Jayanth Dasamantharao

Data Scientist & Machine Learning

Piscataway, NJ • (848) 313-5617 •
jayanthdasamantharao@gmail.com •
LinkedIn • Github • Portfolio

Available to work immediately • Open to relocation within the US

TECHNICAL SKILLS

Programming Languages	Python (Problem-solving) | C++ | MATLAB | R (Statistics) | SAS | Scala.
Machine Learning (ML)	Regression Models | Classification Models | XGBoost | Gradient Boosting Methods | Ensemble Models | Support Vector Machine | Time Series Analysis.
Deep Learning (DL)	BERT | NLP | Large Language Models | Word/Image Embedding Techniques | LSTMs, GRU | Gen AI | Retrieval-Augmented Generation | Langchain | Prompt Tuning | GPU Architecture.
Computer Vision	Object detection | Image classification | Semantic segmentation | Neural Networks | ADAS, Data Augmentation | Vision Transformers | Image Registration.
ML and DL Tools	Pytorch | TensorFlow | Keras | HuggingFace | OpenCV | NLTK | Spacy | Numpy | Pandas | Matplotlib | Seaborn | scikit-learn/sklearn | Scipy | CUDA | XGBoost | H2O.ai | Plotly.
Data Engineering & cloud	SQL | PostgreSQL | PySpark | Hive | Hadoop | Azure | Amazon Web Services | Google Cloud Platform | Data Warehousing | Kubernetes | MLOps | Airflow | Robotic Process Automation.
Data Visualization & other tools	Tableau | Power BI | Hypothesis testing | Microsoft Excel (Vlookups, Power Pivot, Macros & VBA) | JIRA | CUDA | Git | GitLab | Stata | Google Analytics | Salesforce | Marketing Cloud.

WORK AND RESEARCH EXPERIENCE

AI Engineer, EliteUS Software Solutions - Fedway, Piscataway, NJ **Aug 2024 – present**

Developing an **AI chatbot** for Fedway system instructions using **NLP and AI**, integrating **text and image data processing**. Leveraged **Language models** from **AWS Bedrock, Azure** for **AI deployment** and **Streamlit** for dynamic interaction.

- Developed a Streamlit-based prototype chatbot integrated with **Chroma, vector databases,** utilizing **PyMuPDF** for PDF manipulation, **Hugging Face Sentence Transformers for NLP embeddings**.
- Deployed AI models using **AWS Bedrock with LLaMA 3.1v** for response generation and **Pillow for image processing**.
- Implemented a similarity matching algorithm using the **CLIP model** to compare **text and image embeddings**, utilizing **NumPy** for handling similarity matrix operations.
- Displayed matched text-image pairs, aligning image paths to corresponding textual instructions based on **cosine similarity**.
- **Tech Stack**: Python, AWS, Azure, Llama, OpenAI, Hugging face, Artificial Intelligence, Model Deployment, NLP.

Data Science Intern, Harvest Software Solutions, LLC, Jacksonville, Florida **June 2023 – Sept 2023**

Developed and implemented strategies using large-scale datasets for **optimizing real-time ad campaigns** which involved leveraging **H2O's AutoML, XGBoost, K-fold cross validation** to enhance performance by 70% and predictive accuracy by 40%.

Utilized Aniview **APIs for real-time data extraction** and managed a dataset of over 5 million rows. Conducted **Exploratory Data Analysis (EDA) using python**, integrating CPM metrics for detailed performance evaluation and actionable insights.

- Employed **hypothesis testing, statistical analysis** to gain insights for data-driven decisions and optimize ad spend.
- Developed **predictive models** to minimize error, using cross-validation to ensure robustness and prevent overfitting.
- **Deployed models into production** using **FastAPI** for predictions & used H2O's AutoML for model selection, development.
- Applied advanced natural language processing techniques using **Hugging Face Transformers** to analyze textual data from ad campaigns for **sentiment analysis and topic modeling** to understand customer engagement and feedback.
- **Tech Stack:** Python, Data Analysis, Hypothesis Testing, Machine Learning, Predictive Modeling, Model Deployment, NLP.

Database Analyst, Accenture Solutions Pvt. Ltd., Hyderabad, India **June 2021 – Aug 2022**

Optimized **ETL processes** to cut support calls by 50%, and created **automated data pipelines with python**, boosting **efficiency by 40%**. Streamlined **data warehousing and DDL updates**, ensuring seamless data integration, performance with **SQL and Tableau**.

- Utilized **Amazon Athena and Redshift** for efficient querying and analysis of datasets over 50 million rows, leveraging serverless SQL capabilities and **scalable data warehousing**.
- Optimized **ETL workflows**, reducing support calls by 50% and enhancing data processing efficiency by 40% through **automated pipelines**. Optimized **DDL updates**, improving data management & system reliability by 15%.
- Conducted **data analysis** and reporting with **Tableau and PowerBI**, facilitating actionable insights and business intelligence.
- Designed and executed complex **SQL** queries using SQL Workbench for detailed **data manipulation** and analysis.
- Coordinated cross-functional teams for seamless **deployments on production servers**, ensuring data integration.
- Enhanced system performance through rigorous **testing, tuning, and monitoring** to ensure reliable operations in production.
- **Tech Stack**: Python, Automation, AWS(Athena, Redshift, S3, Sagemaker), ETL, JIRA, SQL, Tableau, Data Warehousing.

Research Assistant, Andhra University, Visakhapatnam, India **Jan 2021 – May 2021**

Contributed to the creation and deployment of self-driving systems with autonomous navigation, lane tracking, and obstacle detection capabilities, emphasizing improvements in driving intelligence through **deep learning and computer vision** methods.

- Developed and optimized **Convolutional Neural Networks** for lane detection and object recognition. Leveraged architectures like **YOLO** and **Faster R-CNN** to enhance accuracy and speed of real-time object detection and classification.
- Applied image preprocessing techniques like **Gaussian blurring and color space conversion**, along with data augmentation, to enhance dataset quality and improve model generalization across various driving scenarios.
- Applied the **Hough Transform** for precise lane detection and vanishing point tracking to estimate road curvature, enhancing vehicle navigation. Integrated **multi-sensor data** from radar, lidar, and cameras to improve neural network accuracy.
- Applied **Neural Network Regression** for precise vehicle positioning and localization. Developed a real-time simulation environment using **Unity3D** with a **Flask backend** for dynamic interaction.
- **Tech Stack:** Python, Neural Networks, Computer Vision, Image Processing, Data Augmentation, Simulation, Flask.

EDUCATION

Master of Science in Data Science (Statistics), Rutgers University, New Brunswick, NJ **Aug 2022 - May 2024**

Coursework: Statistical Modeling and Computing, Statistical Learning, Probability and Statistical Inference, Neural Networks, Machine Learning, Data Structures and Algorithms(Python), Data Wrangling (R), Natural Language Programming, Data Mining.

School of Arts and Sciences | Research Assistant | Statistics Dept. **July 2024 - Present**

- Extracted & aggregated data from various lawyer websites to build a dataset, used for machine learning model development.
- Implemented machine learning models on the extracted data and developed efficient pipelines for automated data processing and integration. This approach optimized data flow and enhanced the overall performance of the machine learning systems.
- **Skills:** Python, BeautifulSoap, Machine Learning, Puppeteer Node.js, Data Wrangling, Pandas, ETL, RESTful APIs.

Bachelor of Technology in Electrical and Computer Engineering, Andhra University, India **June 2017 - June 2021**

PROJECTS

[Bayesian Networks, PyMC3, Informative Prior Distributions] Developed Bayesian logistic regression models using PyMC3 to predict diabetes status, incorporating informative priors. Evaluated performance with recall and f1-score metrics, comparing outcomes between uniform and normal distribution priors. [Bayesian Classification on diabetes data]

[Hugging Face transformers, Naive Bayes, Logistic Regression, BERT, N-Grams, GloVe embeddings] Developed using Naive Bayes, Logistic Regression, and DistilBERT, achieving 89% baseline accuracy globally. Enhanced precision to 94% with DistilBERT, incorporating n-grams analysis. [Language Identification]

[ML models - Random Forests, KNN, Logistic Regression, SVM, Decision Tree, and XGBoost] Executed diverse ML strategies and time series analysis for auto insurance fraud, prioritizing reduced false negatives. Chose K-Nearest Neighbors (KNN) with an impressive 97% recall. [Auto Insurance Fraud Detection]

[PostgreSQL, MongoDB, Python, Sorting techniques, Search Engine Optimization, Django] Created a user query-responsive tweet retrieval system by integrating dual databases: Postgres and MongoDB. Optimized data storage and retrieval efficiency by 35% while implementing advanced caching strategies for superior system performance. [Twitter Search Application]

LEADERSHIP

- Served as VP Membership at Toastmasters International (2018-2021). Analyzed trends to improve retention, led membership drives, and implemented programs to enhance participation and inclusivity.
- Mentored a candidate at Accenture, providing guidance, sharing expertise, and fostering their professional growth, resulting in enhanced technical skills and career advancement.
- Led cross-functional teams for seamless project delivery, managed end-to-end deployments to ensure timely and efficient project delivery, and resolved conflicts to maintain a productive environment.