Weicheng Zhang

wzhang81@jhu.edu | (+1) 443-319-3931 | GitHub | Portfolio

EDUCATION

Johns Hopkins University
MS in Information Security (Concentration in Computer Science, GPA 4.0/4.0 for tech courses)
August 2016 - May 2018, Baltimore, MD
Coursework: Natural Language Processing, Algorithms, Independent Research, Cryptography and Coding, Security Analytics, Security & Privacy in Computing, Advanced Topics in Computer Security

Beijing University of Posts and Telecommunications
B.Eng. in Communication Engineering (GPA 3.5/4.0) with Scholarship
September 2012 - July 2016
Honorable Mention, Mathematical Contest in Modeling, 2015
Coursework: Speech Recognition, Machine Learning and Pattern Recognition, Data Structures, Computer Networks, Linux Operating System, Image Processing, Discrete Mathematics, Linear Algebra

WORK EXPERIENCE

Center for Language and Speech Processing
Jun 2017 - Present, Baltimore
ML Research Assistant (Elasticsearch, Deeplearning4j, Word2Vec, TF-IDF) (Java, Python)
- Built a robust Lucene/ Elasticsearch based search engine for Cross-lingual IARPA MATERIAL. Currently indexing large datasets like Wiki, TREC and CACM.
- Monitored indexed data on server in real-time using Kibana.
- Accomplished using word embedding for query expansion. The model outperformed baseline by 12%.
- Implementing pseudo relevance feedback to improve the performance of Elasticsearch.

RESEARCH

Tsinghua University – Natural Language Processing and Computational Social Science Lab
Apr 2015 – Jun 2016, Beijing, China
NLP Researcher and Engineer (Neural Network, DeepWalk, SVM, t-SNE) (Java, Python, Matlab)
- Focused on information extraction, induction, classification, and exploration.
- Designed and implemented matrix factorization based highly discriminative representation learning method called Max-Margin DeepWalk. Paper accepted by IJCAI 2016 (Acceptance Rate 18.9%).
- Designed and optimized Online Neural Network based Max-Margin DeepWalk with better efficiency.
- Significantly outperformed other state-of-the-art representation learning methods like LINE by 16.2%.

PUBLICATION

Cunchao Tu, Weicheng Zhang, Zhiyuan Liu, Maosong Sun. Max-Margin DeepWalk: Discriminative Learning of Network Representation. *International Joint Conference on Artificial Intelligence (IJCAI 2016)* (Co-first author). [pdf]

PROJECTS

Spam Filter with Deep Learning and Information Retrieval (Capstone Project)
2018, Baltimore, MD
Software Developer (GRU, LSTM, CNN, TF-IDF, Elasticsearch) (Python, Java)
- Used Word2Vec trained word embedding for LSTM and CNN based spam classification.
- Used Elasticsearch for noise reduction by implementing weighted KNN according to scores.
- Outperformed baseline model by 30%, reaching nearly 99% of precision.

Real-time Detection of Social Network Attacks Based on Machine Learning
2017, Baltimore, MD
Software Developer (Crawler, Twitter REST/Streaming API, SVM) (Java)
- Implemented a machine learning based detecting system for real-time social media attacks.
- Succeeded in detecting 95% real-time malicious posts in testing using SVM for classification.

Developing De-anonymizing Attack and Defense to Social Networks
2017, Baltimore, MD
Software Developer (MLE, Jaccard) (Java)
- Implemented a de-anonymizing MLE attack towards social networks to reveal personal data in real world.
- Designed and accomplished a defense by changing the topological graph of the social network.

SKILLS

Java (Jblas, Liblinear, Deeplearning4j), Python (TensorFlow, Scikit-learn, Pandas, Numpy), Matlab, C++, HTML+CSS, C, JavaScript, RESTful API, Git, Tcpflow, Wireshark, Web Service, Algorithm Design, Unix/ Linux system, Elasticsearch, Lucene, Kibana, Android SDK, Web design, OO design