:

|                      |  |                         |
|----------------------|--|-------------------------|
| <b>Président:</b>    | <i>M<sup>r</sup></i> BELFEDHAL Alaa Eddine | Maître de conférences B |
| <b>Examineur:</b>    | <i>M<sup>me</sup></i> NAOUM Hanae          | Maître assistante A     |
| <b>Encadrant:</b>    | <i>M<sup>r</sup></i> Nadir MAHAMMED        | Maître de conférences A |
| <b>Co-encadrant:</b> | <i>M<sup>me</sup></i> Badia KLOUCHE        | Maître de conférences B |

Année universitaire 2021/2022

# Résumé

Les réseaux sociaux occupent une grande place dans nos vies. Les gens les utilisent pour socialiser et s'informer. Les avantages sont nombreux, mais les inconvénients également par exemple les faux profils qui se multiplient rapidement et diffusent de fausses nouvelles et informations. Ils jouent un rôle important dans les menaces persistantes avancées et sont également impliqués dans d'autres activités malveillantes.

Ce travail examine différents algorithmes d'apprentissage automatique et la façon dont ils aident à résoudre les problèmes liés à la détection de faux profils sociaux, dans un premier temps. Par la suite, une proposition sous la forme d'une application Web est présentée pour y remédier.

*Python* est utilisé comme langage de programmation, avec l'environnement *Jupyter Notebook* avec diverses bibliothèques d'analyse de données et de *machine learning* comme *Pandas*, *Sklearn*, *Numpy*, etc. Des algorithmes d'apprentissage automatique (supervisés et non-supervisés), à savoir les machines à vecteurs de support (SVM), Random Forest, Naive Bayes, KNN, decision tree et K-means clustering avec différentes versions, sont utilisés dans ce travail. Pour aller plus loin dans la proposition de résolution du problème de détection des faux profils sur les réseaux sociaux, nous avons proposé d'une hybridation entre les algorithmes bio-inspirés et les algorithmes d'apprentissage non-supervisés. Les résultats obtenus sont encourageants. Ils ont été comparés sur la base de différentes métriques tirées

de l'état de l'art de l'utilisation du *machine learning*.

**Mots clés :** Web 2.0, réseau social, faux profil, apprentissage automatique, algorithmes d'apprentissage supervisés, algorithmes d'apprentissage non-supervisés, algorithme d'optimisation par essaim.

# Abstract

Social networks are a big part of our lives. People use them to socialize and get information. The advantages are many, but so are the disadvantages, such as fake profiles that multiply rapidly and spread fake news and information. They play an important role in advanced persistent threats and are also involved in other malicious activities.

This work examines different machine learning algorithms and how they help solve the problems related to fake social profile detection, initially. Subsequently, a proposal in the form of a web application is presented to address this. *Python* is used as a programming language, with the *Jupyter Notebook* environment with various data analysis and machine learning libraries such as *Pandas*, *Sklearn*, *Numpy*, etc. Machine learning algorithms (supervised and unsupervised), namely Support Vector Machines (SVM), Random Forest, Naive Bayes, KNN, decision tree and K-means clustering with different versions, are used in this work. To go further in the proposal of solving the problem of false profile detection on social networks, we proposed a hybridization between bio-inspired algorithms and unsupervised learning algorithms. The results obtained are encouraging. They have been compared on the basis of different metrics from the state of the art of machine learning.

**Keywords** : Machine learning, Web 2.0, Social networking, Fake profiles, Supervised

learning algorithms, Unsupervised learning algorithms, Particle Swarm Optimization Algorithm