Yali Bian

Seattle, WA • (540) 998-0256 • yali@vt.edu

LinkedIn • Portfolio Website • GitHub • Publications

SUMMARY

Research scientist working on visual analytics, human-in-the-loop machine learning, interactive deep learning. Software engineer with 5 years of industry work experience in full-stack web, big data, and ML Infrastructure.

WORK EXPERIENCE

Senior Research Scientist AI/ML - Intel Labs

Santa Clara, CA   *Jun. 2022 - Present*

- Establishing an AI-assisted image annotation system to reduce annotation efforts and accelerate DL model building for object detection tasks, including real-time defect detection and working activity tracking.
- Building an exploratory data analysis platform with interpretable DL to help engineers understand and debug the manufacturing process and improve the yield and quality of Intel products.

Senior Software Engineer (ML Infra) - Instacart

San Francisco, CA   *Jan. 2022 – Jun. 2022*

- Built the ML pipeline to standardize the ML development life cycle and fully utilize existing ML platforms.
- Built a Feature Store (Marketplace) system to manage, monitor, and serve batch and streaming features.
- *MLOps, K8s, Snowflake, MLFlow, Airflow (Amazon MWAA), Sagemaker, Datadog, FastAPI, Kafka, S3*

Software Development Engineer - Akuna Capital

Chicago, IL   *May. 2019 – Mar. 2020*

- Led a team in building trade surveillance and compliance application to monitor and analyze live trading behaviors from the message streaming platform, and optimized the system processing efficiency by 72%.
- *Kafka, Solace, ZMQ, Node.js, React.js, CI/CD, Jenkins, Docker, asynchronous and multiprocessing*

Machine Learning Research Engineer Intern - Facebook

Seattle, WA   *May. 2018 – Aug. 2018*

- Built a two-tower sparse neural network between billions of Facebook groups and users, and improved the FB search system in recommending alumni groups.
- Extracting and analyzing billions of alumni user-group tuples using big data mining methods, including collaborative filtering, label propagation, and rule-based filtering.
- *Hive, Spark, Presto, DataSwarm, Digraph, FBLearner, GBDT, Word2Vec, Data Pipeline, Apache Airflow*

Machine Learning Research Intern - Nokia Bell Labs

Murray Hill, NJ   *Jun. 2017 – Aug. 2017*

- Developed a full-stack human-in-the-loop anomaly detection system for a temporal Telecom dataset.
- Built an end-to-end data processing system to render and explore millions of time points in real time.
- *Python, Flask, Scikit-Learn, D3.js, React.js, SVM, Interactive Machine Learning, Active Learning*

EDUCATION

- **Ph.D. in Computer Science and Application** − Virginia Tech (GPA: 3.97/4.00)   Aug. 2016 – Feb. 2022
- **M.S. in Computer Science and Application** − Virginia Tech (GPA: 3.97/4.00)   Aug. 2016 – May. 2019
- **M.S. in Computer Science and Technology** − Zhejiang University (GPA: 3.78/4.00)   Sep. 2013 – Mar. 2016

SKILLS

Programming Languages: Python, Java, JavaScript, C, C++, C#, Go, Clojure, Scala, R, MATLAB
Data Science and Analytics: PyTorch, TensorFlow, Scikit-Learn, Transformers, Pandas, Matplotlib, Numpy
Data: MySQL, MongoDB, Redis, Hadoop, Hive, Spark, Apache Graph, Kafka, Snowflake, S3, Databricks
Technologies: K8s, Airflow, Docker, Linux, AWS, Apache, Spring, Tomcat, FastAPI, Lucene, Jenkins, Latex
Machine Learning: Few-Shot, Self-Supervised, Semi-Supervised, NLP, BERT, CV, ResNet, Embedding