Sai Vittal Ayyalasomayajula

Boston, MA | ayyalasomayajula.sai@gmail.com| [LinkedIn](#) | 857-675-0628

EDUCATION:

Northeastern University, Boston, MA, United States (CGPA: 3.569) Expected December 2025
Master of Science in Data Analytics Engineering
Coursework: Data Management, Data Mining, Computation and Visualization, Foundations of data Analytics, Natural Language Processing

Gandhi Institute of Technology and Management University, India (CGPA: 3.66) August 2019 - June 2023
Bachelor of Engineering in Computer Science Engineering
Coursework: Machine Learning, Cloud Computing, Artificial Intelligence, Probability and Statistics, Data Structures and Algorithms

TECHNICAL SKILLS:

Data Analytics: Machine Learning, Statistical Analysis, Business Insights, Data Models.
Visualization Tools: Tableau, Power BI, Microsoft Excel, Flourish.
Machine Learning: Neural Networks, Deep Learning (CNN), Classification, Regression.
Programming Languages: Python, SQL, C, Java, Cypher, JavaScript.
Cloud: AWS (Redshift, S3, CloudWatch, EC2, Lambda, RDS, QuickSight, Data Pipeline, Glue), IBM, Azure.

WORK EXPERIENCE:

Phoenix Global January 2022 - May 2022
Data Analyst Intern
- Collected and analyzed **500,000+** user engagement records, including click-through rates, geolocation, and device/browser statistics.
- Processed and cleaned **100,000+** datasets, removing duplicates and validating URLs to maintain a 98% data accuracy rate.
- Analyzed user engagement data, providing insights through data transformation and visualization to improve system efficiency and user behavior understanding.
- Conducted trend analysis on **1M+** URL clicks and user behavior patterns that increased engagement by 30%.
- Provided data-driven recommendations that improved server scalability and reduced downtime by 20%, leading to a 15% increase in user retention.

PROJECTS:

Data-Driven Zip Code Ranking for Suffolk County September 2024 – December 2024
- Conducted a data analysis of 50+ zip codes in Suffolk County, evaluating 6+ key factors like population, income, safety, education, housing costs, and public services.
- Applied MinMax scaler to normalize 5+ attributes, assigning weighted scores (e.g., income 30%, education 20%) for a comprehensive data review and report generation
- Developed an interactive dashboard in Tableau with customizable filters, allowing real-time adjustments of weights for key parameters.
- Identified top 10% high-ranking zip codes, with 02116 and 02109 scoring the highest in income, schools, and livability.
- The project enhances decision-making for 10,000+ potential movers, enabling dynamic data-driven insights to choose optimal locations.

Fitness Tracker January 2024-April 2024
- Developed a Fitness Tracker Database for **5000+ users**, enabling tracking of workouts, nutrition, and biometric data for goal-oriented fitness management.
- Improved user engagement by **30%** through comprehensive data tracking and personalized insights, facilitating better fitness decisions.
- Designed and implemented interactive dashboards in Tableau/Power BI, enhancing data visualization and real-time analytics.
- Used Python for data processing and SQL for managing **100,000+** entries, ensuring secure, scalable, and efficient database operations.

Detecting Malicious URLs Using Machine learning February 2023-April 2023
- Trained and evaluated six models (Logistic Regression, SVM, CNN, Decision Tree, Random Forest, XGBoost) for malicious URL classification, achieving **92%** accuracy with Random Forest.
- Developed and deployed a **real-time web service**, reducing cyber threats by 40% through automated URL analysis.
- Applied feature engineering, model optimization, and comparative analysis, leading to a 20% increase in model performance for URL classification.
- Conducted data analysis and audit processes, improving **data integrity by 25%**, and generated detailed reports on threat patterns.

Credit Card Fraud Detection System September 2022 - November 2022
- Developed and tested an SVM model (95.8% accuracy) and a CNN model (98.5% accuracy) to classify credit card fraud transactions, showcasing deep learning superiority.
- Achieved 99.7% accuracy using a deep learning model with multiple layers, improving predictive performance across 500,000+ transactions.
- Designed a modular fraud detection system, increasing detection efficiency by 15% and enabling future integration of advanced deep learning techniques.
- Conducted extensive data analysis on imbalanced datasets, refining models for better fraud detection accuracy and reducing false positives.