

Sandeep Polavarapu Venkata Naga

• Hyattsville, MD • sandeepvvn@gmail.com • +1(240)906-7966
• [linkedin.com/in/sandeepvvn/](https://www.linkedin.com/in/sandeepvvn/) • github.com/sandeepvvn • [Portfolio](#)

PROFESSIONAL SUMMARY

Data Scientist with 4+ years of experience, proficient in developing data-intensive applications and end-to-end pipelines, leveraging expertise in data analysis, machine learning, NLP, deep learning, data engineering, feature engineering, model development, deployment, monitoring, and continuous training using Python, R, SQL, PyTorch, Tableau, and AWS.

EDUCATION

University of Maryland – College Park, MD Masters in Machine Learning	Aug 2021 - Dec 2022 GPA: 3.93
PES University - Bangalore, India Bachelors in computer science and engineering (Minor in Management Studies)	Aug 2014 - July 2018 GPA: 8.31

EXPERIENCE

Craftle, Inc - Detroit, MI <i>Machine Learning/Data Scientist</i>	Jan 2023 - Present
<ul style="list-style-type: none">Lead data science initiatives' design, development, and deployment, collaborating with UI and business using Agile practices.Developed an end-to-end, modularized Generative AI application using pre-trained ControlNet (Controllable Stable Diffusion) that enables users to generate a redesigned room from an image and text prompt using python and HuggingFace.Built a Lookalike finder using computer vision techniques to assist users in finding similar-looking furniture products at a lower price scrapping various websites and APIs, SQL, and NoSQL databases.	
NASA LCLUC Research, University of Maryland - College Park, MD <i>Machine Learning Research Assistant</i>	Jan 2022 - Dec 2022
<ul style="list-style-type: none">Developed Time series crop classification on multispectral satellite images from large data sources (>10TB) using stacked Bi-directional LSTM to study CO2 emissions of sugarcane burning in Thailand, achieving a producer accuracy of 96% and user accuracy of 92%. (Project Link)Implemented a Computer Vision Semantic Segmentation model using DeepLabV3+ to map burnt area in sugarcane residue images, resulting in an IOU (Intersection over Union) score of 77.7% and presented the findings of the above study as a report at the NASA LCLUC conference. (Project Link)Devised a statistical model using SciPy numerical optimization on soil and weather data for winter wheat Phenology dates prediction. This was used to make logistic and farming decisions. (Project Link)	
Akamai Technologies - Bangalore, India <i>Machine Learning Engineer 2</i>	July 2020 - July 2021
<ul style="list-style-type: none">Developed and evaluated Anomaly Detection ML pipelines using A/B testing on Decision Trees, SVM, XGBoost, CNN-LSTM, and Autoencoders to detect and mitigate DDOS (Distributed Denial of Service) attacks. Deployed models with Docker, AWS ECR, SageMaker, and Kinesis streaming. Monitored data drift and performed continuous training, reducing critical event reaction time by 30% and achieving an 82% recall score.	
Machine Learning Engineer	July 2018 - July 2020
<ul style="list-style-type: none">Engineered ETL pipelines, developed and deployed a secure NLP-based Topic modeling Document Information Retrieval system using LDA (dimensionality reduction), K-Means clustering and knowledge graphs on hybrid cloud (AWS S3). This resulted in : (a) Semantic Similarity score 0.71, (b) MAP score 0.5, (c) Click Time Score 0.2.	

SKILLS

Languages:	Python, R, C, C++, MATLAB, Linux(Bash), SQL, NoSQL, JavaScript, Java
Machine Learning:	Scikit-learn, Pandas, NumPy, SciPy, NLTK, spacy, OpenCV, AWS Sagemaker, Classification, Regression
Data Visualization:	Tableau, Matplotlib, ggplot2, R-shiny, plotly, ArcGIS, seaborn, Tensorboard.
Data Engineering:	JIRA, Docker, CI/CD Jenkins, Git, Flask API, GDAL, Rasterio, GlueETL
Deep Learning:	Keras, TensorFlow, PyTorch, HuggingFace, Transformers, GAN, VAE, Diffusion models
Cloud/Big Data:	AWS, (S3, Lambda, Kinesis, Glue, EC2, ECR, Redshift), Azure, GCP, PySpark, Salesforce, MongoDB