SIDDHARTH SOLANKI

siddharth.solankie97@gmail.com | (+1) 4086905298 | LinkedIn | Portfolio | GitHub

PROFESSIONAL SUMMARY

Experienced data analyst with a passion for data science and 2+ years of professional expertise, dedicated to advancing accuracy, efficiency, and productivity through data-driven strategies. Proficient in developing AI-powered solutions complemented by hands-on experience in diverse projects, showcasing a commitment to innovation and impactful solutions.

EDUCATION

Master of Science in Data Science | San Jose State University, San Jose ***Aug 2022 – May 2024***

Bachelor of Engineering in Computer Science | JECRC University, Jaipur, India ***Aug 2017 - July 2021***

EXPERIENCE

Data Analyst | First Alliance Bank ***Aug 2021 – June 2022***

• Spearheaded a customer churn analysis project for the Zambian Revenue Authority, achieving a 20% increase in customer retention, utilizing **R** for analysis and **SQL** for database querying.

• Enhanced Oracle Database query performance by 25%, enabling faster data access and informed decision-making. Leveraged **PowerBI** for advanced visualization.

• Implemented a mobile Point of Sales application with **Java**, as well as created user documentation, and conducted user training sessions reducing customer support queries by 36%.

Data & Systems Analyst Intern | Surya Biofuels Ltd. ***Jan 2021 – July 2021***

•Improved the vehicle tracking system through targeted data analytics using **Python**, achieving a notable 31.5% increase in supply chain accuracy and operational efficiency.

• Executed strategic optimizations to the company's **ERP** system, utilizing data-driven approaches to realize a 22% increase in system performance and throughput.

• Conducted thorough data evaluations to monitor and assess system performance, leveraging technologies such as, **anomaly detection algorithms**, and **statistical analysis** methods to identify anomalies and provide actionable recommendations for process improvements.

Systems Analyst Intern | Airtel Africa ***July 2019 - Aug 2019***

• Collaborated with the network team to analyze customer usage patterns and system performance, leading to a 13% improvement in network efficiency and enhanced stability through informed upgrades and maintenance planning.

SKILLS

♦Python	♦MATLAB	♦PowerBI	♦Snowflake	♦Neo4j	♦Keras	♦NLP	♦OpenAI ChatGPT
♦Java	♦Excel	♦Hadoop	♦AWS	♦MapReduce	♦XGBoost	♦Computer Vision	♦Jira
♦SQL	♦Tableau	♦GCP	♦MongoDB	♦TensorFlow	♦PyTorch	♦Docker	♦Git/GitHub

PROJECTS

Master's projects ***Aug 2022 – May 2024***

• **Household Waste Management System:** Developed a high-accuracy (90%, target >95%) **AI-powered** waste categorization system, combining **Machine Learning** and deep learning models, with actionable recommendations (**ChatGPT** integration) and user-friendly interface (including integration with **Google Maps API**), leading to a projected 40% reduction in household waste mismanagement.

• **SMPD Challenge: Social Media Post Popularity Prediction**: Outperformed first place solution in the Social Media Popularity Prediction Challenge with a 17% increase in accuracy by developing a **Deep Learning** model in Python to forecast post popularity across platforms. Applied **statistical analysis** and machine learning techniques, including **TensorFlow, NumPy, Keras** and **Scikit-learn.**

• **Bay Area Flood Prediction**: Led the creation of a machine learning model for flood forecasting in the Bay Area with 85% accuracy, integrating **real-time** weather forecasts and historical flood records on cloud **(AWS)** to deliver zip-code level forecasts through a **CI/CD pipeline**, reducing data processing time by 50% and significantly improving local flood preparedness and response strategies.

• **Air Quality Forecasting (Predictive Analysis):** Developed an Air Quality Index predictive model using Python and AirNow API data, employing **regression** and **hypothesis testing** for accuracy. Using data at latitude and longitude granularity increased prediction accuracy by 21%. Leveraged big data tools like **Hadoop** and **Google BigQuery** for data processing and environmental trend analysis.