MEGHA GANDHI

Principal Data Scientist at Verizon | Lean Six Sigma Green Belt Certified | 214-940-9433| Dallas, TX



**P R O F E S S I O N A L S U M M A R Y :**

* Data driven professional with over 12 years of experience in the field of analytics and m ore than 8 years of experience in **designing, interpreting and deploying machine learning algorithms using Python. Utilized various libraries and frameworks like Keras, TensorFlow, NLP, PYOD for designing & developing ML models**
* Proficient knowledge of **Statistical Modeling**, Machine Learning, Data Visualization, SQL and Analytical Tools like **Tableau. QlikSense, PowerBI**
* Experience in analyzing large datasets on distributed databases and developing Machine Learning algorithms to gain operational insights and present them to the leadership
* Extensively involved in **Data preparation, Descriptive Analysis, Exploratory Analysis, Feature Engineering and Predictive Analysis** using Supervised Learning, Unsupervised Learning, Deep Learning and NLP algorithms.
* Experience in creating interactive dashboards using Tableau, MS Excel and IMS Clear Decision
* Actively involved **in SAFe ARTs implementation**, participated **in SAFe Product Increment (PI) and iteration planning events**
* Excellent understanding of business operations and methodologies for effective **project management**

### Second Runner-Up for Verizon Data & Analytics Single Pane of Glass Gen AI Hackathon 2024

**E D U C A T I O N A L Q U A L I F I C A T I O N :**

* **Masters of Science in Marketing Research and Analytics;** University of Mumbai; June 2013
* **Bachelor of Engineering in Information Technology;** University of Mumbai; May 2008

### Lean Six Sigma Green Belt Certified: May 2024

**Certifications:**

### GCP - Google Cloud Professional Associate Cloud Engineer- In Progress

* + **Customer Analytics**; The Wharton School, University of Pennsylvania
  + **Python for Data Science and AI** - IBM Data Science Professional Certificate
  + **Machine Learning with Python** - IBM Data Science Professional Certificate
  + **Introduction to Data Science** - Microsoft Data Science Professional Certificate
  + **Developing Big Data Solutions with Azure Machine Learning** - Microsoft Data Science Professional Certificate
  + **Analyzing and Visualizing Data with Power BI** - Microsoft Data Science Professional Certificate
  + **Neural Networks and Deep Learning** – DeepLearning.AI

**T E C H N I C A L E X P E R T I S E :**

|  |  |
| --- | --- |
| **Programming:** | Python, R Programming, SQL, Azure ML |
| **Algorithm & Techniques:** | Supervised Learning, Unsupervised Learning (Clustering), NLP, Deep Learning, Sentiment Analysis, Dimensionality Reduction (Principal Component Analysis), Feature Extraction /Transformation techniques, Anomaly reduction (outlier detection), NLTK, Keras, TensorFlow, Scikit-learn, Customer Analytics, Brand  Equity Management, Business Intelligence & Reporting, Gen AI |
| **Database:** | Oracle 10g, PostgreSQL, SQL, Hive, GCP Big Query, Teradata SQL Assistant |
| **IDE:** | Jupyter, Spyder, PyCharm, GCP, |
| **Visualization:** | Tableau, Qlik Sense, Power BI, Kibana, Matplotlib, Seaborn, Plotly, Advanced  Excel |
| **Methodologies:** | SAFe, Agile |
| **Others:** | Apache Pulsar, Apache Flink, REDcap,Apache Spark, JIRA, Azure Cloud Services, IBM  Cloud Services, Tableau Desktop, GitLab, Slackbot/Chatbot |

**W O R K E X P E R I E N C E :**

# Principal Data Scie ntist | Jan 2021 – Till Date

## Verizon, Irving TX

### Project 3: Created a Classification Model to predict the probability of attrition at the retail store level (current)

* + Built a **CatBoostClassifier** model that predicts the probability of attrition at the store level with a **precision of 72% and recall of 85%**
  + Due to the imbalanced nature of the dataset, I applied **oversampling, undersampling and a mix of oversampling under-sampling techniques** such as **SMOTE, Tomeklinks, NearMiss, etc** to make sure that we

are feeling the right data into the model for training and predictions. Decided to go with **SMOTE** to retain information in the data

* + Performed rigorous data preprocessing, and statistical and exploratory analysis to check for correlations and collinearity in the features
  + Performed image classification on using tensorflow and pytorch
  + Developed Tableau visualizations and dashboards using Tableau Desktop.
  + Performed feature extraction and feature engineering before passing the final data for modeling**. Currently in the phase of optimizing the model by feeding in new data**
  + Produced a production-ready model that helps businesses make important and timely hiring decisions

### Won a hackathon for creating a Single Pane of Glass view for VP & Above business users, utilizing the power of GEN AI

### Experience working in AI and Deep learning platforms such as PyTorch, TensorFlow, CNN, RNN

* + **Extensively worked on LLM models like Copilot, AIbel, and Vegas (Verizon's in-house LLM Model). These models are used to assist business users get answers to questions they have on the SPoG report**
  + Implemented **a network of 28 functions that are scalable to build a 5-fold Anomaly Detection platform** to generate multifold insights.
  + Utilized PYOD library designed for **anomaly detection in multivariate as well as SHAP** to describe the contribution of features in the anomaly

### Project 2: Created an integrated, user-friendly tactical tool that can track real-time and historical performance, while providing meaningful insights and visibility into Field Engineers' day-to-day activities

* + Worked hand in hand with business partners to plan and deploy a **Tactical Operational tool** that provides **insights on activities of a largely autonomous workforce** of 1,500 employees that maintain and operate the Verizon Wireless Network through a push/pull model

### Performed and presented complex descriptive analytics on completed and open tasks and impacts on expected SLAs - all the data resides in GCP and has been used extensively to perform these tasks

* + **Tableau report** presents **Historical data** to **spotlight team/individual trends** over specified periods, outliers, types of tickets at market, submarket, category, and subcategory levels

### Working closely with the business and AI&D team for the tool's adoption in a predictive model for capacity planning and forecasting

* + Actively presenting, explaining, and taking feedback from Executive Directors and Managers to Increase the adoption of this **report** and deliver KPIs like Average ATTR, Bucketed durations (open and resolved tickets), Average segment durations, Ticket Volumes, etc
  + Working on **message bus (Apache Pulsar) and Apache Flink framework** to perform real-time streaming analytics and consume the same in the tool
  + Developing optimized and scalable machine learning algorithms capable of performing predictive modeling using PyTorch, Spark, SparkML, Scikit-learn and other packages.
  + Being SME for Ticket data and Ticket lifecycle, helped multiple cross-functional teams by educating on remote tasks, on-site tasks, and automated tasks
  + Working on creating a **Planner Prioritization Algorithm** that would recommend the next most important task to be picked up by an engineer from his planner for having the **best network service** at all times thereby providing the **best-in-class customer experience**
  + Created RESTful APIs using frameworks like Flask in Python and consumed them from a .NET application
  + Worked closely on **Digital Twin Initiative** to demonstrate end-to-end **DILO (Day in Life of an Engineer)** by taking into account the **tasks in his queue, site location, and engineer's GPS location** to evaluate various metrics that can recommend tasks to be performed next in real time

### Project 1: Determine the Field Engineer’s efficiency, efficacy and utilization by predicting the Effort spent to complete Tasks and Tickets for the Network Assurance segment

* + Designed, developed, and trained multi-layered **TensorFlow Keras Sequential Machine Learning** comprising of **Preprocessing Raw Data, Outlier Detection, Model Trainer, and Deployment Modules**
  + Extensively utilized Python libraries and modules to perform end-to-end Machine Learning processes involving **Univariate Analysis, Bivariate Analysis, Data Transformation, Feature Engineering , and Fine- Tuning Accuracy**
  + Analyzed records of Field Engineer’s geo-tagging and work profiles based on Task Segment, GPS, and IVR location using Oracle SQL functionalities to achieve accuracy in predictions
  + Produced **graphs and plots** at various ML stages to show the Correlation, Dependencies, Outliers, and

### Explainability of the Deep Learning output

* + Created **Tableau Dashboard** for the business owners to visualize historical data versus current data at Task, Sub-Task, Market, Sub-Market, and Engineer Level
  + Designed, and developed a multi-layered model in TensorFlow Keras Functional API with Multi-Output to get two predicted outcomes from single input DataFrame
  + Deployed Math Function using Haversine Distance to **determine Engineer’s Site Presence and Repeat Site**

**Visit count** in Pandas

* + Designed and deployed an API to **Auto Pause Engineer’s Sub-Task** once he falls out of the Cell Site’s radius using GPS and Action Data. Created python functions using Pandas, Numpy, and Stored Procedure in Oracle to update the Sub-Task End Status
  + Worked on **Oracle SQL and Postgresql databases** to churn out and collate relevant data for analysis.
  + Participated in **Quality assurance (QA) and created them on JIRA** for updating the DDPA Tool. Worked closely with SMEs to gain an understanding of business expectations and assisted in a successful deployment
  + Actively involved in SAFe ART and SAFe Product Increment (PI) and iteration planning events

# Data Scientist |Aug 2017 – Oct 2020

## SilverXis INC, Ir ving TX (Client - Cuna Mutual Group : Insurance Company)

* + Performed **data extraction, manipulation, cleaning, analysis, statistical modeling, actuarial analysis, and data mining** using Azure, and Python to predict the time taken in the underwriting process to minimize the overall time for approval
  + Designed 10+ dashboards in Tableau for sales managers with instant access to personalized analytics portal, so they can access key business metrics resulting in increased customer satisfaction and delivering optimum performance.
  + Utilized Power BI for descriptive statistics for use cases and integrated that with Azure for harnessing the power of ML & AI that it provides. Performed end-to-end ML development to deployment using Azure ML
  + Constructed the model by **iteratively** revisiting variables and features. Performed **Feature Engineering** by creating a new variable to improve the model performance and Out of Sample Accuracy
  + Used **One Hot Encoding and Labeled Encoding** to convert categorical variables for better predictions by ML algorithms.
  + Applied concepts **of R-squared, R.M.S.E, Chi Square Test**, etc. in the evaluation stage to extract interesting findings through comparisons.
  + **Performed Dimensionality Reduction** by removing Noise, Redundant, Duplicate, and Feeder Variables to get better accuracy of the algorithm
  + **Built Logistic Regression, Ensemble Model and Decision Tree machine learning algorithms** in Python (Pandas, Numpy, Scipy, Scikit-Learn) to compare the risks and to recommend the most appropriate model resulting in better customer retention
  + Worked on writing complex SQL queries in performing Data analysis using window functions, and joins, improving performance by creating partitioned tables.
  + Developed POC for Slack integration with NLP using Azure ML

# Senior Business Dat a Analyst | Ap r 2016 – Jul 2017

## SilverXis INC, Ir ving TX ( Client - HGS)

* + Worked for customers to deliver various data-driven strategies performing cognitive and predictive analytics. **Delivered Interactive visualizations/dashboards using Tableau** to present analysis outcomes in terms of patterns, anomalies, and predictions.
  + Developed ML model by performing end-to-end analytics to provide a desirable outcome that helped the company to retain the employees and reduce volatility in the business.
  + Proficient with using scripting languages such as Python and data manipulation/analysis libraries such as Scikit-learn and Pandas for analyzing and modeling data
  + Prepared the data by **detecting outliers, treating missing values, and transforming variables.**
  + Created **SQL queries** to validate and perform Data Transformation. Involved in writing queries in SQL on large datasets to derive meaningful insights.
  + Conducted and facilitated sessions with various departments to understand business needs and challenges faced. This helped to define detailed business requirements for solutions and development of the predictive models
  + Worked with Data Engineers, Business Teams, and HR Teams to provide respective analysis on an ad-hoc basis
  + Assisted in the analysis and design of new business processes around requirements definition, information gathering, and consistent documentation methodologies
  + Collaborated with stakeholders to address critical issues and implemented process improvements

# Senior Executive - Marketing Analytics | Jul 2013 - Dec 2015

## Nielsen Corporation, Mumbai MH

### Worked on Indian Readership Survey (world’s largest continuous readership study with an annual sample

**size exceeding 330,000 households)** to cover over 600 publications, 71 product categories (at penetration and

brand level) and product consumption behavior along with the Indian Demographic Report at a pan -India level ***(you can view the project here :*** [***IRS Topline Report***](https://mruc.net/uploads/posts/b8856b4f30c91fc0c156ce2ea9961873.pdf)***)***

* + Analyzed data on customer demographics, preferences, needs, and buying habits to identify potential markets and factors affecting product demand
  + Involved in **sampling and clustering of data** and trained the data collections team on weekly basis to capture right data with minimum noise.
  + **Manipulating, cleansing & processing data** using Advanced Excel
  + Created detailed functional specifications documents, including data samples, use cases, key project assumptions, report samples, data flows and data models
  + Leveraged **Tableau’s built-in graphical capabilities** to create Bar charts, Tree maps, Geo maps, Area charts
  + Build several proofs of concepts dashboards in Tableau and IMS Clear Decisions to lay framework of analytics, find areas of improvement in the extensive process of data gathering and suggest future actions

### Created Executive level KPI dashboards with interactive visualizations

* + Worked on a project starting from gathering Requirements to publishing it to the client. Involved in assisting and training the client at every step of the project

# Senior Data Analyst | Oct 2008 - May 2013

## Various (Nielsen Corporation, Future Bridge, Aditya Birla Ret ail), Mumbai MH

* + **Lead the product strategy and brand equity development** for Holcim across 75 countries measuring marketing effectiveness and brand equity index value relative to competition.
  + Increased 4+ points share in 12 countries and overtook competitors as an innovative and reliable brand by addressing consumer needs.
  + **Provided key economic and industry analysis** which enabled accurate need vs demand forecasts. Developed and implemented information analysis for Itella in Finland, maintaining profitable momentum in the market
  + Recommended optimum pricing, competitive positioning and service features to drive customers’ purchase
  + **Assisted marketing manager in implementing actionable insights** . Increased the sales by 30% reformulating and reframing the products placement in the stores