



République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université des Sciences et de la Technologie Houari Boumediène

Faculté d'Informatique

Département d'Intelligence Artificielle et Sciences des Données

Mémoire de Licence

Filière : Informatique

Spécialité : Informatique Académique Générale

Thème

Evolution du profil social par apprentissage automatique

Encadré par : Mme BOULKRINAT Nour El Houda

Organisme : Centre de Recherche sur l'Information Scientifique et Technique

Soutenu le : 06/06/2023

Réalisé par :

Devant le jury composé de :

SAADA Samir

Mme BENSAOU Nacera **Présidente**

Mr GOT Adel **Membre**

N° Projet : ACAD_E_011

Année Universitaire : 2022/2023

Dédicaces

Je souhaite tout d'abord dédier ce travail à Dieu, qui m'a donné la détermination et la force de continuer à travailler malgré les difficultés rencontrées tout au long de mon parcours.

À mes chers parents, Ils m'ont donné la vie, élevé avec amour et dévouement, soutenu dans toutes les étapes de ma vie et m'ont toujours encouragé à poursuivre mes rêves. Je ne serais pas l'homme que je suis aujourd'hui sans eux et je leur suis infiniment reconnaissant pour tout ce qu'ils ont fait pour moi. Leur amour inconditionnel, leur soutien constant et leur confiance en moi ont été les pierres angulaires de mon épanouissement personnel.

À mon frère Nabil, qui a toujours été présent pour m'encourager et me conseiller.

À toute personne qui a eu un impact positif sur ma vie.

Je dédie ce travail.

Remerciements

Avant tous nous tenons à remercier ALLAH, le miséricordieux, de nous avoir donné la force, le courage, et la patience pour compléter ce modeste travail.

Nous commençons, par remercier et exprimer notre profonde gratitude envers notre encadrante, Mme. BOULKRINAT Nour El Houda pour ses judicieux conseils, sa disponibilité, sa bienveillance, son expertise et ses encouragements, nous ont permis d'avancer de manière efficace et de dépasser les difficultés rencontrées.

Nos vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils portent à notre travail, d'avoir accepté de l'examiner et d'avoir consacré leur temps pour l'évaluer.

Une grande reconnaissance à tous nos professeurs qui nous ont transmis leurs savoirs en toute conscience, et qui ont été à la hauteur avec leur honorable noble métier.

Enfin, nous adressons nos sincères remerciements à nos familles respectives, plus essentiellement à nos parents qui nous ont aidés, soutenus et encouragés tout au long de la réalisation de ce mémoire.

Résumé

Les réseaux sociaux sont devenus un élément incontournable de la vie quotidienne, offrant une plateforme pour l'interaction sociale, le partage d'informations et l'expression de soi. Les utilisateurs interagissent avec ces plateformes en fonction de leurs intérêts personnels, créant ainsi une multitude de données pour les analystes de données. Cependant, les utilisateurs ne sont pas toujours actifs sur les réseaux sociaux et leurs intérêts peuvent évoluer rapidement. Le développement des réseaux sociaux entraîne une croissance rapide du volume d'informations disponibles, rendant difficile pour l'utilisateur de trouver les informations pertinentes. Cependant, les réseaux sociaux collectent constamment des données sur les utilisateurs, offrant ainsi une opportunité d'évoluer leurs centres d'intérêts.

L'objectif de notre travail consiste à exploiter les données et les activités sociales de l'utilisateur en vue de l'évolution de ses centres d'intérêt afin de l'aider à trouver le contenu qu'il cherche, ainsi que d'enrichir son profil social. À cet effet, notre projet vise à modéliser l'évolution des centres d'intérêt des utilisateurs passifs et actifs en exploitant leurs postes du dataset Reddit et cela à l'aide des techniques d'apprentissage automatique, telles que les arbres de décision, la régression logistique, les réseaux de neurones CNN, RNN et ANN. Nous avons effectué une étude comparative entre ces techniques et notre système d'évolution fondé sur le réseau de neurones CNN, qui a fourni de meilleurs résultats.

Mots clés : Réseau social, Profil social, Evolution des intérêts, Réseau de neurone, CNN, Centre d'intérêt.

Abstract

Social networks have become an essential element of the daily live, offering a platform for social interaction, information sharing, and self-expression. Users interact with these platforms based on their personal interests, creating a multitude of data for data analysts. However, users are not always active on social networks, and their interests can evolve rapidly. The development of social networks leads to a rapid growth in the volume of available information, making it difficult for the user to find relevant information. However, social networks constantly collect data on users, offering an opportunity to evolve their interests.

The objective of our work is to leverage user data and social activities to evolve their interests in order to help them find the content they are looking for and enrich their social profile. To this end, our project aims to model the evolution of interests for both passive and active users by leveraging their posts from the Reddit dataset, using machine learning techniques such as decision trees, logistic regression, CNN, RNN, and ANN. We conducted a comparative study between these techniques and our evolution system based on the CNN neural network, which yielded better results.

Keywords : Social network, Social profile, Evolution of interests, Neural network, CNN, Interests.

Tables des matières

Introduction générale.....	1
Chapitre 1 : Réseaux Sociaux et Evolution Du Profil Social	
1. Introduction.....	3
2. Réseaux Sociaux	3
2.1 Définitions	3
2.2 Types des réseaux sociaux.....	3
2.3 Les composants d'un réseau social	4
2.3.1 Groupes.....	4
2.3.2 Tags	4
2.3.3 Nœuds.....	4
2.3.4 Liens	4
2.3.5 Graphe de contenu social.....	5
2.3.6 Les ressources	5
2.4 Type d'utilisateurs.....	5
3. Le profil utilisateur.....	5
3.1 Définition profil utilisateur	5
3.2 Représentation des profils utilisateurs	6
3.2.1 Représentation ensembliste.....	6
3.2.2 Représentation conceptuelle	6
3.2.3 Représentation multidimensionnelle	6
3.3 Détection des données d'un profil utilisateur	7
3.3.1 Détection explicite.....	7
3.3.2 Détection implicite	7
3.4 Evolution du profil utilisateur	7
3.4.1 Approche par sélection d'instance	7
3.4.2 Approche pondérée	8
3.4.3 Mise à jour du profil.....	8
4. Le profil social	8
4.1 Profilage social.....	8
4.2 Définition profil social.....	8
4.3 Détection des données utilisateurs	9
4.3.1 Producteurs de données	9
4.4 Evolution du profil social	9
4.4.1 Les tags.....	10

4.4.2	L'analyse de réseau social	10
4.4.3	Les techniques de l'apprentissage automatique	10
5.	Conclusion	11

Chapitre 2 : Les Techniques D'apprentissage Automatique

1.	Introduction.....	12
2.	Apprentissage automatique.....	12
2.1	Qu'est-ce que l'apprentissage automatique ?	12
2.2	Types d'apprentissage automatique	12
2.2.1	Apprentissage supervisé	12
2.2.2	Apprentissage non supervisé	14
3.	Apprentissage profond (Deep Learning)	14
4.	Réseaux de neurones	15
4.1	Qu'est-ce qu'un réseau de neurone artificiel ?.....	15
4.2	Les réseaux de neurones artificiels (ANN).....	15
4.3	Réseaux de neurones convolutifs (CNN)	16
4.4	Réseaux de neurones bouclés ou récurrents (RNN).....	17
4.5	Réseaux de neurones Long-Short Term Memory (LSTM)	17
5.	Réseaux de neurones et évolution du profil social.....	18
6.	Conclusion	20

Chapitre 3 : La Conception De La Solution

1.	Introduction.....	21
2.	Positionnement du problème	21
3.	Présentation du dataset	22
3.1	La plateforme reddit	22
3.2	Description du dataset	22
4.	Prétraitement des données	23
4.1	Label Encoding	23
4.2	Conversion des valeurs utc en datetime	23
4.3	Génération du nouveau dataset	24
5.	Modèle de deep learning.....	25
5.1	Choix du modèle	25
5.2	Features and labels	25
5.3	La standardisation	25
5.4	La division du dataset.....	25
5.5	Architecture du modèle	26
5.6	Compilation du modèle.....	27

5.7	Entrainement du modèle	27
5.8	Evaluation du modèle	29
6.	Comparaison des modèles	29
7.	Conclusion	30

Chapitre 4 : Implémentation et Mise en Œuvre

1.	Introduction.....	31
2.	Architecture du système	31
3.	Les diagrammes UML.....	32
3.1	Le diagramme de cas d'utilisation	32
3.1.1	L'identification des rôles	32
3.2	Le diagramme de séquences	33
3.3	Le diagramme de classes	34
4.	L'environnement de développement.....	35
4.1	L'environnement matériel.....	35
4.2	L'environnement logiciel.....	35
4.2.1	TensorFlow	35
4.2.2	Python.....	35
4.2.3	Django	35
4.2.4	SGBD SQLite	35
4.2.5	VS Code.....	36
4.2.6	Google Collaboratory	36
5.	Interface de l'application	36
5.1	Page d'accueil de l'application	36
5.1.1	Inscription	36
5.1.2	Authentification	37
5.2	Espace utilisateur.....	37
5.2.1	Consultation de profil	38
5.2.2	Créer un post	38
5.3	Espace administrateur.....	39
5.3.1	Visualiser la liste des utilisateurs	39
5.3.2	Suivre l'interaction de chaque utilisateur	39
5.3.3	Lancer l'évolution	40
6.	Conclusion	40
	Conclusion générale	41
	Bibliographie.....	42