

Saloni Thakkar

College Station, Texas | (979) 422-7858 | saloni123@tamu.edu | [GitHub](#) | [LinkedIn](#)

EDUCATION

Texas A&M University, Mays Business School

College Station, Texas

Master of Science in Management Information Systems | GPA: 4.0/4.0

May 2021

Awarded Graduate Teaching Assistant and Tuition fee waiver

Relevant Courses: Database Management Systems, Data Warehousing, Statistical Computation with Big Data

Gujarat Technological University, VGEC

Ahmedabad, India

Bachelor of Engineering in Electronics and Communication Engineering | GPA: 3.7/4.0

June 2019

TECHNICAL SKILLS AND COURSES

- **Languages** – Python, R, SQL, SAS
- **Visualization/Big data Tools** – Tableau, Power BI, Hadoop, Looker, Spark, Hive, Advanced Excel
- **Databases/ETL tools** – MySQL, MS SQL Server (SSIS, SSRS, SSMS), MongoDB, Cassandra, Redshift, S3
- **Other Skills** – AWS (EC2 and S3), MS Office Suite, Git

PROFESSIONAL EXPERIENCE

HP Inc.

San Diego, California

Data Scientist Intern

May 2020 – August 2020

- Designed an ETL data pipeline for processing more than 70GB of data to automate sales data validation using pandas, and chunking data optimized batch performance by 40%
- Tuned SQL reporting queries and updated indexes for sales data to reduce memory usage by 17%
- Developed an econometric model to assess sales growth of products on Amazon using hypothesis testing for pre-post test, improved data quality by removing outliers
- Collaborated with analytics team to communicate 34% gross sales margin growth with 95% confidence interval caused by the change in product appearance on product detail page
- Analyzed product-level sales daily data using Amazon API to determine COVID effects on product advertising and created managerial report using PowerBI
- Implemented a predictive model using search keyword data with an accuracy of 83%; collaborated with digital marketing team to optimize the advertisement bidding process

SoftVan

Ahmedabad, India

Intern

February 2018-August 2018

- Visualized customer retention for client's eCommerce web application by building interactive Tableau dashboard; identified key metrics to lower the customer churn rate by 33%
- Spearheaded team of 7 employees for software development project in Agile environment to gather requirements, defined scope, track timeline, tasks, and resources; enhanced client satisfaction by 27%

ACADEMIC PROJECTS

Log Analytics Solution | Amazon Kinesis and Elasticsearch

December 2020-February 2021

- Integrated AWS tools to collect, ingest, and process real-time batch streaming data to perform log analytics and visualize server access activity which can be used for digital marketing, fraud analytics, ad tech and IOT
- Utilized Kibana to visualize server activity at edge and provide live insights based on log data

Credit Card Fraud Detection | Python | Tableau

November 2020- January 2021

- Performed exploratory analysis and utilized SMOTE technique & KNN imputation to identify key variables, patterns and handle missing values in dataset; KNN performed best with 82% accuracy

Dominick's Finer Food (DFF) Chain | MS SQL Server (SSIS, SSAS, SSRS)

September 2020-November 2020

- Designed data warehouse for retail chain and implemented an OLAP system by creating ETL pipeline in SSIS
- Generated report using SSAS, SSRS and Tableau to communicate sales trend and promotions

Data Warehouse Deployment | Amazon Redshift

August 2020-October 2020

- Configured Amazon Redshift to connect with SQL Workbench/J client assembling a scalable and high-performance data warehouse solution to analyze actionable insights from the stored data

Airline Recommendation System | Python | MongoDB | Tableau

January 2020-March 2020

- Mined 10GB data, performed ETL using Pandas and developed application using MariaDB in a Galera cluster
- Migrated from MariaDB Galera cluster on AWS EC2 instance to NoSQL to facilitate user's decision making with state-wise Tableau dashboard to communicate trends