

الجمهورية الشعبية الديمقراطية الجزائرية  
République Algérienne Démocratique et Populaire  
وزارة التعليم العالي والبحث العلمي  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

المدرسة العليا للإعلام الآلي - 08 ماي 1945 - بسبدي بلعباس  
Ecole Supérieure en Informatique  
-08 Mai 1945- Sidi Bel Abbès



## MEMOIRE

En vue de l'obtention du diplôme de **Master**

Filière : **Informatique**

Spécialité : **Systeme d'Information et Web (SIW)**

## Thème

---

Fouille de web basée sur les méthodes d'apprentissage automatique et  
d'apprentissage profond

---

Présenté par :

- Mr Brahim FELOUH
- Mlle Fadia Nouna BOUSBA

Soutenu le : **26/06/2019**

Devant le jury composé de :

- |                          |           |
|--------------------------|-----------|
| - Mr Nabil KESKES        | Président |
| - Mme Rim Sakina BENNABI | Encadreur |
| - Mme Nadia ELOUALI      | Examineur |

Année Universitaire : 2018 / 2019

## Résumé

L'apprentissage en profondeur est une technique d'apprentissage automatique qui vise à extraire des représentations de haut niveau à partir de données brutes en utilisant plusieurs couches neuronales. Le succès récent de l'apprentissage en profondeur a montré qu'il surpassait les systèmes conventionnels en traitement d'image. La fouille du Web utilise les algorithmes de fouille de données pour détecter les modèles à partir de la toile. Il est divisé en trois axes : fouille du contenu du web, fouille de l'usage et fouille de la structure du web. L'usage de l'approche d'apprentissage profond peut offrir un potentiel considérable pour améliorer la fouille du Web, et ce en exploitant les grandes masses de données disponibles sur la toile d'une manière efficace. L'utilisation des méthodes conventionnelles en fouille de données sur ces grands volumes de données se trouve souvent limitée, n'exploitant pas au mieux ces bases d'information. Ceci nécessite donc un apprentissage en profondeur afin d'extraire de la connaissance nouvelle et non évidente. Dans ce travail, on dresse un état de l'art des travaux existants dans l'usage des algorithmes d'apprentissage automatique et apprentissage profond dans l'exploration du Web. Ceci permet de comparer ces travaux et de les synthétiser pour avoir une vision globale sur cet axe de recherche.

**Mots clés** : la fouille de données, l'apprentissage profond, l'exploration de web.

## Abstract

Deep learning is an automatic learning technique that aims to extract high-level representations from raw data using multiple neural layers. The recent success of deep learning has shown that it outperforms conventional image processing systems. Web search uses data mining algorithms to detect models from the web. It is divided into three areas : searching the content of the web, searching the usage and searching the structure of the web. The use of the deep learning approach can offer considerable potential for improving Web browsing by leveraging the vast amount of data available on the web in an efficient manner. The use of conventional methods in data mining on these large volumes of data is often limited, not making the best use of these databases. This requires deep learning to extract new and unseen knowledge. In this work, we draw up a state of the art of existing works in the use of machine learning algorithms and deep learning in the exploration of the Web. This allows to compare these works and to synthesize them to have a global vision on this line of research.

**Keywords** : Data mining, deep learning, web mining.