Wenbo Zhang

wenbozhangjs@gmail.com | +1 814-852-9399

EDUCATION

Pennsylvania State University   **Pennsylvania, USA**
Doctor of Philosophy in Informatics   *Aug 2021 – Present*
- **Research**：*AI framework to predict user's medical emergency based on code-mixed (English and Swahili) text*

University of Southern California   **California, USA**
Master of Science in Electrical Engineering   *Aug 2016 – May 2018*
- **Research**：Machine learning for phenotypic pattern identification of adolescents with drug usage

University of Electronic Science and Technology of China   **Sichuan, China**
Bachelor of Engineering in Renewable Energy Materials and Devices   *Sept 2011 – Jul 2015*

QUALIFICATIONS

- More than two years working experience related to NLP and speech processing (especially speech synthesis)
- Experience with deep learning model construction and parameter tuning with Tensorflow and Pytorch
- Proficient in Python, Matlab, R; familiar with C, SQL, shell, git
- Familiar with development under Linux and platforms (such as google cloud platform, Alibaba cloud)
- Excellent critical thinking skills, academic writing skills, and teamwork ability

RESEARCH EXPERIENCE

Project 1: A tele-triage framework to lower the risk of maternal and neonatal death in Kenya   *Oct 2021 – June 2022*
- Proposed an NLP framework to predict the user's medical emergency level based on code-mixed SMS messages
- Transformed the emergency level prediction into a text classification problem with multi-lingual pretrained models
- Applied multi-lingual pretraining and continual pretraining to deal with the code-mixed text automatically
- Achieved the F1 score of 0.774 and deployed inside the PROMPTS (system developed by Jacaranda Health)
- Reduced the monthly system management cost by 22.8% and PROMPTS helpdesk's workload by ~12%

Project 2: Health data processing and pattern recognition   *Feb 2018 – Jul 2018*
- Identified potential students who may use alcohol, cigarette or marijuana in the future
- Cleaned original data and applied sampling techniques to solve unbalanced problem in the data
- Implemented feature selection techniques to reduce the complexity of machine learning models
- Constructed multiple classification models via Python for pattern recognition and applied ensemble methods
- Compared the results from different classification models and integrated results from different models

PUBLICATIONS

- **Wenbo Zhang**, Hangzhi Guo, Prerna Ranganathan, Jay Patel, Sathyanath Rajasekharan, Nidhi Danayak, Manan Gupta, Amulya Yadav. A Continual Pre-training Approach to Tele-Triaging Pregnant Women in Kenya. In the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI), February 2023 (under review).

WORK EXPERIENCE

Algorithm Engineer at Kingsoft AI Lab, Beijing, China   **Jan 2019 – Jul 2021**
Projects related to knowledge graph
- Applied NLP techniques to develop text processing modules inside knowledge graph

Project 1: Open domain knowledge graph construction for Kingsoft electronic notebook website   *Jan 2021 – Jul 2021*
- Designed modules (name entity recognition and relation extraction) for Chinese knowledge graph construction
- Conducted data cleaning on the public dataset and built baseline for the module of relation extraction
- Preprocessed and cleaned Chinese Wikipedia data into BIO format for the task of name entity recognition
- Developed name entity recognition module with the model (BERT+CRF) and achieved F1 score of 0.82
- Optimized name entity recognition module through acceleration tools and increased 10% of inference speed

Projects related to speech synthesis

- Applied NLP techniques to design text processing modules to extract semantic information from input sentences
- Employed deep neural network to construct the Chinese speech synthesis system based on semantic information
- Improved the naturalness of generated speech by introducing prosody and polyphone pronunciation prediction
- Extended the single-speaker English speech synthesis system to a multi-speaker speech synthesis system

Project 1: English multi-speaker speech synthesis system for novel website   *Jul 2020 – Dec 2020*

- Developed a system which generated speech with someone's tone through few minutes' voice recordings
- Constructed an end-to-end English single-speaker speech synthesis system via seq2seq framework
- Introduced speaker identity to transform model into a multi-speaker system which supports diverse human voice
- To verify speaker's voice, introduced speaker encoder with speech as input and speaker embedding as output
- Utilized the generalized end-to-end loss for training, Equal Error Rate (EER) of encoder reduced from 13% to 4%
- Created a prototype for audiobook reading on English novel translation website to support multiple human voices

Project 2: End-to-end framework for Chinese polyphone pronunciation prediction   *April 2020 – Jul 2020*

- Built end-to-end framework for pronunciation prediction of Chinese polyphone with multi-phonemic values
- Data acquisition of large amount of sentences from Weibo with extensive data cleaning and polyphone capture
- Explored neural network with polyphone position encoding and BLSTM layer for context information extraction
- Achieved 94% accuracy of polyphone prediction with the established comprehensive neural network model

Project 3: NLP based text prosody prediction   *Jan 2020 – Mar 2020*

- Considered the prosody (short pronunciation break among words) prediction as the sequence tagging problem
- Preprocessed data with word segmentation and Part-of-Speech (POS) tagging, generated new feature-label pairs
- Realized prosody prediction by Conditional Radom Field (CRF), experimented combinations of feature template
- Enhanced CRF model by introducing BLSTM model considering both past and future information
- Achieved F1 score of 0.93 through CRF and improved the synthesized speech naturalness and quality

Project 4: End-to-end Chinese speech synthesis   *Jan 2019 – April 2020*

- Implemented end-to-end Text-to-Speech (TTS) system through Natural Language Processing (NLP) techniques
- Realized text processing module through regular expression and tokenization to transform texts into specified format
- Built seq2seq based models (Tacotron, Deepvoice, series of Transformer) to improve the synthesized voice quality
- Used length regulator to replace the attention mechanism to reduce the stammer problem in synthesized speech
- Replaced vocoder modules (WORLD, or deep learning model: LPCNet, MelGAN) to enhance speech naturalness
- Created speech synthesis API through Django and applied on the policy Question Answer (QA) system

Intern at Comprehend Information Technology Co., Ltd., Suzhou, China   Jun 2017 – Aug 2017

NLP-based data mining on the traffic data (electronic checkpoints data) accessed from Suzhou City Brain

- Partitioned city into different function regions (education area, central business area etc.)
- Constructed a database via Python and SQL to transform raw data into data records indicating vehicle paths
- Utilized the topic model from Natural Language Processing to segment the city into various function regions
- Evaluated model performance by comparing the result with official city planning diagram