Tianjian Li

Website: truthbutcher.github.io
Github: github.com/truthbutcher

Email: tli104@jhu.edu
Mobile: +86-18610290164

Education

- **Johns Hopkins University** — Baltimore, MD
  M.S. in Security Informatics — *August 2022 - May 2024(Expected)*
- **New York University** — New York, NY
  B.A. in Computer Science/Mathematics — *August 2017 - September 2021*
  Courses: *Operating Systems, Data Structures, Basic Algorithms, Machine Learning, Theory of Computation, Numerical Analysis*

Internship Experience

- **Baidu Inc.** — Beijing, China
  Machine Learning Engineer(Intern) — *Aug 2021 - Feb 2022*
  - Built a classification model on the influence of small paths on customers' driving experience with XGBoost.
  - Optimized route ranking model by experimenting with two strategies: 1D-CNN and multihead self attention in modeling sequential trajectory data.
  - Designed a Spatial-Temporal Graph Neural Network model to further improve the performance of the route ranking model to anticipate and dodge traffic jams.
- **Huawei Technologies Co., Ltd.** — Beijing, China
  Software Development Engineer(Intern) — *Jun 2020 - Aug 2020*
  - Refactored the code from MyBatis to Spring Data JPA and designed junit-based tests to compare data transmission throughput.
  - Researched compression algorithms and designed REST APIs based on GZIP and MLP to enable large files over 2GB to be efficiently transmitted between microservices.

Research Experience

- **Tsinghua University - Knowledge Engineering Group(KEG)** — Beijing, China
  Research Intern, Advisor: Jie Tang — *Mar 2022 - Present*
  - **Multilingual Language Model**: Pretrained GLM language model on 19TB of multilingual corpora then finetuned with task-specific prompts on sentence classification, named entity recognition and question answering.

Projects

- **Baidu AI Studio Regular Challenge 8/1035 (Machine Learning, Graph Neural Networks)**: Implemented Graph Convolution Networks with residual connection and label smoothing for academic paper classification contest. Rank 8 out of 1035 teams.
  Project Link: github.com/truthbutcher/arxiv-competition
- **Re-implementation of STFGNN model for traffic jam forcasting(Machine Learning, Spatio-Temporal Graph Neural Networks)**: Re-implemented STFGNN model in PaddlePaddle, modified model architecture to achieve an accuracy of over 80% in forecasting the time and severity of traffic jam in the next 48 hours.
  Project Link: github.com/truthbutcher/STFGNN-PaddlePaddle
- **Other Projects**: Self-study online course projects and research notes are maintained and regularly updated at github.com/truthbutcher/studymaterials

Honors and Awards

- New York University College of Arts and Sciences(CAS) Scholarship
- First Prize in National Olympiad in Informatics Provincial(NOIP)
- Global Top 5% in American Mathematics Contest(AMC12)

Skills Summary

- **Languages**: Python, JAVA, C/C++, SQL, Shell Scripting, Unix Commands(grep, sed)
- **Frameworks**: PyTorch(Distributed Training), TensorFlow, Keras, PaddlePaddle
- **Tools**: Docker, GIT, MySQL, Hadoop streaming, Spark, Vim, LaTeX