SATISH VENNAPU

240.467.6844 | satish@terpmail.umd.edu | GitHub: Satish-Vennapu | LinkedIn: satishvennapu

Data Scientist Summary

- Data Scientist with 2+ years of experience leveraging big data, data analysis, and advanced machine learning algorithms to drive innovations in diverse sectors, including research, e-commerce, service, environmental safety, and automotive industries
- Directed an ML team to engineer SOTA CNNs, achieving a breakthrough of 95% accuracy in wildfire detection and segmentation
- Boosted revenue by 20% ($765K) through predictive data models and optimized pricing strategies while reducing operational costs
- Co-founder of 2 startups: Crossfire (specializing in AI-driven drones and systems) and Munching Boxes (in the food industry)

Skills

Programming Languages: Python, C++, SQL, R (basic), Linux Shell
Frameworks and Computing Tools: Pandas, Tableau, Scikit-Learn, TensorFlow, PyTorch, Streamlit, Hadoop, Spark
Data Technologies: Big Data distributed structures, Statistical Modeling, Hypothesis Testing
Cloud Technologies: Google Cloud Platform, AWS, Git, Docker
Machine Learning: Regression, Classification, Clustering, Neural Networks, SVM, Random Forest, Ensemble, Boosting Methods
Artificial Intelligence: CNNs, Object Recognition, Segmentation, Natural Language Processing, LLMs, Transformers, LangChain

Certifications

- AWS Certified Machine Learning Specialty & Cloud Practitioner
- Google Professional Data Engineer
- Stanford Machine Learning Specialization
- Google Project Management

Work Experience

XPRIZE $ 11 Million Competition and NSF I-Corps Internship — **College Park, MD**
Machine Learning Engineer and Project Manager — June 2023 – Present

- Spearheaded a team of 18, implementing deep learning architectures (CNNs) with Scikit-Learn and PyTorch for wildfire detection, achieving an accuracy of 95%, and slashed false negatives to less than 1% using transfer learning
- Collaborated with 120 stakeholders to launch a $300K-funded startup, pioneering the development of AI-driven drones equipped with state-of-the-art convolutional neural networks (CNNs) for enhanced wildfire detection and suppression
- Optimized model performance and efficiency by executing rigorous A/B testing and hyper-tuning the parameters and engineered automated CI/CD pipelines that streamlined the deployment process, cutting the model deployment time by 65%
- Established the Scrum framework as a Scrum Master, boosting sprint deliverables by 15% and achieving an optimal burn-down rate by prioritizing actionable tasks using Kanban and Fibery
- **Domain & Skills:** Research, Agile Methodology, Model Development, Validation, Production Launch

University of Maryland (iSchool) — **College Park, MD**
Data Scientist (Intern) — August 2022 – May 2023

- Engineered a menu item recommendation engine for food production, addressing 4+ key pain points, increasing customer satisfaction (CSAT) by 35%, and containerizing model functions into reusable Docker components
- Drove a 55% increase in operational efficiency by pioneering the use of univariate or bivariate analysis and machine learning algorithms with TensorFlow and Scikit-Learn, extracting actionable insights from large datasets and fuelling strategic decision-making
- Streamlined data preprocessing workflows using SQL, Python, NumPy, Pandas, Matplotlib, and Seaborn, reducing times by 33% through advanced exploratory data analysis (EDA) and the development of interactive dashboards
- **Domain & Skills:** Customer Service Sector, Anomaly Detection, AWS, Statistical Analysis, Hypothesis Testing

Righteous Technologies — **Hyderabad, India**
Data Scientist — September 2020 – August 2022

- Led a $765K revenue surge, a 20% increase, by architecting end-to-end machine learning systems using Scikit-Learn to refine pricing strategies and implementing time series forecasting techniques that accurately predicted order surges and optimized delivery times
- Created 25+ ETL pipelines with SQL, Python, NumPy, and Pandas and designed 20+ interactive dashboards in Tableau to visualize key performance indicators (KPIs), leading to a 23% improvement in decision-making processes
- Constructed clustering and random forest models to precisely identify at-risk customer segments and implemented tailored product recommendations driven by historical order data, effectively reducing customer attrition by 10%
- **Domain & Skills:** eCommerce, Pharma, Trends, Churn Analysis, Monitoring KPIs, Demand Forecasting

Tata Motors — **Kakinada, India**
Data Analyst (Intern) — May 2019 – July 2019

- Developed a sophisticated demand forecast model to predict 3-year vehicle demand, automating data mining and data cleaning processes, leading to a 50% reduction in data prep time and highlighting 6+ KPIs to monitor supply and demand variables

- Presented data-driven prediction summaries using Matplotlib, showcasing a 15% revenue growth potential; influenced senior leadership's strategic review and decision-making for the upcoming fiscal quarter
- **Domain & Skills:** Automotive Industry, Quantitative Analysis, Regression Modeling

Education

University of Maryland (A. James Clark School of Engineering) — **College Park, MD**
Master of Engineering(M.Eng) in Robotics | GPA 3.83/4.00 — August 2022 – May 2024
Coursework: Machine Learning, Deep Learning, Artificial Intelligence, Computer Vision, Reinforcement Learning, Programming

Vellore Institute of Technology — **Vellore, India**
Bachelor of Technology in Mechanical Engineering | GPA 8.69/10.00 — July 2016 – September 2020

Technical Projects[Link]

Natural Language Processing (LLM) based Chatbot, a user-friendly News Insight Retrieval Tool (OpenAI API, Vector database, Streamlit)
- Developed a bot enabling effortless information retrieval by loading URLs or text files, processing content with **LangChain** and **OpenAI** embeddings, and leveraging **FAISS** for swift and effective similarity search, resulting in a **90% reduction** in reading time

PoseFusion: Multi-View Human Pose Integration for Comprehensive Action Recognition (Keras, TensorFlow)
- Leveraged Graph Convolutional Networks (**GCN**) and **Transformers** to aggregate semantically rich partial information from multiple views of human poses, resulting in robust recognition, achieving up to **82% accuracy** and **82% F1 score** on the occluded dataset

Early Wildfire Detection and Segmentation using Deep Learning (OpenCV, TensorFlow)
- Conceptualized and incorporated a wildfire detection system using a **customized Xception Lite** network and **U-Net** segmentation, achieving **81.91%** accuracy and further boosting accuracy, precision, and recall **metrics** by **15%** using **Transfer Learning**

Real-Time Traffic Light Detection System (OpenCV, Roboflow, PyTorch)
- Developed and implemented a **YOLOv8**-based traffic light detection system, achieving **95% accuracy** and a mean average precision (**mAP**) of **92%**, surpassing pre-trained model benchmarks, setting a foundation for integration into self-driving cars

Exploratory Data Analysis of YouTube Channel, COVID-19, and Car Sales (Python, Pandas, NumPy, Matplotlib, Plotly, Scikit-Learn)
- Applied **data mining** and data analytics to extract and interpret data, contributing to the **30% growth** of the YouTube channel, a **20% increase** in the precision of **COVID-19** infection spread predictions, and a **16% boost** in **sales** revenue in the Indian car market