

DARSHEE BHUPENDRA MACHHAR

☎ 347-941-7854 ✉ darshee.m013@gmail.com 🔗 [linkedin.com/in/darshee-machhar](https://www.linkedin.com/in/darshee-machhar) 📁 [Portfolio](#) 🐙 github.com/Darshee-m

Education

Rutgers- The State University of New Jersey

Sep. 2022 – May 2024

Masters in Computer Science, GPA: 4/4, Concentration in Machine Learning & AI.

New Brunswick, NJ

K.J. Somaiya College of Engineering

Aug. 2016 – Oct. 2020

Bachelors in Computer Engineering, GPA: 3.6/4, IoT Research Assistant.

Mumbai, India

Won prestigious hackathons including WicHacks 2021, IBM Hack Challenge 2019, and Mumbai Hackathon 2019.

Experience

Rutgers Institute for Health | Data Analyst | New Brunswick, NJ

Nov. 2022 – Present

- Spearheaded research on predictive cognitive disorder detection by building a pipeline to collect keystroke and accelerometer data from smartphones, achieving a **RMSE score of 0.18** using **recurrent neural networks**.
- Achieved **86% accuracy** in predictive time series ECG data analysis through advanced machine learning algorithms.

Chera Health | Software and Data Intern

Jul. 2023 – Sept. 2023

- Built an **Android and iOS** personalized food diary mobile app aiding 500 users with eating disorders in accessing tailored diet plans and specialized licensed therapists.
- Designed in-app forms to handle user feedback, gather daily health checks and raise support tickets using **Java** and **C#**.
- Leveraged users' diet and mood data to optimize patient care plans and identify mental health triggers with an **accuracy rate of 78%** by deploying diurnal analysis using machine learning algorithms like **Support Vector Machines**.

ZS Associates | Data Engineer

Aug. 2020 – Aug. 2022

- Adept at utilizing **Jira Scrum Agile** development for daily workflows.
- Utilized cloud microservices with **Flask & AWS Lambda** to streamline solutions, **reducing response times by 50%**.
- **Dockerized** applications firm-wide, enhancing deployment speed, scalability & quicker migrations.
- Maintained responsive **100 TB large big data systems** and **ETL pipelines** for data warehousing using tools like **Oracle and Hadoop**.
- Reduced overhead by 45% by leveraging **GraphQL APIs** for search & data discovery purposes.
- Implemented **RESTful APIs** using **Python and Flask** in conjunction with **Postgres** and **MongoDB** of large critical software programs while leading code reviews and providing mentorship to junior developers.
- Employed continuous integration and deployment (**CI/CD**) pipelines and **Machine Learning** to automate build and deployment processes, reducing release cycle times by 20%.

Colgate Palmolive | Security Engineer

Jun. 2019 – Aug. 2019

- Developed and maintained a software scan portal using **Django** to validate software installation by deploying **VirusTotal APIs** which resulted in a **30% reduction** in security tickets.
- Monitored safety incidents using vulnerability scanners like **Qualys** and **OWASP ZAP** and mitigated security issues using the a SOAR (Security Orchestration, Automation, and Response) platform like **Siemplify**.

Projects

Metadata Marketplace | Python, Flask, Neo4j, AWS, Azure, GraphQL, Data Catalogs, Control-M

- Engineered a robust ETL pipeline leveraging Airflow to generate, enrich, and publish metadata from structured data sources to a knowledge graph, facilitating critical insights for **over 1000 employees daily**.
- Implemented data cataloging techniques to organize and classify metadata, enhancing accessibility for downstream applications. Collaborated with stakeholders to define **250** metadata standards and ontologies.
- Designed and implemented monitoring and alerting mechanisms within **Kibana** to track metadata quality and performance metrics, enabling proactive maintenance and troubleshooting with **96% uptime**.

Snowflake ETL Pipeline | Snowflake, Azure, Airflow, PySpark

- Developed end-to-end ETL pipeline leveraging **Snowflake's cloud-native** data warehouse capabilities with Azure.
- Implemented automated data processing workflows using **Apache Airflow**, ensuring robust scheduling and monitoring.
- Leveraged **PySpark** for complex data transformations and processing tasks, optimizing performance by **39%**.

Medical Product Sentiment Analysis | NLP, Opinion Mining, BERT, Python, TensorFlow

- Engineered a sentiment analysis pipeline using NLP techniques, including **BERT and Opinion Mining**, to analyze physician interview transcripts regarding medical products.
- Implemented **TensorFlow** models to perform sentiment analysis on the transcripts, achieving an accuracy rate of **83%**.

Skills

Languages: Python, JavaScript, TypeScript, Shell, Java, C, C++, C#, R, Go, Solidity, SQL

Clouds/Platforms: AWS, Azure, Google Cloud, IBM Cloud, Windows, Unix, Linux, MacOS, AndroidStudio

Database/Tools: MySQL, Oracle, Postgres, MongoDB, Neo4j, Elasticsearch, Docker, Kubernetes, Airflow, Git, Jira, Hadoop

Frameworks: Flask, Django, Nodejs, React, Angular, Vue.js, Next.js, Express, Spring Boot, Crow, Pandas, NumPy

Generative AI Technologies: LLama2, Mistral, OpenAI, Langchain, ChromaDB, Pinecone