Xingkun Yin

Email :yinxingk@bu.edu | Github: https://github.com/troyyxk

EDUCATION

- **Boston University** — Boston, US
  Master of Science in Computer Science — *2021.09 - 2023.01*
- **University of Toronto** — Toronto, Canada
  Honours Bachelor of Science, Computer Science with minor in Statistical Science — *2016.09 - 2021.06*
  - Full time internship from 2019.05 to 2020.08

PUBLICATION

- **Yin, X.**, Yan, Da., Almudaifer, A., Yan, S., Zhuo. Y, Forecasting Stock Prices Using Stock Correlation Graph: A Graph Convolutional Network Approach, The International Joint Conference on Neural Networks 2021 (IJCNN 2021)

PATENT

- **Yin, X., Liu, Y., Zhang, Z., "A Time Series Prediction Method for Anomaly Alerting in Distributed Databases", (in submission)**:
  - Proposed a method for anomaly detection in distributed database using time series models.
- **Liu, Y., Yin, X., Zhang, B., Liu, H., "A Highly Available and Scalable Monitoring Data Collection and Storage Solution", (in submission)** :
  - Developed a high-availability, scalable, and highly observable monitoring data collection and storage solution.

RESEARCH EXPERIENCE

- **Forecasting Stock Prices Using Stock Correlation Graph: A Graph Convolutional Network Approach (2021)**:
  - Proposed a stock price forecasting model combining Gated Recurrent Units (GRU) and Graph Convolutional Networks (GCN). The model used GCN to cluster highly related stocks and piped the information to GRU to form predictions.
  - Implement model with TensorFlow, data cleaning and preprocessing using Pandas and Numpy.
  - Trained the model using ETF and DOW data across multiple time dimensions. Achieving result of 5% accuracy improvement over baseline model.
- **Forecasting distributed database abnormality using time series models (2023)**:
  - Developed and evaluate time series models, including Informer and LSTM, predicting key distributed system figure surrounding load balancing to enhance forecasting accuracy of database anomaly.
  - Designed customized training process and loss functions based on characteristics of the distributed database.
  - Designed anomaly warning system using forcasting model.
  - Successfully deployed trained models on Agricultural Bank of China' s platform, sending predicted anomaly warning to system admins.

PROFESSIONAL EXPERIENCE

- **Agricultural Bank of China** — Shanghai, China
  Database Engineer — *2023.6 Present*
  - Designed and deployed database architecture, establishing the technology stack for the TDSQL distributed database.
  - Customized, designed, and tested the DBbridge, a data migration tool tailored for TDSQL based on the deployment experience of Agricultural Bank of China.
  - Designed the TDSQL Distributed Database AI Alert Project, integrating models such as Informer and LSTM to predict key metrics, generate anomaly alerts, and applied for patents related to these innovations.

- **Amazon Web Service, AWS** Boston, US
  Software Develop Engineer Intern *2022.05 - 2022.08*
  - Enhanced the migration recommendation solution by incorporating internal communication costs of server clusters as a parameter input into the recommendation algorithm.
  - Improved the recommendation algorithm to reduce the number of calls to the DynamoDB backend, increasing efficiency and performance.
  - Developed three new API recommendation functionalities to assist customers in migrating on-premise servers to AWS.

- **Boston University** Boston, US
  Teaching Assistant *2022.01 - 2022.05*
  - Taught DS 120 (Introduction to Data Science) tutorial session weekly to help students review the content of the week.
  - Went over topics in Linear Algebra, Calculus and Probability during office hours.
  - Grading assignments and exams.

- **Ceridian HCM, Inc.** Toronto, Canada
  Software Engineer Intern *2019.05 - 2020.08*
  - Developed and implemented functionalities for both frontend and backend components of the company's accounting platform.
  - Participated in the migration of the platform from Silverlight to modern frameworks including DojoJS, ReactJS, and ASP.NET.
  - Created an internal accounting tool to streamline financial operations.

PROJECT

- **Cache System for Flink State Access Workloads**: Spring 2022
  - Designed and implemented a cache layer between Flink operator and embedded RocksDB, supporting user-customized policies.
  - Reduced processing time by 50% under random synthetic workloads and by 15% with Google cluster data.
- **Raft Based Fault Tolerant Distributed KV Storage**: Spring 2022
  - Implemented Raft concensus algorithm using Golang
  - Built a fault-tolerant distributed key-value storage system on top of Raft, ensuring resilience against minority failures.
  - Implemented sharding to guarantee linearizability while dynamically moving data shards across workers.
- **SimpleDB**: Fall 2021
  - Implemented B+ Tree indexing with bulk loading to enhance database query performance.
  - Developed a System-R based query optimizer that leverages system statistics for efficient query planning and execution.

SKILL SUMMARY

- **Programming Language**:
  - Python, Java, Golang, C Language, C++, JavaScript, HTML, CSS, SQL, C#, Haskell, R
- **Frameworks and Tools**:
  - TensorFlow, Keras, PyTorch, Flink, SQL Server, MySQL, NodeJS, ReactJS, MongoDB, GIT, DojoJS, MVC.NET