ARPITA DESHMUKH

College Station, TX | (979)739-8643
arpitapd@gmail.com | linkedin.com/in/arpitapd | github.com/arpitapd | people.tamu.edu/~arpitapd/

EDUCATION

Texas A&M University, Mays Business School — College Station, TX
Master of Science in Management Information Systems (GPA:3.55) — ***August 2018 - May 2020***

University of Pune — Pune, India
Bachelor of Engineering in Information Technology — ***August 2014 - May 2018***

SKILLS

Languages: Python (Pandas, NumPy, Matplotlib, Flask) | SQL | R programming | SAS | HTML | CSS
Data Visualization: Tableau | Targit | PowerBI | Advanced Excel (PowerPivot, VBA)
Statistical Algorithms: PCA | Bagging | Boosting | Linear Regression | Logistic Regression | KNN | Decision Trees | Clustering
Databases: Neo4J | MariaDB | SQL Server | DynamoDB | MongoDB | Cassandra | MS Access
Tools: AWS (EC2, S3, Redshift, RDS) | AirFlow | SAS Miner | SSIS | SSRS | SSAS | MS Visio | XL Miner

EXPERIENCE

Strategic Materials, Inc — **Houston, TX**
Data & Business Intelligence Analyst Intern — ***May 2019 – Aug 2019***

- Conducted requirements elicitation from C-suite executives, third-party consultants and stakeholders for creating drill-down reports, real-time interactive dashboards; delivered BI projects 4 months prior to deadlines
- Reduced data warehouse refresh time from 2 hours to 30 minutes by implementing incremental load strategy; enhanced ETL solutions, updated dimensional data model in Targit while coordinating with developer
- Researched and analyzed industry trends, consolidated reports to compare company and industry standards associated with rail freight costs; resulted in potential annual savings of $5 million

6-Pak Dashboard Project

- Developed 20 KPIs for real-time inventory monitoring by creating dashboard with relevant metrics for 50 global locations

Tracking on-Time shipments

- Created real-time dashboard to track status of shipments across global facilities for past 6 months; incorporated email alert system for notifying carriers about delayed shipments

PROJECTS

Movie Recommender System using NoSQL DB & SQL DB — **Graph – Neo4J | MariaDB with Galera clustering on AWS**

- Built recommender system using Netflix and IMDB data (3GB) which enables users to search movies, view watchlist of their friends based on users rating, personal preferences and user history using SQL, Cypher Language, APIs

Data Warehouse Implementation for Retail Store Chain — **ETL | SSIS | SSAS | SSRS | MS EXCEL**

- Built data warehouse solution to provide analytics & answer key business questions for reduction in sales by developing data marts

Zoom Travel — **Python-Flask | AWS -RDS | AWS Elastic Beanstalk | HTML | CSS | JS | Bootstrap CSS**

- Created web app to assist nervous fliers to have easy flight experience by tracking trip progress using dynamic checklist

Classification of Risk Mortality for Patients Above 65 Years — **Tableau | Decision Trees | Naïve Bayes | Logistic | SAS Miner**

- Performed dimension reduction, added derived variables to train models & correctly classified severity of risk mortality of patient
- Conducted Hyper-parameter tuning on Naïve Bayes; resulted in decrease of misclassification error rate from 52% to 47%

Equipment Failure Prediction (Conoco Phillips TAMU Datathon) — **Random Forest | Logistic | AdaBoost | Gradient Boosting**

- Implemented classification machine learning techniques by analyzing 172 sensor data with Random Forest emerging as best model
- Performed resampling and cross validation to deal with imbalanced data; led to increase in F1 score from 95.25% to 99.32%

LEADERSHIP & CO-CURRICULAR ACTIVITIES

- **TAMU Data Analytics Club (Corporate & VP of Treasury team) -** Reduced events expenditure by 50% — **Jan – Present**
- **Runner up position out of 15 Teams at Keystone Case Competition, TAMU -** Analyzed multiple B2C/B2B eCommerce platforms to provide financially viable solution — **Dec 2018**