Mitali Bharali

469-380-3996; mitali11bharali@gmail.com;
https://www.linkedin.com/in/mitalibharali/; https://github.com/mitalibharali; https://public.tableau.com/profile/mitali.bharali#!/

EDUCATION

The University of Texas at Dallas 2018 - 2020
Masters, Business Analytics; Data Science 3.6

The LNM Institute of Information Technology, (NAAC A) 2012 – 2016
Bachelors, Electronics and Communication; Computer Science 3.8

CERTIFICATIONS & TECHNICAL SKILLS

Certifications:	Machine Learning with GCP – Google Cloud-ID: JV88G9HDX5DL, Bayesian Statistic (R)-ID: CVGKUBSEUX56, Programming for Data Science with Python, Deep Learning with Tensorflow – UT Dallas
Data Analysis Tools:	**Python**, **SQL**, **PostgreSQL**, R, SAS, Stata, MATLAB,
Big Data Technologies:	**PySpark**, **Scala**, Hadoop, **SparkML, GCP- DataProc**
Programming:	**Python**, SQL, C, R, **GIT**
Visualizations and Libraries:	**Tableau**, **PowerBI**, QlikView, Plotly, **D3.js, Gephi**, Keras, NumPy, Pandas, scikit-learn, Spacy
Databases and Cloud:	**AWS**, **GCP, PostgreSQL**, MS SQL Server, MySQL
Statistical Techniques:	Predictive Modeling, Time Series Modeling, Bayesian Statistic, **Markov Chain Monte Carlo**, Hidden Markov Model

BUSINESS EXPERIENCE

CITIBANK, NA July 2020 – Present
Model Development & Data Scientist – Economic Forecasting

- Develop Model Validation python scripts for timeseries – Switching regression in **SAS** and Bayesian Network in **python** for forecasting of GDP variables for FED's recession prediction
- Automate Data loading process of 1723 databases using **python** and exporting using **PyODBC** for its **SQL** connections

The Beneficent Company Group L.P, Dallas June 2019 – Nov 2019
Data Science Intern

- Built risk models to forecast key performance metrics which led to 20% improvement in error rates potentially resulting in $200,000 savings using **Xgboost** in **Python** in **GIT**
- Forecasted Net Asset Values using **Timeseries modeling** and **sequential forecasting** in Python for low data clients and helped contribute additional revenue of $70,000
- Improved performance of querying data by 14% using **PostgreSQL** and maintained databases over **AWS S3** buckets
- Identified funds into asset classes with multiclass classification and cluster analysis using **sampling**, MICE, Knn and **feature engineering**
- Evaluated risk and performance benchmarks and estimated Risk contributors with **hypothesis testing**, **Monte Carlo simulations**
- Developed dashboards in **PowerBI** and visualizations in **Plotly** to monitor increasing net asset values and cashflows of funds
- Collaborated with Risk, Underwriting and IT leaders to assist on cross-functional duties for CRM development and **PowerBI** integration

Iron Mountain Inc August 2017 - July 2018
Operations Data Analyst

- Improving data qualities by removing 67% discrepancies in the data by performing **root cause analysis** with **SQL procedures**
- Designed and developed **ETL process**, stored procedures, views for data analysis, extraction, and complex mapping of **Salesforce** data.
- Built dashboards in **Tableau** to identify gaps in revenue potential and recognize cross sell opportunities leading to 4-11% increase in sales
- Strategized an $8 million gain based on **Power BI** and **Tableau** dashboard insights to tap on potential clients from churn data

SilverPush March 2017 – July 2017
Data Science Analyst

- Formulated a price optimization framework to fine tune the product pricing strategy leading to a 25% increase in monthly revenue using **Multilinear models** in **Python**
- Improved CTR of the recommended products by 58% by building a custom recommender system by analyzing web traffic information of similar user's user-flow in **Shopify** through python
- Achieved a 20% process optimization developing product SKUs and tracking user-flows through BI reports and dashboards

UrbanClap June 2016 – Feb 2017
Data Analyst

- Collaborate between Product and Engineering team to develop automation tool to reduced TAT by 43% by recognizing blockage in **CRM**
- Develop solutions for customer acquisition system through CRM and debug system reducing **churn rate** of customers by 11%

PROJECTS

Conversational Chat Bot for Customer Experience using vectors and embedding in Natural Language Processing April 2020

- Labelled a dataset with topic modeling, trained it to return closest match articles of an input question using GUI chat bot - accuracy 80%

Machine Learning algorithm over Google Cloud Platform Using Pyspark MLlib November 2019

- Made VM instance, cluster and firewall rules on GCP to instantiate Jupyter notebook to predict people with strokes from a highly imbalanced dataset using Pyspark MLlib libraries with AUC area score of 0.98