# Faisal Baig

STUDENT · FULL STACK · ML

*1219 Wetlands Ct. 30044 - Atlanta GA*

678-770-5117  |  baig.faisal@outlook.com  |  www.faisalbaig.com  |  baig-faisal

## Education

**Georgia Institute of Technology**                                                                 *Atlanta, Georgia*

B.S. IN COMPUTER SCIENCE – ZELL MILLER SCHOLAR + DEANS LIST                                          *Grad May 2023*

**Georgia Institute of Technology**                                                                 *Atlanta, Georgia*

M.S. IN COMPUTER SCIENCE                                                                             *Grad May 2024*

## Experience

**Balyasny Asset Manangement**                                                                      *Chicago, Illinois*

SOFTWARE ENGINEERING INTERN                                                                          *June 2023 – August 2023*

- Build systems and data platforms that process and prepare unique data sets for research and generate insights for our investment teams.

**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 and secured 15+ classifier models using Databricks/MLflow/Snowflake, implementing data encryption, access controls, and secure logging with Datadog integration.
- 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, & IAM for ML Infrastructure and data pipelines (Databricks, Github, 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, using data protection best practices(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

**Georgia Tech AI-Based Research and Discovery Lab**                                                *Atlanta, Georgia*

DEVELOPER INTERN                                                                                     *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

RESEARCH INTERN                                                                                      *January 2021 - May 2021*

- Developed time-series model to study spatiotemporal factors influencing COVID-19 spread; utilized time-series cross-validation for optimization
- Reviewed recent research papers on COVID-19 transmission to identify potential new features, aiming to augment datasets and enhance the training of our predictive models, compare predictive capacity among multiple models types, & present findings.

## Projects

PHISHING DETECTOR | PYTHON, PYTORCH, NLTK, SCIKIT-LEARN, REST

- Analyze email content and metadata w/NLTK, providing real-time flagging of suspicious emails and alerting users to potential phishing attempts

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, PASSPORT.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 and employs encryption and access controls to ensure data confidentiality and privacy.

## Skills

**Languages & Frameworks:** Python, Java, C++, PyTorch, PySpark, Springboot, Apache Beam, pwntools, Express.js, Django, SQL, scikit-learn, OWASP, TCP/IP

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

**Coursework:** Data Structures  Algorithms, Algorithmic Analysis, Computer Organization, Operating Systems, Software Design, Databases, Systems & Networks, Statistics, (MS) Machine Learning, (MS) Computer Vision - **Enrolled:** (MS) Network Security, (MS) High Performance Computing, (MS) Natural Language

**Certifications:** Google Analytics Individual, Google Cloud Digital Leader