



République Algérienne Démocratique et Populaire



Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene

Département Informatique

Mémoire de Master

Spécialité

Systemes Informatiques Intelligents

Thème

Prédiction des intérêts des utilisateurs dans les réseaux sociaux par apprentissage profond

Organisme d'accueil : CERIST



Présenté par :

- Mr. BOUDHAR Mohamed El Amine
- Mr. GUELLATI Mehdi Anis

Proposé par :

Mme. BOULKRINAT Nour El Houda

Devant le jury composé de :

Mme. BOUGHACI Dalila Présidente
Mme. BENSEBAA Amina Membre

Binôme N° : 036 / 2023

Résumé

L'avènement des réseaux sociaux a entraîné une profonde transformation de la communication, du partage d'informations et de la découverte de contenus en ligne. Avec des millions d'utilisateurs dans le monde entier, ces plateformes ont généré une quantité colossale de données. La prédiction des intérêts des utilisateurs est désormais cruciale pour personnaliser les recommandations de contenu et éviter d'être submergé par des informations qui ne correspondent pas à leurs intérêts futurs. Cependant, peu de recherches et de travaux ont été axés sur la prédiction des intérêts des utilisateurs en utilisant l'apprentissage profond. Notre objectif est donc d'exploiter cette abondance de données pour répondre de manière précise et pertinente au besoin de prédire les intérêts des utilisateurs en utilisant des réseaux de neurones.

Pour atteindre cet objectif, nous avons utilisé la dataset Yoochoose que nous avons adapté à nos besoins spécifiques. Nous avons exploité les informations relatives aux recherches effectuées par les utilisateurs, pour construire notre modèle basé sur un réseau neuronal récurrent, plus précisément un modèle LSTM (Long Short-Term Memory). Ensuite, nous avons comparé les performances de notre modèle LSTM avec celles d'autres techniques populaires telles que le KNN, le RNN, les arbres de décision et la régression logistique. Les résultats de nos expérimentations ont révélé que notre modèle LSTM surpassait les autres méthodes en termes de précision et de capacité prédictive.

Mots clés : réseaux sociaux, apprentissage profond, prédiction d'intérêts, LSTM, modèle de prédiction, profil social.

Abstract

The advent of social networks has brought about a profound transformation in communication, information sharing, and online content discovery. With millions of users worldwide, these platforms have generated an enormous amount of data. Predicting user interests has become crucial in personalizing content recommendations and avoiding being overwhelmed by information that does not align with their future interests. However, only a few studies and works have focused on predicting user interests using deep learning. Therefore, our goal is to leverage this wealth of data to accurately and relevantly predict user interests using neural networks.

To achieve this objective, we utilized the Yoochoose dataset, which we adapted to our specific needs. We exploited information related to user searches to construct our model based on a recurrent neural network, specifically a Long Short-Term Memory (LSTM) model. Subsequently, we compared the performance of our LSTM model with other popular techniques such as KNN, RNN, decision trees, and logistic regression. The results of our experiments revealed that our LSTM model outperformed the other methods in terms of accuracy and predictive ability.

Keywords : social networks, deep learning, interest prediction, LSTM, prediction model, social profile.

Dédicaces

Je dédie ce modeste travail

À mes parents qui, depuis mon plus jeune âge, ont toujours fait leur maximum en consacrant leur temps et leur argent pour m'éveiller et m'encourager dans mes passions.

C'est grâce à vous et pour vous que j'ai réalisé mon mémoire.

Aucun mot sur cette page ne saurait exprimer ce que je vous dois, ni combien je vous aime.

À mon binôme Mehdi,

à mon cher frère, à ma chère sœur,

à ma promotrice Nour El Houda Boulkrinat,

à ma grand-mère et à ma tante, les personnes les plus importantes de ma vie et les véritables supportrices tout au long de ma carrière, je vous dédie ce mémoire en témoignage de mon amitié sincère.

À tous ceux qui m'ont soutenu, qu'ils trouvent ici l'expression de ma profonde reconnaissance.

Mohamed

Dédicaces

Je souhaite dédier ce travail à toutes les personnes qui ont joué un rôle important dans ma vie et dans la réalisation de ce mémoire.

Tout d'abord, je dédie ce travail à Allah, qui m'a accordé la force, la persévérance et la sagesse nécessaires pour accomplir cette étape de ma vie.

À mes parents, vous êtes mes piliers, ma source de soutien et d'amour inconditionnel. Votre dévouement, vos sacrifices et vos encouragements constants ont été les moteurs de ma réussite. Je vous suis infiniment reconnaissant pour tout ce que vous avez fait pour moi.

À ma promotrice, Mme Nour El Houda Boulkrinat, je tiens à vous remercier pour votre encadrement, votre expertise et vos conseils précieux. Je vous suis profondément reconnaissant pour votre investissement.

À Manel, ma meilleure amie, tu as été ma confidente, mon soutien et ma source d'inspiration. Ta présence et ton soutien inconditionnel ont été inestimables. Je te suis reconnaissant pour tous les moments partagés et pour ton amitié véritable.

À mon binôme, Mohamed, tu as été mon partenaire de travail tout au long de cette aventure. Notre collaboration et notre soutien mutuel ont été essentiels pour mener à bien ce projet.

À ma famille, mes frères, mes tantes et mon grand-père. Votre amour, vos encouragements et votre présence ont été d'une importance capitale pour moi.

À mes amis, votre amitié sincère et votre présence à mes côtés ont rendu cette aventure encore plus spéciale.

Mehdi

Remerciements

En préambule à ce document, nous tenons à exprimer notre gratitude envers ALLAH qui nous a accordé l'aide, la patience et le courage tout au long de cette année d'études.

Nous souhaitons adresser nos sincères remerciements à toutes les personnes qui nous ont apporté leur aide et ont contribué à l'élaboration de ce document, ainsi qu'à la réussite de cette année universitaire exceptionnelle.

Nos remerciements vont tout d'abord à notre encadrante, Mme BOULKRINAT Nour El Houda, pour sa disponibilité constante durant la réalisation de ce document. Nous lui sommes reconnaissants pour son inspiration, son aide précieuse et son suivi attentif.

Nos remerciements vont également à Mme BOUGHACI Dalila et Mme BENSEBAA Amina qui ont accepté bienveillamment de participer à l'évaluation de ce document.

Nous n'oublions pas nos parents, dont la contribution, le soutien et la patience ont été inestimables tout au long de cette étape importante de notre vie.

Enfin, nous souhaitons exprimer nos sincères remerciements à tous nos proches et amis, qui nous ont constamment encouragés tout au long de la réalisation de ce document.

Nous tenons à remercier du fond du cœur toutes les personnes qui ont contribué de près ou de loin à notre réussite. Merci à tous et à toutes.

Table des matières

Introduction Générale	1
-----------------------------	---

Chapitre 1 : Réseaux sociaux

1. Introduction	4
2. Définitions	4
3. Types des réseaux sociaux	5
4. Catégories des réseaux sociaux numériques	5
5. Les éléments d'un réseau social	6
5.1. Les liens	6
5.2. Les nœuds	6
5.3. Groupes	6
5.4. Les graphes de contenu social	7
5.5. Tags	7
5.6. Ressources (Contenu)	7
6. Type d'utilisateur	7
7. Profilage social	8
7.1. Notion de profil utilisateur	8
7.2. Définitions du profilage social	9
7.3. Méthodes d'acquisition des données	9
7.3.1. Méthode explicite	9
7.3.2. Méthode implicite	9
7.4. Construction du profile social	9
7.4.1. Modélisation comportementale	9
7.4.2. La modélisation des intérêts	9
7.5. Evolution et mise à jour du profil social	9
8. Prédiction des intérêts des utilisateurs	10
8.1. Définition	10
8.2. Solutions existantes	10
8.2.1. Prédire à partir des clics	10
8.2.2. Prédire à partir des tags	11
8.2.3. Prédire à partir des ressources	11
9. Conclusion	12

Chapitre 2 : Apprentissage automatique

1. Introduction	14
-----------------------	----

2. Définition	14
3. Type d'apprentissage automatique	15
3.1. Apprentissage supervisé	15
3.2. Apprentissage non supervisé	16
3.3. Apprentissage semi-supervisé.....	16
3.4. Apprentissage par renforcement	17
4. Tâches d'apprentissage automatique.....	18
4.1. La classification	18
4.2. La régression.....	19
4.3. Le clustering (Regroupement)	19
5. Modèles superficiels	20
5.1. Arbre de décision (Decision Tree)	21
5.2. Machine à vecteurs de support (SVM).....	21
5.3. K plus proches voisins (KNN).....	21
5.4. K-means.....	21
6. Apprentissage Profond (Deep Learning)	21
6.1. Types de modèles basés sur le Deep Learning	24
6.1.1. Perceptron multicouche (MLP)	24
6.1.2. Réseaux de neurones convolutifs (CNN).....	24
6.1.3. Réseau de neurones récurrent (RNN)	25
6.1.4. Réseaux de neurones Long-Short Term Memory (LSTM).....	26
7. Réseaux de neurones et prédiction d'intérêts	27
8. Conclusion	29

Chapitre 3 : Conception du modèle

1. Introduction.....	32
2. Positionnement du travail	32
3. Description du dataset.....	33
4. Prétraitement du dataset.....	35
4.1. Nettoyage de données.....	35
4.2. Traitement du temps.....	36
4.3. Attribution des intérêts	36
5. Préparation des données.....	38
5.1. Sélection des données d'entrée et de sortie	38
5.2. Reshape des données	38
6. Choix du réseau de neurones.....	39

7. Modèle LSTM	40
7.1. Architecture	40
7.2. Compilation	41
7.3. Entraînement du modèle	42
8. Architecture générale du système	42
9. Diagrammes UML représentatifs du système	43
9.1. Diagramme des cas d'utilisation	44
9.1.1. Description de diagramme des cas utilisation	44
9.2. Diagramme de séquence	45
9.2.1. Définition	45
9.2.2. Scénario d'authentification	46
9.2.3. Scénario de recherche	47
9.2.4. Scénario de prédiction des intérêts futurs	48
9.3. Diagramme de classes	49
9.3.1. Définition	49
9.3.2. Passage du diagramme de classes au modèle relationnel	49
10. Conclusion	50

Chapitre 4 : Réalisation et mise en œuvre

1. Introduction	52
2. Résultats du modèle	52
2.1 Extraction des intérêts à partir des résultats	52
2.2. Résultats obtenus	52
2.3. Evaluation du modèle	53
2.3.1. Métriques d'évaluation	53
2.3.2. Comparaison avec d'autres techniques de prédiction	54
3. Environnement de travail	55
3.1. Environnement matériel	55
3.2. Environnement logiciel	55
3.2.1. Python 3.10.5	56
3.2.2. Flask	56
3.2.3. SQLAlchemy	56
3.2.4. Jinja2	56
3.2.5. HTML	56
3.2.6. CSS	56
3.2.7. JavaScript	57

3.2.8. Jupyter	57
3.2.9. Visual Studio Code	57
4. Interfaces de l'application	57
4.1. Inscription	57
4.2. Authentification	58
4.3. Espace utilisateur	58
4.3.1. Profil	59
4.3.2. New Post	59
4.3.3. My Posts.....	60
4.3.4. Recherche	61
4.4. Espace administrateur.....	61
4.4.1. Delete	62
4.4.2. Posts	62
4.4.3. History.....	63
4.4.4. Predict	63
4.4.5. Future interests	64
5. Conclusion	64
Conclusion générale	71
Bibliographie.....	66