Ziyu Yang

Philadelphia, PA
-Email me on Indeed: http://www.indeed.com/r/Ziyu-Yang/9c9eb34292e4f359

Work Experience

Data Scientist/Machine Learning Engineer (Research Collaboration)

Penn Mutual Life Insurance Company

April 2019 to Present

• Data mining with real health care data (Attending Physician Statement).
• Built an Information Extraction system with Neural Networks (LSTM, BERT), which automated data processing.
Research Natural Language Processing/Understanding: Named Entity Recognition, In-
Interests formation Extraction, Entity Linking, Text Summarization/Simplification, Classifica- tion, EHR Data Analysis
Research Aspect Terms recognition with Distant Supervision Data (Scripting)
• Built a BERT-partial CRF model to solve the sequence labeling task without human labeled data.
• Integrated pre-trained models and knowledge bases to generate weak labels for model training.

Research Data Scientist

Industry ADP
• Applied Natural Language Processing (NLP) methods to human resources data.
• Developed a pipeline for extracting information from job postings.
• Built a Siamese LSTM with pre-trained MIMICK neural network to learn the similarity between raw and canonical job titles.

Radiology Sentence Simplification with Low-resource Labeled Data (Ongo- ing)
• Generated synthetic labeled data with GPT-3 and Bio-Clinical BERT.
• Retrained a radiology-based BART model with MIMIC III dataset.
• Fine-tuned a BART model with synthetic and limited labeled data.

Education

Ph.D. in Computer Science

Temple University

May 2023

M.S. in Statistics

Temple University

June 2017

B.S. in Statistics

University of Science and Technology of China

June 2015

Skills

- R
- SQL
- Data science
- Deep learning
- Statistics
- Machine learning
- Java
- Computer science
- Data mining
- Computer networking
- Human resources
- Python
- Scripting
- Shell Scripting
- PyTorch
- Research
- Linux
- Natural language processing
- Data analysis skills