Wen-Yuh (Ken) Su

Chicago, IL | 312-927-2532 | wsu23@uic.edu

linkedin.com/in/suwenyu | github.com/suwenyu

WORK EXPERIENCES

Particle Media Inc., Mountain View, USA *May.2019-Aug.2019*

Machine Learning Intern *[Python, Word2Vec, XGBoost, FastText, Tensorflow]*

- Applied embedding similarity features by training word2vec to improve disambiguation of names for US cities and counties with F1 measure of 78.48 percent
- Implemented XGboost model to recommend the news to users living in the related area with accuracy of 90 percent
- Built the data pipeline for the push-notification system with extracted time and action with predicted event type by using SUTime and the dependency parser

University of Illinois Chicago, Chicago, USA *Mar.2019-May.2019*

Research Assistant *[Python, TensorFlow, FastText, Word2Vec, LSA]*

- Implemented Hierarchical LSTM which combined the word embedding sequence and the clause embedding to achieve high accuracy of 78.6 percent on classification task with 4 labels on Twitter HPV-related articles
- Compared effectiveness of the semantic representation by applying latent semantic analysis, Word2Vec, and FastText on the different level of words in HPV-vaccine-related data

Cathay Financial Holding Co. Ltd., Taipei, Taiwan *Aug.2017-May.2018*

Machine Learning Intern *[Python, Shell Script, Tensorflow, XGBoost]*

- Applied Lambdamart model on financial data such as financial behavior of users, truncation record, and customer profiles to find important features
- Implemented the Deep-Wide model with those features to predict the rank of the customers
- Predicted the average variation in the amount of credit card transaction of customers and the features from the Deep-Wide model with an accuracy of 21 percent (ndcg@10) by XGBoost

CLIP Labs, National Chengchi University, Taipei, Taiwan *Aug.2017-May.2018*

Research Assistant *[Python, C++, JavaScript, Tensorflow, Flask, MongoDB, D3.js, Restful API]*

- Emphasized risk detecting for word-level financial reports by training a self-attention LSTM model on the collection of financial reports with financial sentiment phrases, resulting in high accuracy of 88 percent (appeared in IEEE'20, 2nd author)
- Built a labeling system by using Flask framework, JavaScript and MongoDB for users to add annotations of multiword expressions for financial reports with visualization which increased the efficiency of labeling
- Visualize the financial statement by highlighting the strong words which the deep learning model learned

EDUCATION

University of Illinois Chicago, Chicago, USA *Aug.2018-May.2020*

Master of Science in Computer Science, GPA: 3.40

National Chengchi University, Taipei, Taiwan *Sep.2013-Jun.2017*

Bachelor of Science in Computer Science, GPA: 3.70

SELECTED PROJECTS

Data Mining and Text Mining, Chicago, USA *[FastText, Tensorflow, Python]* *Mar.2019-May.2019*

- Implemented the classification model by attention LSTM and pre-trained word embedding from FastText for the Twitter dataset and surpassed the baseline of 10 percent on the 3 labels classification task

Introduction to Machine Learning, Chicago, USA *[Tensorflow, Python]* *Dec.2018*

- Implemented a self-attention LSTM on Quora insincere questions classification task and resulted in high accuracy of 64.1 percent which improved 5 percent from the LSTM model

Digital Humanities Project, Taipei, Taiwan *[Python, JavaScript, Flask, MongoDB, Restful API]* *Feb.2017-Jun.2017*

- Improved the digit humanities website into parallel resulted in reducing 20 percent of users' the waiting time
- Implemented the data pipeline and Restful API contribute to the visibility and scalability of the system
- Built visual analysis and charts tools for users to refer the text mining results (Link: https://clip.csie.org/DHPNew)