

Mansi Borole

mansi.borole@rutgers.edu | +1 7323745143 | [linkedin/mansi-borole](https://www.linkedin.com/in/mansi-borole) | [github/MansiGit](https://github.com/MansiGit) | mansigit.github.io

EDUCATION

Rutgers University, New Brunswick, New Jersey

Master of Science in Computer Science

Aug 2022 - May 2024

GPA: 3.91/4.0

University of Pune, India

Bachelor of Engineering in Computer Science

Aug 2015 - Jul 2019

GPA: 8.61/10

TECHNICAL SKILLS

Languages: Python, SQL, HTML, CSS, JavaScript, Shell Scripting, YAML

Cloud & Databases platforms: PostgreSQL, MySQL, MongoDB, AWS (S3, EC2, Lambda, RDS, API Gateway, CloudFormation)

Big Data Hadoop: HDFS, Apache Spark, Pyspark, Scoop, ETL, Enterprise Data Ingestion and Warehousing

Tools & Frameworks: Git, Serverless, Boto, Node.js, React, Pytest, Tableau, Grafana, Microsoft Suite, JIRA, Agile, CI/CD

PROFESSIONAL WORK EXPERIENCE

Michigan Health Information Network (MiHIN) | Software Development Engineer Intern

May 2023 - Present

- Worked on backend development of a state-sponsored project "Social Determinants of Health". Developed and deployed **AWS Lambda** functions using Python Boto library for processing, validating, and storing Electronic Health Records in **AWS RDS**.
- Created mock objects and implemented unit tests to validate the functionality of services and endpoints. Took a **test-driven design** approach to ensure code quality and improved **code coverage from 20% to 80%**.
- Optimized **PostgreSQL** database performance using indexes and views, resulting in **80% faster processing**.
- Created a PoC Project to evaluate **Serverless architecture** using AWS Lambda, SNS, CloudFormation and showcased it to the leadership.

HSBC Bank | Software Engineer (Big Data Analytics)

Jul 2019 - Aug 2022

- Built "BigData360", a real-time alerting, monitoring system and designed dashboards for 80+ Hadoop dev and prod clusters using **Grafana** and **Prometheus** by collecting metrics from **Cloudera APIs** and **Node Exporters**.
- Generated monthly cost reports to support Big Data Capacity Planning and forecasting demand. Built a self-serving data cataloging tool to analyze tenant usage metrics and trends in real time using **NiFi**, **MySQL** and groovy saving ~150 work hours.
- Virtualized control nodes by migrating critical Hadoop Master services (NameNodes, YARN, Solr, Zookeeper, Hive) from on-premise servers to ESXi virtual machines to enhance cluster efficiency. Achieved **cost avoidance of \$700k** on buying new expensive servers while saving 400 work hours in maintenance upgrades.
- Developed ETL pipelines and scripts in Python to archive cold data. Leveraged Erasure coding to achieve savings of 50% in total storage space.
- Created **Ansible playbooks** and Python scripts to build RedHat servers with Cloudera(CDP) Hadoop capabilities.
- Mentored and onboarded new SWEs. Conducted knowledge sharing sessions with my team.

RELEVANT PROJECTS

MovieSync

AWS, Node.js, React, JavaScript, MongoDB

- Created a MERN stack app for synchronized movie playback using **Dynamic Adaptive Streaming over HTTP (DASH)** protocol. Deployed application on AWS EC2 and used S3 and CloudFront for storage and latency improvements.
- Utilized Selenium for comprehensive testing among 50 concurrent users to study buffering rate, bitrate, latency metrics. [\[link\]](#)

Recommendation system evaluation

Pytorch, Tensorflow, Keras, Pandas

- Designed Groceries Recommendation Systems using collaborative filtering techniques - Matrix Factorization, Slope One, and SVD, on the Grocery and Gourmet Food subset of the Amazon Review Dataset. [\[link\]](#)
- Evaluated system accuracy using metrics like RMSE and MAE, and assessed performance through Precision, Recall, and F1 Scores. Report: [\[link\]](#).

Data analysis and Price Prediction of Airbnb listings data

Sklearn, Pandas, Matplotlib

- Used linear regression, XGBoost and multivariate adaptive regression splines (MARS) to predict AirBnB prices based on city, availability, reviews per month and room type. [\[link\]](#)
- Cross validation, feature engineering to optimize model parameters and improve prediction accuracy by minimizing RMSE.

CERTIFICATIONS, ACHIEVEMENTS & SCHOLARSHIPS

- **Grace Hopper Scholarship (2023)**
- **HSBC Star Performer Award (2021,2022)**
- **AWS Cloud Practitioner (2023)**
- **Certified Associate in Python Programming (2021) PCAP**