Ritwik Kumar Singh

ritwiksingh39@gmail.com | +1- 2246164078 | linkedin.com/in/ritwiksingh28 | github.com/ritwik-singh

Education

Northwestern University, Illinois **Sept 2021 - Dec 2022 (Expected)**

Master of Science in Computer Science and Minor in Engineering Management (GPA – 3.91/4.00)

Coursework: Algorithms, Machine Learning, Statistics, Databases, Computer Networking, Web Development.

Experience

Software Engineer Intern | inLumon **Jun 2022 – Present**

- Built a multithreaded data network system connecting MySQL server to AWS using Turbodbc and Apache arrow, to construct ETL Pipeline that could read over 500M rows of SQL data per second.
- Engineered a multi-processing data analysis framework in Vaex, Numba, and CUDA to visualize application behavior and provide insights to minimize application process by 30%.
- Designed MLFlow-driven pipeline leveraging a BERT+ AST model (accuracy 0.75) to reduce code review time by 20%.

Tech Stack: Python, gRPC, PySpark, Vaex, Numba, CUDA.

Research Intern | Habits Lab – Northwestern University **Jan 2021 – Aug 2021**

- Devised a data tool capable of converting videos into frames, removing glitches, and producing optical flow images to develop a dataset for 1000+ Mukbang videos.
- Generated hand trajectory using media pipe and performed time series imputation applying Kalman filter with 70% confidence.
- Prototyped an LSTM based Seq2Seq model to detect eating activity with an accuracy of 75%.

Tech Stack: Python, TensorFlow, Scikit-Learn.

Co-Founder and Technical Lead | Examly App [App Link] **Apr 2020 – Oct 2020**

- Coordinated and created a flutter android app to simulate an examination experience for students garnering over 4K+ downloads.
- Introduced caching to decrease network overhead by 40%. Consumed Provider framework to implement data state management.
- Boosted user engagement and retention by introducing Realtime dashboard visualizing user performance.

Tech Stack: Flutter, Firebase, NoSQL, Dart.

Projects

Trading bot [GitHub] **Aug 2022 – Oct 2022**

- Prototyped asynchronous data feeder to fetch historical and livestream data from Binance with a latency of 4ms.
- Trained an ensemble ML model (Gradient Boost, KNN, Random Forest) to predict buy/sell signals over past OLHCV data.
- Designed a gRPC based ML microservice to reduce model inference latency to 75% for live data.
- Constructed a Grafana dashboard (Watchdog system) in Python to monitor latency of backend server and a system logger.

Tech Stack: Python, React, gRPC.

Game server [GitHub] **May 2022 – Jun 2022**

- Implemented an asynchronous architecture server with multi-threading to handle multiple clients in Asio C++ module.
- Incorporated a thread safe message dequeue for client communication and to lower server load by 50%.

Tech Stack: C++.

Publications

- "A Case-Based Approach to Data-to-Text Generation" Proceedings of the 29th ICCBR 2021 Spain [Paper Link]
- "Accident Detection using Time-Distributed Model in Videos" Proceedings of the 5th ICICT 2020 London [Paper Link]

Skills

- Languages: C++, Python, JavaScript, SQL.
- Tools and Frameworks: Flask, TensorFlow, React, Git, Linux.

Achievements

- Ranked 63 for ACM-ICPC regionals in 2018 & 84 for ACM-ICPC regionals in 2019.
- Ranked 3rd in NU WildHack (Hackathon) among a pool of 50+ teams.