

Université de Franche-Comté

Projet d'Initiation à la recherche
L2 - CMI

Complétion (semi-)automatique

BOITEUX Elouan
BENALI Samia



Table des matières

Introduction	2
1 Les différentes approches utilisées	3
1.1 Modèles basés sur des règles	3
1.2 Modèles basés sur des statistiques	3
1.3 Modèles basés sur l'intelligence artificielle	3
1.4 Modèles basés sur le deep learning	3
2 Les algorithmes de calcul de distance	5
2.1 Distance d'édition	5
3 Chaîne de Markov	6
4 Notre outil de complétion (semi-)automatique	7
Conclusion	8
Résumé	9

Introduction

Dans le cadre de notre projet de Recherche du CMI informatique de l'Université de Franche-Comté, encadré par Monsieur Héam, nous avons travaillé sur la complétion (semi-)automatique.

Ce projet nous a permis de découvrir ce qu'était la complétion automatique et la complétion semi-automatique et de comprendre sur quoi ce repose ces deux notions.

La complétion automatique est un processus par lequel un système va prédire et compléter une entrée en fonction de certaines données et contextes. Cependant, la complétion semi-automatique, est une assistance permettant au système de proposer des options tout en laissant à l'utilisateur la décision finale.

La complétion (semi-)automatique peut être utilisée dans de nombreuses applications : une saisie sur clavier, une complétion de code, une recherche sur un moteur de recherche, une assistance virtuelle etc. . .

Ce rapport va nous permettre, dans un premier temps, d'étudier les différentes approches qui existent ainsi que les différents algorithmes de distance. Ensuite nous parlerons des chaînes de Markov pour la gestion d'historique et enfin vous retrouverez l'application que nous avons créer permettant une complétion semi-automatique.

Chapitre 1

Les différentes approches utilisées

1.1 Modèles basés sur des règles

Pour proposer des suggestions, ce modèle utilise des algorithmes simples basés sur des règles préprogrammées comme la correspondance de préfixes ou une séquence que nous allons donner en amont par exemple. Ces algorithmes vont être gérés principalement grâce à des dictionnaires statiques ou des listes. Cette implémentation est plutôt rapide et simple à mettre en place, elle est cependant très peu flexible et empêche donc une utilisation complexe. . .

1.2 Modèles basés sur des statistiques

Pour proposer des suggestions, ce modèle utilise des statistiques fournies grâce aux données historiques. Cela permettra de prédire des séquences comme avec le modèle de Markov ou la méthode TD-IDF. Cette implémentation permet d'obtenir des résultats rapidement, on a cependant aucune compréhension sémantique donc les suggestions ne conviendront que rarement au contexte. . .

1.3 Modèles basés sur l'intelligence artificielle

Pour proposer des suggestions, ce modèle utilise des algorithmes basés sur l'intelligence artificielle et les réseaux neuronaux. Avec l'apprentissage supervisé et non supervisé, ces modèles apprennent des motifs complexes à partir des données. Ils peuvent inclure des algorithmes comme les forêts aléatoires ou les régressions pour fournir des prédictions plus contextuelles. Cette implémentation permet d'être efficace face à des problèmes très complexes et de répondre à des demandes rares. Cependant, ce genre d'implémentation nécessite énormément de temps de calculs et de ressources. . .

1.4 Modèles basés sur le deep learning

Pour proposer des suggestions, ce modèle utilise des algorithmes basés sur l'amélioration en temps réel. C'est à l'utilisateur de faire des choix et ces mêmes choix sont mémorisés pour une utilisation personnalisée et plus précise. Cette implémentation est donc très adaptable et permet des réponses précises avec un sens sémantique. Cependant, cette implémentation est complexe et très longue à mettre en place puisque les

choix de l'utilisateur sont nécessaires et les réponses dépendront totalement des données collectées. . .

Chapitre 2

Les algorithmes de calcul de distance

2.1 Distance d'édition

Chapitre 3

Chaîne de Markov

Chapitre 4

Notre outil de complétion (semi-)automatique

Conclusion

Résumé

Le résumé.