SAI KRISHNA REDDY

Data Analyst

| Email-Skagetala08@gmail.com | Phone- +1 (919) 299-1079

SUMMARY

- **Data Scientist with 4+ years of expertise** in **statistical modeling**, **predictive analytics**, **machine learning**, and **data visualization** using **Python**, **R**, and **Big Data** technologies like **Hadoop**, **Spark**, and **Hive**.
- Skilled in the full data science lifecycle, including **data acquisition**, **cleansing**, **engineering**, **feature selection**, **model development**, evaluation (ROC, cross-validation), and production deployment.
- Proficient in **data manipulation** and **visualization** tools such as **Pandas**, **NumPy**, **Matplotlib**, **Seaborn**, **Tableau**, **Power BI**, and **Qlik** to drive actionable business insights and interactive dashboarding.
- Strong experience in designing and implementing machine learning models, including **Decision Trees**, **Random Forests**, **SVMs**, **Logistic/Linear Regression**, **Clustering**, **PCA**, and **NLP** for both structured and unstructured data.
- Adept at translating complex business requirements into scalable data-driven solutions, with proven success across various industries using **SQL & NoSQL Databases**, **JIRA**, and cloud-based data platforms.

SKILLS

Languages:	Python, R programming, SQL
Libraries & Model:	Pandas, NumPy, Seaborn, Matplotlib, TensorFlow, PyTorch, Keras, Scikit-learn, Beautiful Soup, Hugging Face, NLTK, Spacy, Transformers (BERT, GPT, RoBERT)
Databases:	MySQL, PostgreSQL, SQL Server, Oracle, MongoDB
Data Analysis:	Exploratory Data Analysis, Data Warehousing, Data Wrangling, Data Mining, Statistical Analysis, SSAS, Synapse Analytics
Data Engineering:	PySpark, Airflow, Databricks, ETL/ELT Pipelines, Data Modeling, Hadoop, Teradata, BigQuery, Snowflake
Cloud Services:	AWS, Microsoft Azure, GCP
Statistical Analysis:	ANOVA & Hypothesis Testing, F-test, t-test, Time Series Analysis, Regression Analysis
Artificial Intelligence:	Supervised & Unsupervised Learning, Reinforcement Learning, Deep learning Architectures (Neural networks, Transfer Learning), Computer Vision, Natural Language Processing (Semantic Analysis, Text segmentation), ML Ops, LLMs
Tools:	Databricks, Tableau, Power BI, Fabric, Looker, QuickSight, Spark, Docker, GitHub, Jira, CI/CD Pipelines
Methodologies:	SDLC, Agile, Scrum, Kanban, Waterfall

EDUCATION

Master's in Computer Science
Harrisburg University, United States | ***Aug 2025***

Master's in Data Science & Analytics with Advanced Research
University of Hertfordshire, London, United Kingdom | ***Mar 2022***

Bachelor's in Electronics & Communication EngineeringJa
GITAM University, India | ***Mar 2019***

EXPERIENCE

Black Rock, USA
Data Analyst **| Jan 2025 – Current**

- Engineered predictive models in Python (Scikit-learn, TensorFlow) using Logistic Regression, Support Vector Machines (SVM), and Neural Networks to forecast customer churn, improving retention campaign targeting by 25%.
- Conducted rigorous statistical analysis (ANOVA, t-tests, and F-tests) to identify key drivers of hospital readmissions, resulting in a 15% reduction in irrelevant variables and a more accurate, actionable model for the patient care team.

- Executed a seamless data integration pipeline, connecting to OLAP cubes via ODBC drivers and extracting multidimensional data directly into the Python environment for advanced analysis.
- Built robust data workflows to unify disparate data sources, ensuring data integrity & accessibility for analytics and reporting.
- Developed interactive Tableau dashboards to communicate predictive insights and key performance indicators (KPIs) to stakeholders, enabling data-driven decision-making and reducing the time to insight by 30%.
- Translated complex model outputs into strategic recommendations, providing data-backed reports to management that identified opportunities for cost reduction and performance improvement.

TCS, UK
Data Analyst **| Jan 2022 – Aug2023**

- Conducted exploratory data analysis (EDA) to identify patterns and extract actionable insights, supporting strategic initiatives in risk modeling and operational efficiency.
- Utilized Python and Tableau to create visualizations and insight dashboards, helping stakeholders with real-time KPIs and reporting.
- Built and deployed real-time anomaly detection models using Python and Scikit-learn, significantly reducing manual monitoring and enabling faster response to unusual behaviors.
- Developed customized credit scoring models using supervised learning techniques, improving risk assessment accuracy and enabling data-driven decision-making for client-specific verticals.
- Applied Natural Language Processing (NLP) to analyze customer service transcripts, extracting sentiment and intent to improve customer experience and drive an increase in satisfaction scores.
- Leveraged unsupervised learning techniques like clustering and dimensionality reduction to segment credit card customers, enhancing precision in marketing campaigns and product targeting.
- Performed model validation, hyperparameter tuning, and feature engineering to ensure statistical robustness, regulatory compliance, and performance consistency across datasets.
- Designed and executed A/B tests to evaluate the performance of a new recommendation engine, analyzing uplift and conversion metrics for iterative optimization.
- Created interactive Tableau dashboards and automated reporting solutions to provide business intelligence insights to leadership, driving faster, data-informed decisions.
- Collaborated cross-functionally with teams to frame complex challenges into scalable data science use cases and model-driven solutions.

Capgemini, India
Data Analyst **| Jan 2018 – Dec 2019**

- Spearheaded exploratory data analysis using Python's Pandas and NumPy to clean, transform, and analyze complex datasets, uncovering key trends and data quality issues that informed strategic business decisions.
- Engineered new features and scaled existing ones using Scikit-learn to prepare data for predictive modeling, enhancing model accuracy and reliability for critical projects.
- Designed and developed efficient ETL processes using SSIS to migrate and integrate data from SQL Server and Flat Files, reducing data preparation time by 30% and ensuring data availability for analysis.
- Authored complex SQL queries, stored procedures, and functions in Oracle and SQL Server to extract, validate, and prepare data, creating a single source of truth for enterprise reporting and dashboards.
- Built reusable SSIS templates and frameworks to standardize and accelerate the deployment of data integration packages across development, test, and production environments.
- Applied statistical techniques and hypothesis testing to validate assumptions and measure the effectiveness of business initiatives, providing a rigorous, data-backed foundation for model comparison and strategy.
- Transformed complex analytical findings into actionable insights by building interactive Tableau dashboards & SSRS reports (ad-hoc, parameterized, master), enabling stakeholders to track KPIs and make data-driven decisions more accurately and faster.
- Collaborated with cross-functional teams to gather requirements and define key metrics, ensuring that all reports and dashboards directly addressed business needs and delivered clear, actionable intelligence.