# MICHAEL TANG

📞 +852 98160130   ✉️ michaeltang6649@gmail.com   🔗 mike6649.github.io   💼 @mike6649   🐙 @mike6649

## EDUCATION

**Georgia Institute of Technology**   Atlanta, GA
*MS Computational Science and Engineering*   *Aug 2022 – Dec 2023*
- GPA: 4.0/4.0
- Project: Improving the scalability of machine learning model serving in soft real-time systems

**University of Hong Kong**   Hong Kong
*BS Physics*   *Sep 2016 – Aug 2021*
- GPA 3.76/4.0
- Thesis: Using maximum likelihood estimation to measure electron charge flip rate at the ATLAS detector

## EXPERIENCE

**General Motors Financial**   Fort Worth, TX
*Data Scientist Intern*   *May 2023 – Aug 2023*
- Enhanced customer risk scoring model by implementing **SMOTE** oversampling technique, achieving 5% higher AUCROC.
- Conducted research to perform feature selection for customer scoring model with **PySpark**, **XGBoost** and **SHAP**.
- Effectively communicated research findings to both external model auditing groups and internal technical teams.

**Georgia Institute of Technology**   Atlanta, GA
*Graduate Research Assistant*   *Jan 2023 – Present*
- Spearhead **data engineering** efforts for campus meter data infrastructure, encompassing approximately 10000 meters.
- Implement CI/CD and data pipelines utilizing **Docker** and **Apache Airflow** to streamline ETL processes.
- Lead team of 8 in development of dashboards in **React** and **Flask**, providing intuitive data visualizations to stakeholders.
- Assist in geospatial model development for simulating energy trends within Atlanta region using **Postgres & PostGIS**.

**Meta Air Labs Ltd.**   Hong Kong
*Full Stack Developer*   *Jun 2021 – Aug 2022*
- Led team of 4 to design and build high-throughput cryptocurrency transaction systems with **Kafka** and **Python** on **AWS**, supporting a sustained load of 2,000 transactions per second.
- Planned and executed performance tests with JMeter to identify and address underlying system-wide bottlenecks.
- Utilized Grafana to create real-time dashboards, simplifying system monitoring and enhancing user insights.

**Dah Sing Bank Ltd.**   Hong Kong
*Fintech Intern*   *Jul 2019 – Jun 2020*
- Assisted in development of new E-Banking system using **Java Spring Boot** and **Angular** with Agile methodologies.
- Implemented CI/CD pipelines in **Jenkins** to automate release builds, enabling rapid software iterations.
- Undertook critical architectural tasks, including infrastructure design, firewall configuration and data migrations.

## PROJECTS

**Federated Learning of Generative Large Language Models via Knowledge Distillation**   *Oct 2023*
- Enable knowledge distillation of generative LLMs such as Llama-2, GPT-2 within Federated Learning regime
- Allow decentralized and resource-constrained devices to effectively produce a strong language model without loss of privacy

**Power Map – Hackathon Winner** | *Python, Decision Trees, GCP*   *Feb 2023*
- Developed predictive model using **LGBM** tree boosting model on weather and building data, achieving $R^2$ of 0.8.
- Implemented model deployment using GCP onto a website, enabling energy usage predictions of buildings across the US.

**Blockchain Transaction Network Visualization and Anomaly Detection** | *Neo4J, React*   *Nov 2022*
- Transformed the Ethereum transaction network into a graph utilizing **Neo4J**, facilitating comprehensive visualization.
- Employed **Graph Neural Networks** to identify potentially malicious actors within the blockchain network.
- Visualized the findings using **React** and **D3.js** for intuitive and informative representations.

## SKILLS

**Machine Learning:** Probability & Statistics, Data Visualization, Model Serving, LLMs, Deep Learning
**Programming:** Python, SQL, PyTorch, TensorFlow, Scikit-learn, Numpy, Matplotlib, C++, NodeJS
**Platforms & Applications:** Linux, Docker, CI/CD, AWS, Spark, Kafka, RabbitMQ, REST API, SAS, Tableau, Git