

# SOORAJ BHARADWAJ

[soorajbh@buffalo.edu](mailto:soorajbh@buffalo.edu) | (716) 426 8555 | [linkedin.com/in/sooraj-bharadwaj](https://linkedin.com/in/sooraj-bharadwaj) | [github.com/surajbharadwaj17](https://github.com/surajbharadwaj17) | [soorajbharadwaj.com](https://soorajbharadwaj.com)

---

## EDUCATION

**Masters: Data Science & Applications**, University at Buffalo, The State University of New York, Buffalo  
Graduation date: January 2023, GPA 3.75/4.

**Bachelor of Engineering: Computer & Information Science**, R.V. College of Engineering, Bangalore, India  
Graduation date: June 2018, GPA 8.94/10.

## PROFESSIONAL EXPERIENCE

**Software Engineer Intern, Data & ML Engineering, Johnson Controls, Milwaukee, WI:** June 2022 - August 2022

- Designed and developed a framework as a micro service to observe machine learning models in production and also analyze their performance using various statistical tests.
- Devised the framework such that data scientists/engineers were able to observe the performance of the model in production environment using a single button click.
- Contributed in building a microservice to re-train a deployed model on demand.

**Software Engineer II, Aruba Networks (Hewlett Packard Enterprise), Bangalore, India:** April 2021 - August 2021

- Collaborated with an agile team to launch an in-house anomaly detection system to assess end-to-end data and application reliability clusters hosted on AWS using Pyspark, Apache oozie and Cassandra.
- Analyzed the collected time series data and real time insights were visualized using REST APIs and Grafana in the form of a dashboard which helped the QA team to ensure data accuracy and availability by increasing failure detection speed by 100% (4hrs+ to ~real time)

**Software Engineer I, Aruba Networks (Hewlett Packard Enterprise), Bangalore, India:** July 2018 - March 2021

- Built and maintained monitoring applications on Hadoop for data integrity checks using python which helped the QA team to monitor the production failures in real time using an alerting mechanism.
- Orchestrated CI/CD activities to deploy on AWS using automated python scripts, git and deployment tools like Jenkins, Ansible. Developed pipeline was used by dev and QA team for automated rollout of more frequent releases with ease.

## PROJECTS

**Tweeper - Real time Twitter Data Analysis** (*Python, Airflow, NLP, Docker, Postgres, FastAPI*) [\[Github\]](#)

- Designed and implemented a microservice to ingest real time data from Twitter using REST APIs and use natural language processing techniques to display the polarity count, frequency distribution and geographical distribution of the tweets related to a given keyword.

**Soccer Live Updates** (*Python, Airflow, Docker, Postgres, Streamlit*) [\[Github\]](#)

- Created an app with a streaming and batch data ingestion pipeline to ingest updates across all soccer competitions and teams using Airflow and visualize the summary statistics and performance charts across competitions.

**Human Activity Recognition with Smartphones** (*Python, Machine Learning*) [\[Github\]](#)

- Developed 2 prototype applications for recognition of human activity using Random Forest and Logistic Regression models. Prototypes were able to recognize the activity with more than 95% accuracy.
- Created a light-weight python module to automatically create, train and predict based on model type.

## SKILLS & TOOLS

- Data structures & algorithms
- Machine Learning
- Python/scikit-learn
- Cloud computing
- Data visualization
- Data warehouse/ETL
- Agile & Devops
- Microservices (Docker/Kubernetes)
- CI/CD
- RDBMS & SQL (Postgres, Cassandra)
- OOPS & Design patterns
- Version control/Git
- REST API (FastAPI)
- System design
- Distributed systems
- Event streaming/Kafka
- Workflow management/Airflow
- Unix scripting