

Karthik Reddy Mettu | Data Analyst

karthikmettu14@gmail.com | 713 202 0004 | <https://www.linkedin.com/in/karthik-mettu/>

PROFESSIONAL SUMMARY

Data Analyst with around 3 years of experience in driving data-driven decision-making across diverse industries, including healthcare and finance. Expertise in data analysis, visualization, and predictive modelling utilizing a range of tools and technologies such as Python (Pandas, NumPy, Scikit-learn), SQL, Tableau, Power BI, Hadoop, Hive, AWS, Google BigQuery, Jira, Agile methodologies. Proven ability to leverage machine learning algorithms, statistical analysis, and data mining techniques to extract valuable insights from complex datasets, improve operational efficiency, and mitigate risks. Strong understanding of data governance, regulatory compliance, and data security best practices.

TECHNICAL SKILLS

|  |  |
|--|--|
| <b>Data Analysis &amp; Modeling:</b> Scikit-Learn, TensorFlow, Keras, Logistic Regression, Decision Trees, K-means, K-medoids, Anomaly Detection | <b>Cloud Computing:</b> AWS, Google Cloud Platform (GCP), Cloud Data Storage, Scalable Data Solutions                        |
| <b>Data Wrangling &amp; Cleaning:</b> Python (Pandas, NumPy), Data Cleansing, Data Transformation, Data Quality Assessment                       | <b>Project Management &amp; Collaboration:</b> Jira, Agile, Scrum, Project Documentation, Task Management, Knowledge Sharing |
| <b>Data Visualization:</b> Tableau, Power BI, SPSS, Interactive Dashboards, Key Performance Indicators (KPIs)                                    | <b>Statistical Analysis:</b> SPSS, Statistical Reporting, Data & Regression Analysis   |
| <b>Machine Learning:</b> Supervised Learning, Clustering Algorithms, Predictive Modeling, Model Deployment                                       | <b>Compliance &amp; Security:</b> HIPAA Compliance, Data Encryption, Data Privacy Regulations                                |
| <b>Databases &amp; Data Warehousing:</b> SQL Server, Oracle, Google BigQuery, Hadoop, Hive, Data Warehousing, SQL Query Optimization             | <b>Data Mining &amp; Analytics:</b> Customer Segmentation, Credit Risk Scoring, Portfolio Risk Analysis, A/B Testing         |
| <b>ETL &amp; Data Integration:</b> Apache NiFi, SSIS, ETL Processes, Data Flow Automation, Data Extraction, Data Transformation                  | <b>Data Reporting:</b> Reporting Automation, Business Insights, Actionable Reports   |
| <b>Data Standards:</b> FHIR, HL7, Data Interoperability, Healthcare Data Standards   | <b>Version Control &amp; Collaboration:</b> Git, GitHub, Version Control, Collaborative Development                          |
|  | <b>Data Standards Compliance:</b> Regulatory Compliance, Data Privacy  |

EDUCATION

|  |                            |
|--|----------------------------|
| <b>Master of Science in Statistics &amp; Data Science</b> , University of Houston, Houston, TX       | <b>Aug 2023 – Dec 2024</b> |
| <b>Bachelor of Technology in Mechanical Engineering (Minor in CS)</b> , JNTU Kakinada, Guntur, India | <b>May 2019 – Apr 2023</b> |

PROFESSIONAL EXPERIENCE

|  |                             |
|--|-----------------------------|
| <b>Data Analyst   UnitedHealth Group (via VSSOFT LLC) – TX  </b>   | <b>Aug 2024 - Present</b>   |
| <ul style="list-style-type: none"><li>• Spearheaded the integration of diverse healthcare data systems, including EPIC, Cerner and Meditech, utilizing Talend to streamline workflows and achieve a 30% reduction in data retrieval time, thereby enhancing decision-making processes.</li><li>• Designed and implemented interactive dashboards in Tableau, enabling real-time visualization of patient flow and resource allocation, which resulted in a 20% increase in operational efficiency across clinical departments.</li><li>• Rigorously ensured adherence to HIPAA regulations throughout all data handling processes by implementing advanced data encryption protocols and safeguarding sensitive patient information against unauthorized access.</li><li>• Crafted and optimized intricate SQL queries for data extraction and analysis from SQL Server and Oracle databases, enabling precise insights into patient care metrics and operational performance.</li><li>• Developed and implemented efficient ETL processes using Apache NiFi, automating data flow between disparate healthcare systems and enhancing data accessibility for analytical purposes.</li><li>• Applied FHIR and HL7 standards to ensure seamless data integration and interoperability across various healthcare platforms, promoting cohesive data exchange and collaboration.</li><li>• Built and deployed sophisticated machine learning models using TensorFlow and Keras for accurate diagnosis predictions, contributing to improved clinical outcomes and patient safety.</li><li>• Designed and maintained a scalable cloud-based data warehouse in Google BigQuery, optimizing data storage and retrieval processes to support large-scale healthcare analytics.</li></ul> |                             |
| <b>Instructional Assistant   University of Houston – Houston, TX  </b>   | <b>Aug 2023 – Aug 2024</b>  |
| <ul style="list-style-type: none"><li>• Conducted lab sessions in regression analysis, Bayesian methods, and model evaluation, with 87% of students achieving ≥90% scores.</li><li>• Executed comprehensive data wrangling and cleaning of complex healthcare datasets using Python (Pandas, NumPy) and led workshops on Python, R, Scikit-learn, and Matplotlib to enable students to build complete ML pipelines.</li></ul>  |                             |
| <b>Data Analyst   Cognizant (Client: JPMorgan Chase &amp; Co.) – Hyderabad, India  </b>  | <b>Aug 2022 - July 2023</b> |
| <ul style="list-style-type: none"><li>• Implemented a credit risk scoring model using machine learning algorithms (Logistic Regression, Decision Trees) that improved loan approval accuracy and reduced non-performing loans by 8%.</li></ul>   |                             |

- Designed and built a customer segmentation model using clustering algorithms (K-means, K-medoids) to identify high-value customer segments and tailor marketing campaigns, leading to a 20% increase in customer engagement and a 10% uplift in cross-sell/up-sell revenue.
- Developed and maintained a data warehouse using Hadoop and Hive to store and process large volumes of transaction data from various sources, enabling real-time data analysis and faster insights for business decisions.
- Created interactive dashboards using Power BI to visualize key performance indicators (KPIs) such as customer churn rate, loan delinquency and portfolio risk, providing actionable insights to senior management for proactive risk mitigation.
- Automated data ETL processes using SSIS to improve data quality & efficiency, reducing manual effort & minimizing data inconsistencies. Utilized Agile methodologies for project management, ensuring flexibility & adaptability to changing business needs.
- Conducted data quality assessments & implemented data cleansing techniques to ensure data accuracy & integrity, improving the reliability of analytical models & report. Leveraged cloud computing technologies such as AWS to scale data processing capabilities & reduce infrastructure cost.
- Ensured compliance with all relevant regulatory requirements and data privacy regulations, throughout the project lifecycle. Maintained documentation for all data models, analytical processes, and reports to facilitate knowledge sharing and future maintenance.
- Implemented fraud detection models using anomaly detection algorithms and machine learning techniques, reducing fraudulent activities and safeguarding the bank's assets.

## **CERTIFICATIONS**

---

AWS Certified Data Engineer Associate