Nanda Vihari Mylavarabotla

📍 Hayward, California | 📞5103053064 | 📧 msn.vihari17@gmail.com | 🔗www.linkedin.com/in/sainandavihari | 💻 GitHub

SKILLS

Programming & Data: Python | R | SQL | Azure Cloud
ML & Statistics: Supervised & Unsupervised Learning | Feature Engineering | Regression | Classification | Cross-Validation | Hyperparameter Tuning | Model Interpretation | A/B Testing | Anomaly Detection
Frameworks/Libraries: Pandas | NumPy | Seaborn | Matplotlib | Statsmodels | Scikit-learn | XGBoost | LightGBM | TensorFlow | Keras | PyTorch | OpenCV | Mediapipe | YOLO | Spacy | NLTK
NLP & LLMs: Text Preprocessing | Embeddings (BERT/GPT) | Fine Tuning | RAG | Transformers
Deep Learning: CNN | LangChain | CrewAI | Agentic AI Systems| Computer Vision
MLOps & Data Engineering: MLflow | CI/CD Pipelines | Power BI

PROFESSIONAL EXPERIENCE

Data Analyst

Oatmeal AI Aug 2025 - Present

- Collected and integrated agricultural data from APIs into Azure SQL Database, achieving 99% data accuracy after cleaning and transforming data using Python.
- Designed dashboards and reports to track agricultural KPIs, enabling faster and more accurate business insights.
- Collaborated with cross-functional teams to translate requirements into actionable analytics, improving operational efficiency.
- Resolved data quality issues in AI chatbot systems, improving response precision and reducing error rates by 15%.

Data Analyst

Capgemini Technology Services Limited Mar 2022 - Dec 2023

- Generated automated data pipelines using Python, SQL, and Pandas to process 500K+ records daily, cutting manual analysis time by 80% and improving data accuracy to 99.5%.
- Developed forecasting models and Power BI dashboards with CI/CD integration, improving forecast accuracy by 35% across five major product lines.
- Created a marketing performance prediction system using Random Forest and XGBoost, reaching 92% accuracy across 15+ digital channels.
- Launched A/B testing and optimization tools that boosted marketing ROI by 40%, saved $300K annually, and increased conversion rates by 25%.

RELEVANT PROJECTS

Telecom Customer Churn Prediction

- Built a customer churn prediction pipeline with EDA, descriptive statistics, and feature selection, enhancing data quality and generating actionable insights.
- Balanced the dataset using SMOTE and developed models including Decision Tree, Random Forest, Gradient Boosting, and XGBoost to ensure fair evaluation across churn/non-churn classes.
- Optimized models with GridSearchCV, 10-fold cross-validation, MLflow, and fine-tuned an ANN using the Adamax optimizer, achieving 86.1% accuracy and 0.83 F1-score—a 17% improvement over baseline.

Transfer Learning for Image Classification

- Applied transfer learning with TensorFlow and Keras using pretrained models including ResNet50, ResNet50V2, VGG16, VGG19, Xception, and EfficientNetV2 on the ImageNet dataset.
- Performed data augmentation with ImageDataGenerator to expand the dataset and improve generalization.
- Found that Xception and VGG16 consistently outperformed other architectures, achieving 98% accuracy on classification tasks.

Text Generation App using Google Gemini AI

- Developed an interactive Streamlit web app integrating Google Gemini AI API for real-time text generation, reducing user response latency by ~40% compared to traditional API calls.
- Implemented a scalable prompt-to-response pipeline, enabling users to generate contextual AI outputs dynamically, improving app usability and engagement by 60% through an intuitive UI.
- Deployed the Gemini-powered LLM chatbot prototype for automated content generation and Q&A tasks, enhancing accuracy and coherence of generated responses by 25% based on user feedback testing.

EDUCATION

California State University – East Bay Jan 2024 - *Present*
Master of Science in Statistics (Data Science)
Jawaharlal Nehru Technological University Aug 2018-May 2022
Bachelor of Engineering in Electronics and Communication