KRISHNA MARGALI GOPARAPU
DATA ANALYST

Houston, TX | +1 (585) 351-9309 | goparapu99@gmail.com | [Portfolio](#) | [GitHub](#) | [LinkedIn](#)

SUMMARY

- Experienced **Data Science** Analyst with **2.5** years of expertise in **Data Extraction, preprocessing, Data Validation, Exploratory Data Analysis, Feature Engineering, Data Wrangling, Statistical Analysis, Machine Learning** and **Data Visualization.**
- Strong expertise in **Python, Advanced SQL, root-cause analysis across structured and unstructured data sets, SAS, Oracle, MySQL,** and Cloud based technologies like **Google Big Query (GCP), Amazon Web Services (S3, Redshift), Azure**
- Experience in applying statistical modelling and **machine learning** techniques like **Linear Regression**, **Logistic Regression**, **Decision Trees, Bagging, Random Forest, Support Vector Machine, KNN, K-Means Clustering, Sentiment Analysis, Naïve Bayes, XGBoost, Clustering** in **Predictive Analytics, Segmentation Methodologies, Regression**-Based Models
- Proficient in developing data visualizations utilizing **Tableau**, and **PowerBI** to identify trends and patterns, effectively conveying insights to both technical and non-technical audiences.

SKILLS

- **Languages:** Python, R, SQL, SAS, Java
- **Libraries/Frameworks:** TensorFlow, PyTorch, Scikit-learn, Pandas, Spark
- **Data Science/Analysis:** Data Science, Data Analysis, Data Mining, Statistics, Statistical Analysis
  IDEs: Visual Studio Code, PyCharm, Jupyter Notebook
- **Machine Learning/Deep Learning:** Linear Regression, Logistic Regression, Classification, Decision Trees, SVM, Random Forests, Naive Bayes, KNN, K Means, CNN, Depp Neural Networks, Hypothesis Testing, Natural Language Processing (NLP)
- **Cloud Services/Platforms:** GCP (Google Big Query), AWS, Azure (Data Lake Storage, Databricks, Data Factory, DevOps), Snowflake, Cloud Environments
- **Packages/Tools:** NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow, Ggplot2, MS Office
- **Visualization Tools:** Tableau, Power BI, Alteryx, MS Excel (Pivot table, VLOOKUP, Advanced Charting)
- **Database:** MS SQL Server, Oracle, MySQL, PostgreSQL
- **Concepts/Fields:** Data Visualization, Mathematics, Computer Science, Programming
- **Other Skills:** ETL, Data Cleaning, Data Warehousing, Data Wrangling, Presenting to technical and non-technical audience, SAP, SAP ERP, CRM

EDUCATION

Masters in Computer Science (GPA: 3.6) | Aug 2021 – Dec 2023
Rochester Institute of Technology – Rochester, NY

Bachelor of Technology in Electronics and Communication Engineering | Jun 2017 - May 2021
GITAM DEEMED TO BE UNIVERSITY, Hyderabad, India |

EXPERIENCE

Kdex Global, TX   *Feb 2024 - Present*
Data Analyst

- Developed **machine learning**-based predictive models like **Logistic Regression, Random Forest, Gradient Boosting**, and **SVM** using **Pythons' scikit-learn** to identify high-quality leads for Schwab Private Client service, resulting in an **8%** increase in customer acquisition rates.
- Streamlined **SQL** scripts by transitioning from **Teradata** to Google Cloud Platform (**GCP**), achieving an impressive **80%** reduction in processing times for existing dashboards, enabling faster and more efficient data analysis.
- Led end-to-end **A/B testing** campaigns for email marketing and conducting statistical analysis with **Python** libraries, achieving a **15%** increase in email open rates and a **10%** boost in click-through rates.
- Conducted comprehensive exploratory data analysis (**EDA**) using **SQL**, **pandas**, and **Tableau**, unearthing key business trends and insights to guide strategic decision-making processes across the organization.
- Collaborated cross-functionally with marketing teams to integrate data-driven insights into campaigns to monitor and optimize performance.
- Established data-driven frameworks and methodologies, including designing, and implementing **SQL** databases for campaign tracking and performance evaluation, enabling continuous refinement of strategies based on iterative testing and analysis.
- Collaborated cross-functionally to design and develop interactive dashboards, scorecards, and **KPIs** in Tableau, with a focus on building customized reports and scheduling automated report generation using **Tableau Server.**
- Effectively communicated complex data findings and actionable recommendations to **stakeholders** through interactive dashboards, scorecards, **KPIs** created in **Tableau** and engaging presentations, facilitating alignment and informed decision-making.

Kdex Global, TX   *Jun 2023 - Dec 2023*
Data Analyst Intern

- Conducted end-to-end data management tasks, including data collection, cleaning, **Exploratory Data Analysis** (**EDA**), Feature Engineering, and in-depth analysis using **SQL** and **Python**, to derive actionable insights and drive informed decision-making.
- Gathered and integrated data from various sources, including Human Resources (HR) databases, workday and internal tools, to conduct comprehensive data analysis utilizing advanced **SQL** and **Python** to uncover data distributions and patterns.
- Developed and integrated machine-learning models to predict employee-scheduling needs by collecting historical data and this initiative-reduced time-to-hire for critical positions from 60 to 45 days.
- Conducted exploratory data analysis on large datasets using **SQL** and **Python**, resulting in the identification of over **80%** increase in the accuracy of risk assessment by analyzing trends, patterns, and actionable insights.
- Led the integration of Machine Learning models like **Logistic Regression, Random Forest, XGBoost** and sentiment analysis on Twitter data, leading to a significant **25%** improvement in credit risk management capabilities, predicted bankruptcy, and effectively identified high-risk customers, resulting in improved portfolio management and automated reporting to the team using **Power BI.**
- Collaborated effectively with **cross-functional** teams including HR and operations teams to develop capacity planning processes, optimize headcount, and staffing and conduct ad-hoc analyses using customized scorecards/dashboards in **Tableau**, showcasing key performance indicators (**KPI**) insights and key metrics reporting.
- Optimized workforce scheduling processes, created interactive Tableau dashboards for real-time insights into workforce performance for efficient resource allocation, and achieving cost savings of $2.8 M in recruitment expenses.

Idea Bytes, Hyderabad, India   *Jan 2020- May 2021*
Data Science Intern

- Collaborated with data engineers and operations team to implement **ETL** process, wrote and optimized **SQL** queries to perform data extraction to fit the analytical requirements.
- Extracted data from the database using **Excel, SQL** procedures and created python datasets for **statistical analysis, validation**, and documentation.
- Designed and implemented a real-time dashboard using **Tableau** for monitoring key credit risk metrics, enabling quick decision-making, and reducing the time taken for risk assessments by **40%.**
- Executed complex **SQL** queries with joins and transformation rules to extract data from legacy **MySQL** and **SQL Server** databases.
- Utilized **Python** for data blending and data preparation using **Alteryx** and **SQL** for **Tableau** consumption and publishing data sources to the **Tableau Server.**
- Automated **data cleaning and validation** processes by developing **Lambda** functions, resulting in a **10%** improvement in data processing efficiency, which allowed for quicker insights extraction and analysis.
- Utilized **Kubernetes** to deploy scale, load balance scale and manage docker containers with multiple name-spaced versions
- Implemented advanced analytics in **Snowflake** to uncover trends and patterns within the data, conducted deep dive analyses, to provide valuable insights for strategic decision-making, empowering the team with data-driven predictions and trends.

CERTIFICATIONS

- **AWS Cloud Practitioner (2023)**
- **IBM Watsonx.ai Technical Advanced**
- **Snowflake Essentials – Badge 1 and 2**
- **Databricks Accredited Generative AI**
- **Google Data Analytics Professional (2024)**