

EDUCATION

Master's in Business Analytics,

University of Illinois at Chicago, IL

Expected Year of Graduation : December 2022

Coursework: Applied Statistics, Big Data, Machine Learning, Advance Database Management, Data Mining, Marketing Analytics

Bachelor of Technology in Information Technology,

Gandhinagar Institute of Technology, India

August 2015 - May 2019

SKILLS

Languages and Databases:	Python, R, SQL, MySQL, SQL Server, SAS
Machine Learning Models:	Linear/Logistic Regression, Decision Trees, KNN, XGBoost, K-Means
Libraries:	Pandas, NumPy, Scikit-learn, Seaborn, PySpark, TensorFlow, Keras, MLLib
Data Integration and Data Engineering:	Alteryx, Tableau Prep, Spark
Business Intelligence:	Tableau, MS Excel, Power BI, SSIS
Certifications:	Tableau desktop specialist

PROFESSIONAL EXPERIENCE

Genpact / Data Scientist

December 2020 – June 2021

- Experience in using AWS (EC2, S3, RDS, EMR, Sagemaker) and GCP (DataProc, Big Query, GCS, APIs) services
- Successfully transitioned 12 Dashboards from MS Excel to Tableau, which reduced manual work by 60%.
- Created CHAID (Decision trees) models to identify key drivers of new MA (Medicare Advantage) member churn.
- Predicted demand across multiple product segments and geographies which helped to reduce overall inventory cost by 1%.

Online PSB Loans | Data Analyst

August 2019 – December 2020

- Helped a Healthcare client in optimizing there existing workflow and finding new opportunities with the existing data which includes activities like data wrangling, demand forecasting, report for pacing rate, etc.
- Developed 20+ interactive dashboards in Tableau showcasing business KPIs, Center of Influence, and Client Engagement.
- Reduced 80% of manual work by developing python script to automate extraction of data from LinkedIn Elevate and cleaning the data in MS Excel.
- Having hands on experience in building recommendation engine, Churn, Acquisition, and CLTR model
- Built ETL data pipeline to automate the data transformation and load scheduling using SQL Server Integration Services.
- Contributed to development of One Big Score (Credit score) by performing data extraction & feature Engineering

Cognisun | Data Analyst

June 2018 – July 2019

- Created workflows using the Alteryx tool to perform data modeling and transformation for process improvement and optimization of current data analytics procedure, resulting in minimizing turnaround time by 40%.
- Experience in using tools like Jira and Confluence
- Built a pipeline to convert unstructured data into a structured format for analysis for a logistics client
- Undertook end to end delivery of data analysis requirements, using firm standard tools and technologies, including gathering business understanding, managing project delivery and producing results as per defined quality guidelines and requirements.

PROJECTS

Health Effects Of Increasing Toxic Dumping & Poor AQI

- Cleaned and Integrated AQI, TRI, Cancer and Life expectancy data for 65 US counties from publicly available data sources.
- Determined most significant features influencing cancer rates by performing Random forest regression & PCA.
- Predicted future cancer rate with an accuracy of 85% using Multi Output Regressor and ANN for model training.

Flight Fare Prediction (Python, Flask, Heroku)

- Implemented One-Hot Encoding and Label Encoding to handle the nominal and ordinal data.
- Performed feature selection using Extra Tree Regressor and hyperparameter tuning on Random Forest using Random Search CV with 5 folds, achieving an accuracy of ~83% on the test data.
- Completed the End-to-End implementation by deploying the machine learning model in Heroku using Flask.