

# Yali Bian

☎ (540)998-0256 • ✉ yali@vt.edu • 🌐 www.yalibian.com • in yalibian • 📱 yalibian

Ph.D. Student in Computer Science, Virginia Tech, Blacksburg, VA

Driven computer science Ph.D. candidate with 7+ years of published research experience in human-AI collaboration, machine learning, interactive data visualization. Well versed with the software development life cycle from 3+ years of industry work experience in full-stack web development, data analytics, and data mining. Looking to leverage problem solving, collaboration, and communication to contribute to cutting-edge research at a dynamic company.

## EDUCATION

- **Ph.D. in Computer Science and Application** – Virginia Tech (GPA: 3.97/4.00) Aug. 2016 – Dec. 2021
- **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
- **B.Eng. in Computer Science** – Harbin Institute of Technology (GPA: 88.40/100) Sep. 2009 – Jun. 2013

## RESEARCH EXPERIENCE

- **Graduate Research Assistant – Virginia Tech** Aug. 2016 – Now
  - *Human-AI Teaming with Interactive Dimensionality Reduction and Deep Learning*
    - Designed an interactive and interpretable visual analytics framework to reinforce the collaboration between humans and deep learning
    - Implemented an interactive BERT adaptation system with human-in-the-loop for exploratory analysis of textual data
    - Led a team of 3 graduate students on an online system to incorporate expert input into pretrained DL model to interactively assess the quality and health of specialty crops
    - Published 3 papers on Interactive Deep Learning and Visual Analytics in ACM IUI, IEEE VIS
    - Using PyTorch, Transformers, Scikit Learn, BERT/ResNet/MDS/T-SNE, D3.js, React.js, Socket.io, Express.js, ZeroMQ
- **Graduate Research Assistant – Zhejiang University** Jul. 2013 – Mar. 2016
  - *Visual Text Analytics with Topic Modeling for Digital Library*
    - Designed and developed an intuitive visualization system for topic modelling of large collections of journal papers
    - Led a team of 2 developers to design a streamlined visual retrieval system to index 2.4M+ books from the CADAL digital library
    - Published 3 papers on Visual Text Analytics in JCDL, Expert Systems with Applications, Journal of biomedical informatics
    - Used Java, Spring, MySQL, Hibernate, NLTK, CoreNLP, Gensim, Lucene, Processing, D3.js, Canvas, WebGL, P5.js, Racket

## WORK EXPERIENCE

- **Software Development Engineer – Akuna Capital** Aug. 2019 – Jan. 2020
  - *Chicago, IL*
    - Led a team of 2 engineers in building a trade surveillance and compliance visualization application to effectively monitor and analyze live trading behaviours from the message streaming platform
    - Optimized the performance of message streaming platforms' capturing and analysis pipeline by 72% in Agile development environment
    - Used Kafka, Solace, ZMQ, Node.js, React.js, CI, Jenkins, Docker, asynchronous and multiprocessing methodologies
- **Ph.D. ML Research Intern – Facebook** Jun. 2018 – Aug. 2018
  - *Seattle, WA*
    - Executed training of a two-tower neural network between billions of Facebook groups and users to improve alumni group recommendation system
    - Applied data science skills to Big Data by extracting billions of alumni user-group tuples using collaborative filtering, label propagation
    - Used Hive, Spark, Presto, DataSwarm, Digraph, FBLeanner, GBDT, Word2Vec
- **Ph.D. Research Intern – NOKIA Bell Labs** Jun. 2017 – Aug. 2017
  - *Murray Hill, NJ*
    - Developed a full-stack web platform for anomaly detection, applying ML and visualization to millions of time series Telecom data
    - Built an end-to-end data sampling and processing system to render and explore millions of time points in real-time
    - Used Python, Flask, Scikit-Learn, D3.js, React.js, OneClassSVM, interactive ML, and active learning

## SKILLS

- **Programming Languages:** Python, Java, JavaScript, C/C++, Go, Lisp, Clojure, R
- **Data Science and Analytics:** PyTorch, TensorFlow, Scikit-Learn, NLTK, Pandas, Matplotlib, Numpy
- **Web Development:** Spring, Flask, Express.js/Koa.js; React.js, D3.js, Three.js
- **Databases:** Hadoop, Hive, Presto, Spark, DataSwarm, Apache Giraph, MySQL, MongoDB, Redis
- **Technologies:** Unix/Linux, AWS, Tomcat, XAMPP, Lucene, Jenkins, Git, Vim, L<sup>A</sup>T<sub>E</sub>X, Docker