

VARSHA HINDUPUR

(857)-693-9108 | hindupur.v@northeastern.edu | [linkedin.com/in/varsha-hindupur-4aaa83175](https://www.linkedin.com/in/varsha-hindupur-4aaa83175) | ([website](https://github.com/varshahindupur09)) [varshahindupur09.github.io](https://github.com/varshahindupur09)

EDUCATION

NORTHEASTERN UNIVERSITY

Master of Science in Information Systems, ML Engineering & Advance Data Science

[Sept 2022 - July 2024]

Boston, MA

UNIVERSITY OF MUMBAI

Bachelor of Engineering in Information Technology

[May 2012 - May 2016]

Mumbai, India

TECHNICAL SKILLS

Programming: Python, SQL, DBT, Spark, Java (SpringBoot), R, C/C++, JavaScript (Angular, React, Node), Scala, Julia
AI, ML & Data Science: Data Modeling, Deep Learning, Regression & Classification, LLM, Hugging Face, NLP, Scikit-Learn, AutoML, Tensorflow, Keras, PyTorch, SAS, Hadoop, Hive, FlaskAPI, FastAPI, Terraform, Teradata, Selenium
Cloud: AWS, Azure, Kafka [Confluent], GitHub Actions, Docker, Snowflake, Databricks, Airflow, Astronomer
Certification: Microsoft Azure Fundamentals (AZ-900)

WORK EXPERIENCE

DATA ANALYSIS CO-OP

[May 2023 – Dec 2023]

Novelis - An Aditya Birla Company | Tech Stack: Hugging Face, Arize, Azure, Power BI, Salesforce, Snowflake

Atlanta, GA

- Executed predictive modeling using Hugging Face & verified equipment failure in diverse manufacturing plants
- Leveraged Azure to automate model rollout time from **30 mins to 3 hours** creating end-to-end MLOps pipeline
- Contributed to enhancing "EagleGPT," customized **Large Language Model (LLM)** on ChatGPT. Improved conversational AI with domain-specific dialogues, enhanced model responsiveness, and assisted in developing the backend system
- Established, monitored, and improved KPIs calculating metrics of budget allocation, project completion, IP generation, contributing to **15% faster project latency findings** and **20% budget optimization** within Snowflake
- Automated manual search tasks in single Python Django web app, **maximizing searching efficiency and saving 200+ hours**. Utilized Docker and deployed on Azure App Service with Azure DevOps and GitHub Actions
- Extra-Efforts:** Completed building **35 out of 68** web-app components (**51.47%**) integrating React, FastAPI, PostgreSQL, Alembic & Postman **within short span of 3 weeks** working collaboratively in team of **4** members

GRADUATE RESEARCH ASSISTANT (MACHINE LEARNING)

[Feb 2023 - May 2023]

Northeastern University | Tech Stack: Computer Vision, Image Recognition & Engineering, AWS

Boston, MA

- Performed image analysis, feature engineering, & computed **20** photo attributes by leveraging OpenCV's GrabCut & Super-Pixel libraries; generating output in **30 seconds** for **1000+ records** in **50** chosen brands for Marketing Research Academia
- Enhanced identity verification, customer profiling, fraud detection via Deepface ML model via image inputs on AWS SageMaker, leading to **25% boost in precision, recall, F1 score** for age, race, gender, emotion extraction

DATA ENGINEER

[Dec 2020 - Aug 2022]

UST Global (Xpanxion) | Tech Stack: Kafka, NiFi, Restful APIs, Airflow DAGs, GCP, Docker, Snowflake

Pune, India

- Led and Architected ETL data pipelines via Apache Kafka, NiFi, Snowflake accelerating data influx for over **50,000** records by **25 times**, cutting daily execution time by **6 hours** in **3-month** timeframe for **Credit Reporting Company client**
- Engineered APIs for Collibra & Salesforce data integration via Databricks, optimizing data processing for reusability across six different use-cases; orchestrated daily processes using Airflow's DAGs
- Elevated** model accuracy by **30%** with GCP's TensorFlow & AutoML through improving hyperparameter tuning, data augmentation, ensemble learning, batch normalization, pruning, and learning rate scheduling

SENIOR SOFTWARE ENGINEER ANALYST

[Mar 2019 - Dec 2020]

Accenture | Tech Stack: UI/UX, ETL, Hadoop, Kafka, NiFi, Restful APIs, Airflow DAGs, AWS

Mumbai, India

- Re-architect multi-page app with product managers, optimizing user experience, boosting **\$1M yearly revenue**
- Championed creation of in-house UI library, building reusable components & saving **125+ hours per month**
- Revamped downstream Hadoop data preparation, achieving **20% performance boost** and slashing creation time from **12 hour to 1 hour**, making the process significantly faster with performance-tuned ETL jobs in Teradata in Agile Team
- Configured ETL Pipeline - AWS Glue, S3, Redshift achieving **30% reduction** in data processing time in **3 weeks**

SYSTEM ENGINEER

[Mar 2017 - Mar 2019]

Infosys | Tech Stack: PySpark, SQL, Angular, JavaScript, Automation Testing | Java

Pune, India

- Using PySpark (queries, views, stored procedures, advanced windows functions) generated 100+ reports with prompting **75% reporting efficacy for banking client**
- Created online banking, accounts, payment processing features via AngularJS (frontend) & Flask APIs (backend)
- Built testing framework with **30+ distinct cases** using Java, Selenium, TestNG, Karate, RestAssured for CI/CD pipeline

BUSINESS INTELLIGENCE INTERN

[Nov 2016 - Jan 2017]

Rendered Ideas | Tech Stack: Sentimental Analysis, Microsoft Excel, SAS, ETL

Mumbai, India

- Revamped player engagement using **SAS** analytics, cutting ETL time by **40 hours**, driving **20% revenue increased** through predictive modeling, and orchestrating complete BI life-cycle
- Innovatively optimized existing sentimental analysis procedures by dataset construction via feedbacks, leading to **10-fold surge** in engaging **game components** & implementation of cutting-edge **marketing strategies**

ACADEMIC PROJECTS

AIRCAST – MLOPS | ML Engineering | Data as a Service | Machine Learning as a Service | AWS

[Mar 2023 - Apr 2023]

- Crafted **Long Term Short Memory** with (**RMSE of 0.8**) to forecast air quality & host it as decoupled **MLaaS** on **EC2**
- Built end-to-end software with frontend, backend, dashboards using Python **Streamlit**, **Docker**, **Airflow DAGs**, **NGINX**, **PostgreSQL**, **Postman**, **FastAPI**, **AWS CloudWatch** and **AWS RedShift**