

Faisal Baig

STUDENT · FULL STACK · ML

1219 Wetlands Ct. 30044 - Atlanta GA

☎ 678-770-5117 | ✉ baig.faisal@outlook.com | 🌐 www.faisalbaig.com | 📱 baig-faisal

Summary

Final Year BS in CS at Georgia Tech with a specialization in Systems & Networks. Concurrently pursuing a BS/MS in Computer Science at Georgia Tech with specialization in Machine Learning. Interested in developing technology that solves real-world problems.

Education

Georgia Institute of Technology

B.S. IN COMPUTER SCIENCE

- Zell Miller Scholar, Dean's List

Atlanta, Georgia

August 2018 – May 2023

Georgia Institute of Technology

M.S. IN COMPUTER SCIENCE

Atlanta, Georgia

January 2023 – May 2024

Experience

CertiK

New York City, New York

PLATFORM ENGINEERING INTERN

September 2022 – December 2022

- Developed automated pipeline for creating & updating monitors of current data platform using Github Actions/Terraform. Automated migration of 320+ existing agents/monitors to infra-as-code (MonteCarlo/Datadog/Github Actions/Terraform)
- Automated 15+ classifier models using Databricks/MLflow/Snowflake. Integrated Datadog into Databricks for logging (Datadog/Spark)
- Investigated economic impact and proposed optimization for high consumption or high runtime Snowpipe loads & Snowflake queries
- Deployed Terraform-based infrastructure to dev, staging, & live environments using Github Actions CI/CD pipelines for public and internal APIs and set up logging and monitoring (ECS, S3, Docker, Airflow, Terraform)
- Policy creation, environment setup, and permissions management for ML Infrastructure and Snowflake

Equifax

Atlanta, Georgia

SOFTWARE ENGINEERING INTERN

May 2022 – August 2022

- Endpoint development and UI design of B2B portal using Springboot and Canva for the MVP and beyond of a new project (Java/BigQuery)
- Developed transformations to process novel data format in an active data pipeline (Dataflow/Apache Beam)
- Refactor & simplify logic of existing transformations for Beam ETL pipeline processing millions of records (Java/Apache Beam/PostgreSQL)
- Interface with internal development teams on large-scale, multi-tier product development

Research

CT Scan and Medical Imaging

Atlanta, Georgia

AI-BASED DISCOVERY AND INNOVATION GROUP | DEV TEAM

August 2021 – May 2022

- Developed full-stack app using Azure Blob Storage & Node.js serving REST API to allow categorization of CT Scans by radiologists at Emory University
- Test & log metrics for ML models trained on novel data to allow for categorization, localization, and gradation of different conditions lung conditions

COVID-19 Prediction

Atlanta, Georgia

AI-BASED DISCOVERY & INNOVATION GROUP | MEMBER

January 2021 – May 2021

- Created ML models to analyze spatiotemporal factors relevant in the spread of COVID-19, present findings & methodology to research group
- Documented recent & relevant papers on transmission of COVID-19 to improve models by identifying potential helpful data & features. Compare predictive capacity among multiple models types and present findings.

Projects

ENSEMBLE LEARNING TRADING ALGORITHM | PYTORCH, SCIKIT-LEARN, REST

- Ensembles multiple supervised methods to maximize profit on historical trading data. Relevant features & stock groupings using K-Means w/DTW

TECHNOTE | MONGODB, EXPRESS.JS, VUE.JS, NODE.JS, GOOGLE CLOUD FUNCTIONS

- Knowledge management tool helping GaTech students organize class notes & files on web scraped class & professor data. Supports a user-specific dynamically updated autocomplete feature

Skills

Languages & Frameworks: Python, Java, PyTorch, PySpark, Springboot, Apache Beam, React.js, Express.js, Django, SQL, NumPy, scikit-learn, Pandas

Dev Tools & Libraries: Git, GCP, AWS(ECS,S3), AzureML, Postman, Docker, Postgresql, CI/CD, Airflow, DevOps, Databricks, MLOps, MLflow, Terraform, Bash, Linux

Coursework: Data Structures Algorithms, Algorithmic Analysis, Computer Organization, Operating Systems, Software Design, Databases, Systems & Networks, Statistics, (MSCS) Machine Learning, (MSCS) Computer Vision

Certifications: Google Analytics Individual, Google Cloud Digital Leader

Clubs: GT IEEE, Datascience Club, ML Club, Hytech Racing