

# Suvigya Vijay

✉ suvigyavijay@gmail.com | ☎ +1 716 303 9681 | 🌐 github.com/suvigyavijay/ | in linkedin.com/in/suvigyavijay

## EDUCATION

---

- **University at Buffalo** The State University of New York  
*M.S.(Research) in Computer Science and Engineering* Buffalo, New York  
Aug 2023 – Dec 2024
- **Birla Institute of Technology and Science(BITS), Pilani**  
*B.E.(Hons) in Electronics and Instrumentation & M.Sc.(Hons) in Economics* Pilani, India  
Aug. 2015 – May. 2020

## EXPERIENCE

---

- **Flipkart**  
*Data Scientist/Machine Learning Engineer* Jul 2020 - Jun 2023
  - Molecule:** Machine Learning Platform for data science teams.
    - \* Implemented an autoscaling workflow orchestrator with multi-language support, UI, caching, and debugging features. Enabled standardized workflows, collaboration, swift experimentation, and productionization.
    - \* Developed the workflow orchestrator and autoscaler in Python with ZeroMQ using multi-processing module, connected it to a MongoDB replica setup for session storage, and created the UI using Flask for the backend and vanilla JavaScript for the frontend. The system was hosted on GCP and featured fault-tolerant capabilities.
    - \* Boosted team efficiency significantly, resulting in over 30,000 experiments conducted and 50+ models in production. Paper submitted to SIGMOD 2024. [\[link\]](#)
  - 14A Model:** Auto-tuned multi-model approach for low-lead demand forecasting.
    - \* Designed a generic approach for sales forecasting with modules and models to address out-of-stock, outliers, future events, seasonality, recency bias, and COVID-related supply constraints.
    - \* Utilized PySpark and tidyverse packages in R for fetching and wrangling data, and classical time series modeling, along with TensorFlow for modeling Seq2Seq and transformer-based models.
    - \* Achieved an average 17% improvement in wMAPE across 25 product categories, saving 85 million USD annually on the e-commerce platform.*Data Science Intern* May 2019 - Jun 2020
  - New Product Launches:** Intervention-based model for cold-start sales forecasting for new product launches.
    - \* Launched a UI-based platform to forecast sales for new products, by arriving at k-nearest product substitutes based on various factors such as product attributes, price, brand, etc.
    - \* Used Pandas and PySpark to write serverless functions for data processing and feature generation, and connected it to Bootstrap UI via Fast API with authentication.
- **Performance Co-Pilot**  
*Google Summer of Code (GSoC)* May 2017 - Aug 2017
  - Vector:** Netflix OSS' project, an on-host performance monitoring framework.
    - \* Enhanced user functionality by adding a custom widget feature that enables the user to add metrics with tweaks to a new blank widget, written in AngularJS.

## PROJECTS

---

- **Distributed Storage System:** Designed and implemented a robust key-value store as a part of a course project in Go, employing Chandy-Lamport and Raft consensus algorithms to ensure consistency and fault tolerance
- **Chat Application:** Developed a text chat app in C++, featuring login functionality to access the server, list connected clients, and facilitate unicast and broadcast messaging. Implemented message buffering and robust error handling.

## CERTIFICATIONS

---

- **Certified Kubernetes Developer** May 2022
- **ColumbiaX Machine Learning** May 2021
- **Data Scientist with Python** Feb 2018

## TECHNICAL SKILLS

---

- **Programming Languages:** Python, R, JavaScript, Go, Java, C/C++; Also familiar with: SQL, Typescript, Bash
- **Frameworks and Libraries:** Apache Spark, Hadoop, Kafka, Tensorflow, Pytorch, Ionic, Angular, Flask, Tableau