

VARSHA HINDUPUR

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

EDUCATION

Northeastern University

Master of Science in Information Systems, Concentration: Data Science & Machine Learning Operations

Boston, MA

Sept 2022 - May 2024

UNIVERSITY OF MUMBAI

Bachelor of Engineer in Information Systems

Mumbai, MA

May 2016

SKILLS

- Programming Languages: Python, R, C/C++, Spark, SAS, Julia, Shell Scripting, Scala, SpringBoot, Java, JavaScript, Angular JS
- Databases: SQL, MySQL, MongoDB, Cassandra, Redis, Neo4j, Google BigQuery, Azure Data Studio, SSMS, Hive, Teradata
- Cloud: Docker, Kubernetes, AWS (S3, EC2, EMR, Glue), GCP Dataflow, Azure Data Factory, Apache Spark, Snowflake
- DevOps: Git/GitHub, CI/CD, Terraform, CloudFormation, Octopus Deploy, Ansible
- Data Pipeline Orchestration: Apache Airflow, Apache Luigi, AWS Step Functions, Google Cloud Composer
- Machine Learning: Linear & Logistic Regression, Bayesian Statistics, Decision Trees, K-Means Clustering, Multivariate Analysis

WORK EXPERIENCE

NOVELIS - AN ADITYA BIRLA COMPANY

Kennesaw, GA

R&D, Ecosystems Coordinator & Machine Learning Engineering Co-op for Novelis's R&D facility

May 2023 - Present

- Spearheaded data-centric strategies to optimize business, utilizing Salesforce for progress tracking & generating Power BI reports
- Conducting cross-functional R&D ecosystem meetings, wielding expertise in data collection & analysis from conferences, industry events, & market research to drive business growth, overcome obstacles, and identify emerging trends consistently
- Implemented a CNN-based image recognition system to automate facility's Quality Control Processes by training ML model to identify defects in manufactured product, reducing manual inspection time by 30% and enhancing overall product quality

NORTHEASTERN UNIVERSITY

Boston, MA

Graduate Data Science Research Assistant for Marketing Research Group

February 2023 - May 2023

- Employed advanced CV algorithms - GrabCut & Superpixel, to perform Image Feature Engineering and created sophisticated scripts to compute 8 pivotal photo attributes in an expedited manner
- Engineered end-to-end ELT (Extract, Load, Transform) pipeline on AWS cloud, harnessing S3, Redshift, EMR, and Data Lake to accelerate deployment by 30% and thereafter combined with Snowflake to enable high-speed querying and reporting

UST GLOBAL

Pune, India

Data Analyst for Credit Reporting Agency client

December 2020 - August 2022

- Developed a Business Intelligence (BI) tool utilizing advanced machine learning algorithms - random forest, gradient boosting, and neural networks for accurate forecasting, classification, and regression analysis on credit data with 70% faster results
- Implemented configuration management for infrastructure resources across 5+ projects, leveraging Ansible, Chef, and Puppet, and this initiative led to workflow standardization, resulting in 30% increase in delivery productivity with faster project completion
- Recruited, onboarded, and mentored 8 new team members, fueling 50% augmentation in collective team capabilities

ACCENTURE

Mumbai, India

Senior Software Engineer for Client's Sales Production Data

March 2019 - December 2020

- Drafted resilient ETL (Extract, Transform, Load) data pipelines with Apache Kafka and NiFi to ingest data, & process with a 25% boost in data influx, saved 6 hours of execution daily and this was accomplished in a short span of 3 months

INFOSYS

Mumbai, India

Software Engineer for Banking Cliental

March 2017 - March 2019

- Enhanced reporting efficiency by 75% through optimized Spark SQL queries, views, and window functions, reducing reliance on redundant and complex subqueries while working on ad-hoc Data Analytics

UNIVERSITY PROJECTS

AIRCAST - ML-OPS| Tech Stack: AWS, Docker, ML, Airflow, PostgreSQL

March 2023 - April 2023

- Automated daily workflow, assimilating Apache Airflow, Docker containerization, FASTapi, Great Expectations, PostgreSQL, AWS CloudWatch, LSTM ML model, and frontend components to bring to fruition a refined & user-centric website (MaaS)

TEXT ANALYSIS - NLP | Tech Stack: Python, NLTK Libraries, PyTorch

Feb 2023 - March 2023

- Effectuated Text Analysis on 50,000 IMDB movie reviews employing NLP, Neural Networks, Computational & Semantic data modeling to find prevalent themes, & sentiments to gain insights into customer predilections & inform strategic business decisions