

ELLIE LI

<https://evakli11.github.io>

Email : mli7@wpi.edu

Mobile : +1-774-578-8477

Address: 75 elm St, Worcester, MA

EDUCATION

Worcester Polytechnic Institute (WPI)

Worcester, MA

Master of Science in Data Science; GPA: 3.72/4.00

Aug. 2017 – May. 2019

Coursework: Statistics, Database System, Big Data Management, Deep Learning, Artificial Intelligence

Shanghai University (SHU)

Shanghai, China

Bachelor of Management Information Systems; GPA: 3.45/4.00

Aug. 2012 – Jun. 2016

SKILLS

Languages: Python, R, SQL, HTML **Software:** AWS(EMR), Google Cloud, Tableau, MongoDB, MS Excel

EXPERIENCE

Research Assistant, Mathematical Science, WPI

Jan. 2019 – Apr. 2019

- Employed neural networks to do function approximation and parameter estimation of partial differential functions
- Wrote python script to find lower and upper bounds and neural importance of feedforward neural networks which fit Genz integrand families functions and visualized the results

NLP Intern, iFLYTEK, Hefei, China

May. 2018 – Aug. 2018

- Worked in the auto-grading group and aimed to build model to automatically label the role of sentences, such as topic sentence, in TOEFL writings
- Extracted and validated candidate features of sentences in TOEFL writings with experts
- Retrained pre-trained Word2Vec and Glove models for sentence embedding by adding our corpus
- Built a model by combining bidirectional LSTM and CRF with TensorFlow to conduct sequence tagging on TOEFL writings and improve the test accuracy to 0.71

Data Engineering Intern, Westwin, Shanghai, China

Jan. 2016 – Jul. 2016

- Involved into Store Locator project, wrote a web crawler with Python to automatically collect basic information of US retailers and archived the information with a proprietary database
- Participated into ShoppingSite project, checked and updated goods information with SQL queries

PROJECTS

Information Retrieval, Rap Maker, WPI

Mar. 2019 – Apr. 2019

- Aimed to build a model can generate rap lyrics automatically according to user requirements
- Added information retrieval module with keywords extraction and sentiment analysis functions before LSTM in order to get customized training set related to user requirements
- Built LSTM model with Keras to generate lyrics and added rhyme collection part into model to guarantee the lyrics are generated in rhyme
- Adopted pyttxs3 to combine lyrics with beats and implemented a simple UI to show results

Graduate Qualifying Project, United Technologies(UTC), Framington, CT

Sep. 2018 – Dec. 2018

- Worked with UTC HR analytics team to create an internal job/candidate recommend system
- Designed algorithms to clean up noisy words and extract n-grams skills from over 20,000 UTCs job descriptions and trained Word2Vec to obtain semantic embedding for extracted skills
- Adopted KNN to find each skills synonyms to generate extended skill sets and built inverted index through ElasticSearch
- Built UI allowing users to upload resumes, automated the system sending job recommendations via email and implemented similar functions for HR to get internal candidate recommendations

Machine Learning, Revenue Prediction, WPI

Nov. 2018 – Dec. 2018

- Aimed to implement a machine learning system to identify potential customers and predict their generated revenue
- Built pre-classified regression model and stacked regression model to conduct revenue prediction based on extremely skewed dataset ($< 1.3\%$ of non-zero rows in target feature) and reduced the loss(RMSE) from 1.84 to 1.67
- Applied weighted RNN model to predict revenue generated by customers visited more than one time and reduced the loss(RMSE) from 2.2 to 1.9