
Enhancement of Question Answering System Accuracy via Transfer Learning and BERT

— Kai Duan, Shiyu Du, Yiming Zhang , Yanru
Lin , Hongzhuo Wu and Quan Zhang —

BADAYODI Samson

INTRODUCTION

Dans le domaine de la question-réponse basée sur les bases de connaissances, l'amélioration de la précision des systèmes est un enjeu majeur pour garantir des réponses pertinentes et fiables. L'utilisation de modèles avancés tels que BERT et l'apprentissage par transfert offre de nouvelles perspectives pour renforcer la corrélation entre entités et prédicats dans le contexte du Chinese Knowledge Base Question Answering (CKBQA).

Problématique

La complexité de l'entité linking en chinois et la nécessité de renforcer la corrélation entre entités et prédicats posent des défis importants pour les systèmes de question-réponse basés sur les bases de connaissances. Les méthodes existantes présentent des limites en termes de précision et de généralisation, notamment en raison des spécificités de la langue chinoise et de la diversité des entités.

Etat de L'art

Lehmann, J. et al. "**Dbpedia—a large-scale, multilingual knowledge base extracted from wikipedia.**" Semant. Web 2015, 6, 167–195.

Bollacker, K. et al. "**Freebase: A collaboratively created graph database for structuring human knowledge.**" Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data.

Suchanek, F.M. et al. "**Yago: A core of semantic knowledge.**" Proceedings of the 16th International Conference on World Wide Web, 2007.

Duan, N. "**Overview of the nlpcc-iccpol 2016 shared task: Open domain chinese question answering.**" Natural Language Understanding and Intelligent Applications, 2016.

Solution Proposée

Pour répondre à ces défis, l'auteur a proposé un cadre BAT-KBQA basé sur BERT et l'apprentissage par transfert. Ce cadre intègre des modèles de réseaux neuronaux tels que BiLSTM-CRF, CNN et Softmax pour renforcer la corrélation entre entités et prédicats. En combinant ces approches, l'auteur a cherché à améliorer la précision du système de question-réponse dans le contexte du CKBQA.

RÉSULTAT

Les résultats obtenus ont démontré que le modèle développé a considérablement amélioré les performances globales du KBQA, avec des performances supérieures à celles des méthodes existantes. Le modèle a affiché une précision accrue sur l'ensemble de données NLPCC-ICCPOL2016 KBQA, avec un score F1 moyen de 87,74%, soulignant ainsi son efficacité et son potentiel de généralisation.

PERSPECTIVE

Pour les perspectives futures, l'intégration de différentes tâches et l'exploration de méthodes d'apprentissage de représentations de graphes de connaissances pour obtenir des fonctionnalités plus riches pourraient être envisagées. Ces développements pourraient contribuer à renforcer davantage les performances du système de question-réponse et à étendre son application à d'autres domaines.

CONCLUSION

En conclusion, le cadre BAT-KBQA basé sur BERT et l'apprentissage par transfert s'est avéré efficace pour améliorer la précision du système de question-réponse dans le contexte du CKBQA. Les résultats obtenus ont confirmé la pertinence de cette approche et ouvrent la voie à de nouvelles avancées dans le domaine de la question-réponse basée sur les bases de connaissances en chinois.