Kaiwei Cen

Mobile: 031-684885996 | Email: k.cen@students.uu.nl | Linkedin: kaiweicen | Github: github.com/kaiweicen

EDUCATION

Utrecht University   Utrecht, Netherlands
MSc in Artificial Intelligence GPA: 7.31/10   *Sept. 2021 – Present*

- Related Coursework: Advanced Machine Learning (7.5), Research Internship, Reasoning about Meaning in Linguistic Communication (7.5)

Foshan University   Foshan, China
B.E in Network Engineering (Internet of Things track)   *Sept. 2016 – July 2020*

RESEARCH EXPERIENCE

Research Internship   November 2021 – March 2022
Supervised by Dr.Tejaswini Deoskar   *Utrecht University, Netherlands*

- Applied a model which extends Bert by a linear layer at the top for the semantic role classification task. The model only uses sentence-predicate pairs as the input feature but outperforms the traditional methods which heavily rely on feature engineering.
- Preprocessed and applied the model on the Chinese Proposition Bank 1.0 (CPB 1.0) and the English Universal Proposition Bank (English UPB) datasets.
- Conducted quantitative on the performance of the model on the aforementioned datasets.
- Conducted a comparative analysis on the model's differing performance in semantic role label classification on the two aforementioned datasets.

BACHELOR THESIS

Intelligent Question Answering System for Administrative Service   July 2018 – April 2020
Supervised by Professor Weigang Guo, won the excellent bachelor thesis award   *Foshan University, China*

- Knowledge graph construction: applied the BERT-BiLSTM-attention-CRF model for entity extraction and relation extraction on the governmental data crawled on the web. Constructed a knowledge graph with 210823 entities, 657372 relationships and 23862 attributes using Neo4j based on the extracted data.
- Question-answering system development: used the BERT-BiLSTM-CRF model to recognize governmental entities in the input questions. Applied the BERT-BiLSTM-CRF model on matching the input question to the most semantically similar relation in the knowledge graph.
- The accuracy of the question-answering system is 80.02%, 10% higher than the former version of the Guangdong governmental question-answering system.

AWARD

Open-domain Question Answering System in Chinese   January 2018 – June 2018
The 6th Teddy Cup National Data Mining Competition, ranked 1% among 3000 teams   *Foshan University, China*

- Designed a hierarchical semantic extraction model based on LSTM and CNN to extract the semantic meaning of input questions.
- Applied the attention mechanism to integrate the semantics of the factual part into the understanding of the core question, which can increase the similarity of the input question and the correct answer. Compared to the model without the attention mechanism, the precision was improved by 6.7%.
- Trained the model on the Baidu QA dataset, which contains 36,181 questions and 181,661 answers.

TECHNICAL SKILLS

Algorithms: Reinforcement Learning, Causal Inference
Machine learning frameworks: PyTorch
Vision-and-language model frameworks: Visiolinguistic Transformer Architectures (VOLTA)