Puthalapattu Sagar – Data Scientist

Email: Sagarputhalapattu@gmail.com
Mobile Number: 9390026264
LinkedIn | GitHub | Medium

Professional Summary

Experienced Data Scientist with expertise in big data technologies, machine learning algorithms, and data modeling, adept at delivering analytical insights to support data-driven decision-making. Proven ability to analyze large datasets using Spark, Hadoop, and other big data tools, enhancing data efficiency and driving impactful solutions in healthcare and other data-intensive domains.

Education

Sri Venkateswara Engineering College 08/2019 – 03/2023 | Tirupati, India
Bachelor's Degree, Information Technology
CGPA: 70.0 / 10.0

- Relevant Coursework: Machine Learning, Data Modeling, Big Data Analytics
- Projects: Developed a predictive model for healthcare data using K-Means clustering and Random Forest algorithms

Technical Skills

- Programming Languages: Python, R, SQL
- Big Data Technologies: Spark (MLlib), Hadoop Ecosystem (HDFS, Hive, Pig)
- WEB Frameworks: Flask, Fast-API,
- Machine Learning Algorithms: Linear & Logistic Regression, Time Series Analysis, Support Vector Machine (SVM), K-Means Clustering, K-Nearest Neighbors (KNN), Decision Trees, Random Forest, Naive Bayes, Principal Component Analysis (PCA), Singular Value Decomposition (SVD), Artificial Neural Networks (ANNs), Association Rules, Genetic Algorithm, Bagging and Boosting Techniques
- Data Processing & Analysis: Statistical Analysis, Hypothesis Testing, Data Modeling, Schema Design, Data Transformation
- Data Visualization: Power BI, Excel, Matplotlib, Seaborn
- Cloud Platforms: Azure, AWS, Google Cloud Platform (GCP)
- DevOps & Deployment: CI/CD, Docker, Kubernetes, Flask API, FastAPI
- API & Testing Tools: Postman
- Soft Skills: Effective Communication, Decision-Making, Team Collaboration, Problem-Solving, Adaptability

Professional Experience & Projects

[Link] Data Science Intern - Cognifyz Technologies 03/ 2024 – 09 /2024| Hyderabad

- Collected, processed, and managed large datasets, improving data quality and accessibility, resulting in a 25% increase in data processing efficiency.
- Implemented data administration strategies, resolving systemic errors and improving data usability by 20%, providing critical insights for strategic business decisions.
- Delivered actionable insights to stakeholders, aligning project goals and increasing project success rates by 15%, ensuring timely data delivery and optimal performance measurement.
- Collaborated with cross-functional teams to process large-scale datasets, implementing data processing solutions with Spark, leading to a 25% improvement in data accessibility and processing time.

[Link] |Customer Churn Model Using BERT and CI/CD Integration

- Increased churn prediction accuracy by 18% within six weeks, helping the company reduce customer churn by enabling timely interventions.
- Automated the deployment pipeline using Databricks and Azure DevOps, reducing manual intervention time by 40% and improving overall efficiency.
- Provided insights that boosted retention strategies by 10%, directly impacting customer satisfaction rates
- Applied machine learning algorithms like Logistic Regression and Decision Trees to refine churn prediction models, increasing predictive accuracy by 18% and reducing churn through proactive customer retention.

[Link]|Real-Time Object Detection System Using YOLO and Python

- Achieved 90% detection accuracy in real-time applications, delivering performance improvements 20% faster than expected due to Flask/FastAPI integration.

- Reduced latency by 20%, improving real-time performance and allowing the system to be deployed in real-world scenarios with minimal lag.
- Completed the project 15% ahead of schedule, meeting client demands and improving deployment timelines.

[Link]|Q&A System for Finance Documents Using Azure and RAG Model

- Enhanced document processing efficiency by 30%, reducing manual labor by 10+ hours per week within the first month of implementation.
- Increased data retrieval speed by 20%, significantly improving query response times for the finance team.
- Automated document handling, boosting overall team productivity and reducing error rates across departments by 12%.

Awards and Certifications

- ✓ **Hacker Rank Python Certification**
  - Demonstrated proficiency in Python fundamentals, including data structures, algorithms, and problem-solving
  - techniques. Earned a top score, validating strong programming and coding skills.
- ✓ **Hacker Rank SQL (Intermediate) Certification**
  - Proved ability to work with complex SQL queries, focusing on data manipulation, joins, subqueries, and
  - analytical functions. Certified for intermediate-level SQL expertise and database management.
- ✓ **Cisco Certified Data Analyst**
  - Acquired hands-on experience in analysing real-world data sets and performing data-driven decision-making.
  - Conducted multiple projects using statistical analysis, data visualization, and reporting techniques.
  - Gained proficiency in data manipulation, data cleaning, and implementing automation for data workflows.
- ✓ **Generative AI Certification – Great Learning**
  - Completed a comprehensive course on Generative AI, covering techniques like GANs, autoregressive models, and
  - transformers. Gained hands-on experience in building and deploying AI models for various applications.
- ✓ **Big Data Certification**: Demonstrated proficiency in Hadoop, Spark, and other big data tools for large-scale data processing.
- ✓ **Machine Learning Certification**: In-depth understanding of algorithms including SVM, PCA, and Neural Networks, focusing on practical applications in data analysis and model building.