

MANSI BOROLE

(732)-374-5143 mansi.borole@rutgers.edu [LinkedIn link](#) [GitHub link](#) [Portfolio](#)

EDUCATION

Rutgers University - New Brunswick

Aug 2022 – May 2024

Master of Science in Computer Science

GPA: 3.83/4

- Coursework - Algorithms, Linear Algebra, Massive Data Mining, Artificial Intelligence, Design of Internet Services, Deep Learning

University of Pune

Jul 2015 – May 2019

Bachelor's in Computer Engineering

GPA: 8.61/10

SKILLS

Programming Proficiency : Python, SQL, Java, Shell Scripting

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

Database platforms : PostgreSQL, MySQL, Apache Hive, HBase

Data Visualization : Tableau, Grafana, QlikSense, Matplotlib, Microsoft Excel

Libraries/Frameworks : Numpy, Pandas, Keras, Tensorflow, Pytorch, Node.js, Docker, git

WORK EXPERIENCE

Software Engineer (Big Data)

Jul 2019 – Aug 2022

HSBC Bank

Pune, India

- **Cluster Virtualization:** Migrated Hadoop Master services (Name nodes, YARN, Solr, Zookeeper, Hive) from on-prem servers to Virtual Machines, achieving a **cost avoidance of \$0.7M** on buying new expensive servers while saving 400 man hours in maintenance upgrades using Ansible Automation
- Performed **Hadoop/Linux administration** provisioning space on clusters by building and decommissioning servers, Node repurposes from one cluster to another and installing Hadoop components and services on the **200+ on-prem servers**
- Developed **ETL pipelines** and scripts in python to Load Data and support application.
- **Nifi Dataflow:** Set up a one-stop dashboard BigData360 and data-flows to monitor Hadoop services and alerts on clusters using Ambari Metrics/ Cloudera Metrics APIs that **reduced the mean time to repair failures by 1 hour**
- **Ansible Automation:** Created Ansible playbooks and python scripts to get daily FS image from clusters and find the daily organic growth of client's HDFS data on the clusters
- Created a MySQL tool to generate Chargeback reports to determine the cost of Physical Hardware, Licenses, and Supporting Staff for each of **2000+ servers** in the estate based on resource usage and used Qlik Sense to publish the reports.
- Wrote housekeeping scripts that would reduce replication factor of cold data on the cluster, saving 45TB space per month

Communications Leader

Jul 2019 – Aug 2022

HSBC Bank

Pune, India

- Lead a team of 4 colleagues delivered high-quality, impactful quarterly newsletters highlighting project **success stories and KPIs** from the 'Group Data Technology' Department, closely working with Ex Co members in US, UK and Asia-Pacific regions
- Published team's quarterly achievements and highlights via Newsletters to the entire department

PROJECTS

Exploring Business Intelligence - Data Visualization [Tableau Public Portfolio](#)

Tableau

- * Analyzed and created reports for Credit Card Complaints Trends, Netflix movies analysis, Salary Insights in NY State showcasing the Trends, Predictions, Density maps, Race Charts, Scatter Plots, Timeseries.

Machine Learning

Pytorch, Keras, Pandas

- * [AirBnB Price Prediction](#): Used linear regression, XGBoost and multivariate adaptive regression splines (MARS) to predict prices based on city, availability, reviews per month and room type.
- * [Image Similarity Search](#) using a pretrained Convolutional Neural Network (CNN) model and then using KNN classifier to find similar images. Achieved an accuracy of 98%
- * Top 10 Friend recommendations with pySpark using MapReduce jobs on a dataset of 49995 records.

Fast Trajectory Replanning using AI Agent

Python, Matplotlib

- * Evaluated and compared the model performances (paths) generated by 4 flavors of A-star search algorithms like Repeated A-Star, forward/backward A-star, Adaptive A-star [Project Link](#)

Data Analytics

HTML, NodeJS, Gulp, JS

- * Implemented Customer engagement tracking using location heatmaps and website interaction using **Google Analytics/Google Tag Manager** [Project Link](#)

CERTIFICATIONS

- **Tableau** Certified Desktop Specialist (2022)
- PCAP - Certified Associate in **Python** Programming (2021)