

Andong Chen

+86 155-2646-6911 | ands691119@gmail.com
Beijing

EDUCATION

Hunan University	Sep 2015 - Jun 2019
Electrical Engineering and Automation Bachelor Electrical engineering	HUNAN
Hunan University Formula E Racing Team-Leader of the Electric Group, won the National Formula Student Championship in 2019.	
MINZU UNIVERSITY OF CHINA	Sep 2019 - Jun 2022
Natural Language Processing Master Information Technology	Beijing
Third Prize of Huawei Cup Graduate Mathematical Modeling Competition	
School-level academic third-class scholarship	
Queen Mary University of London(Exchange Student)	Jan 2021 - Jan 2022
Big Data Science Master Science and Engineering	London

RESEARCH EXPERIENCE

A Joint Model for Representation Learning of Tibetan Knowledge Graph Based on Encyclopedia (Paper : CCF)	Dec 2019 - Feb 2021
Institution: National Language Resources Monitoring and Research Center	Beijing
Introduction: The thesis mainly constructs the first Tibetan encyclopedic knowledge graph and improves the representation ability of the Tibetan knowledge graph.	
•By crawling Tibetan encyclopedia knowledge, the description information of triples and terms is obtained. •Use the TransE model to train the knowledge graph representation of the triples, and extract the semantic representation of the description information through the CNN model. •Finally, merging the two representations to improve the representation ability of the Tibetan knowledge graph.	
JCapsR: A joint capsule neural network learning model for Tibetan language knowledge graph representation(Paper: CCL)	Jan 2021 - May 2021
Institution: National Language Resources Monitoring and Research Center	Beijing
Introduction: In this paper, we propose a joint capsule neural network (JCapsR) learning model for Tibetan language knowledge graph representation based on the previously constructed Tibetan language knowledge graph.	
• First, we use the TransR model to generate a structured information representation of the Tibetan knowledge graph.	
•Second, the Transformer model, which incorporates multi-headed attention and relational attention, is used to represent the textual description information of Tibetan language entities. • Finally, JCapsR is used to further extract the relationships of triads in the semantic space of the knowledge graph, and to fuse the textual description information of entities with the structured information to obtain a better representation of the Tibetan knowledge graph.	
Intelligent service for automatic generation of Tibetan reading problems in elementary schools (Patent)	Jan 2021 - Mar 2021
MINZU UNIVERSITY OF CHINA	Beijing
Introduction: First, expand primary school Tibetan reading comprehension articles, and then build a question generation model.	
• Extract the features from the Tibetan texts of each grade, and train logistic regression as a classification model • Obtain a large number of Tibetan texts through crawlers and then obtain a large number of texts that can be read by primary school students through the classification model • Trained through the LSTM model and Attention mechanism Problem generation model. Finally, enter the article and answer, and finally generate the question.	

EXTRACURRICULAR ACTIVITIES

Ali Tianchi Speech Recognition Novice Competition	Mar 2021 - Apr 2021
Role: contest designer	HangZhou
Introduction: Voice recognition introductory novice competition initiated by Tianchi-food voice recognition	
• In Ali Tianchi. I provided speech recognition knowledge, written Baseline, and live streaming Tianchi to interpret the competition questions and codes. • During the project, I compiled a Baseline based on CNN for speech recognition and planned a live broadcast of Tianchi with the team learning organizer Datawhale.	
Hands-on data analysis (Open source project)	Jun 2020 - Aug 2021
Role: Project sponsor and content builder	Beijing
Introduction: An open-source data analysis project for everyone to learn by themselves	
• Use the Titanic data set to divide tasks according to data observation, data manipulation, data visualization, data modeling, and model evaluation. • Each task is planned and written according to "problems encountered"- "provide ideas"- "code actual combat"- "summary review". It has been open-sourced on GitHub. There are currently 280 Stars and are recommended by the Datawhale public account. • Open source link: https://github.com/datawhalechina/hands-on-data-analysis	

INTERNSHIP EXPERIENCE

Cambricon Technologies Corporation Limited- Ai Lab	May 2021 - Aug 2021
Processor Platforms Department	Beijing
Introduction 1: The lab mainly does quantification of deep learning models	
• Do quantification of the speech model Waveglow and evaluate the quality of quantification by crowdsourcing. • Build the lab's first NLP quantitative reasoning solution with mentors. • Compare with Nvidia's Transformer quantization solution to	

optimize the company's internal quantization solution.

Introduction 2: Reproduction of GPT-3 model

- Participated in the group's project on reproducing GPT-3 model
- Mainly responsible for tokenization and data cleaning tool construction.
- Optimization of downstream tasks for machine reading comprehension