Huamin Zhang

● (571) 342-1885 ● moqi112358@gmail.com

CORE QUALIFICATIONS

- Technical Skills: Python, PyTorch, R, MySQL, MATLAB, Perl, Spark, C++, Bash, Linux
- Data analysis, Machine learning, Bioinformatics
- Passionate for learning new things and dealing with different information, keen on fast-paced work

EDUCATION

University of Illinois at Urbana-Champaign, Computer Science Department **Champaign, IL, U.S.A.**

- Master of Science Bioinformatics: Computer Science Concentration GPA:3.91/4.0 ***08/2016-05/2018***

Shanghai Jiao Tong University (SJTU), Bioinformatics and Biostatistics Department **Shanghai, China**

- Bachelor's degree Bioinformatics & Biostatistics GPA:3.59/4.0 ***09/2012-07/2016***
- Minor: Business Administration

EXPERIENCE

Genonova Inc. **Champaign, IL, U.S.A.**

Data Scientist Intern ***08/2017-12/2017***

- Created a database of 300K+ literature with MySQL and implemented text mining to detect variants that could explain existing diseases or predict diseases
- Designed a GATK pipeline to detect somatic mutation and germline mutation on AWS

RESEARCH EXPERIENCE (SELECTED)

HackerEarth Deep Learning Challenge #4: Tag Recommendation System ***11/2018- 12/2018***

- Cleaned text data and padded sequences for titles and articles data respectively
- Built a pre-trained Word2vec word embedding with gensim as inputs for neural networks in order to perform NLP tasks
- Implemented TextCNN model and Bidirectional LSTM model with titles and articles data as input using Pytorch
- Implemented model ensemble for TextCNN and Bidirectional LSTM model with equal weights
- Trained the model with approximately 1 million technology-related articles mapped to more than 30,000 different relevant tags and generated relevant tags from the given test set of articles
- Achieved 0.6072 overall F1 score and won the fifth place in the competition

VQA (Visual Question Answering) Challenge ***03/2018- 05/2018***

- Implemented VGGNet16 and ResNet152 model in the image channel that provided image representation with 82,783 images
- Implemented Deeper LSTM model in the question channel that provided question representation with 443,757 questions
- Combined the outputs of image channel and question channel with MLP (MultiLayer Perceptron) and MCB (Multimodal Compact Bilinear Pooling) method which based on element-wise multiplication
- Trained the 2-channel model with 4,437,570 answers in training annotations and achieved 56.6% overall accuracy with our best architecture on the validation data.

Automatic discovery of medical knowledge to mine Adverse Drug Reactions ***03/2018-05/2018***

- Mined adverse drug reactions from FDA Adverse Event Reporting System (FAERS)
- Built a healthcare information system which mined drug side effects from different online health forums and social networks using a probabilistic topic model based on PLSA
- Built a knowledge graph for drugs which contained three types of nodes: drug, adverse reactions and medical subject headings from the MeSH database. Mined knowledge for each edge from PubMed literature data. Made a possible explanation of the relation between drugs and adverse drug reactions using the knowledge graph.

KDD Cup 2017: Highway Tollgates Traffic Flow Prediction ***03/2017-06/2017***

- Cleaned 10076 rows dataset, removed outliers and special data points and created a standard and structured form of data
- Created features about time, traffic volume, weather and road topology to predict traffic flow
- Tuned the parameters, trained the model using best separate parameters for different algorithms, such as gradient boosting decision tree and long short-term memory network with Python and Keras package, compared the results, and analyzed why some algorithms were better than others for a deep understanding of different algorithms and their use cases
- Placed in top 5% of challenge to predict average tollgate traffic volume

LEADERSHIP & EXTRACURRICULAR ACTIVITIES

Captain of Basketball Team, School of Life Science & Biotechnology, SJTU ***06/2015-06/2016***

Team leader of Debate Team, School of Life Science & Biotechnology, SJTU ***06/2013-09/2014***