BALA SWAPNIKA GOPI

GitHub | Linkedin | 240-350-7674 | Swapnika29.com | balaswapnika.gopi29@gmail.com

Aspiring Data Scientist pursuing a Master's in Data Science, with strong skills in machine learning, statistical analysis, and data visualization. Proficient in Python, R, and SQL, with hands-on experience analyzing complex datasets and deriving actionable insights. Skilled in leveraging data-driven approaches to address complex challenges. Excited to apply analytical expertise and contribute to meaningful projects through a summer internship.

Education

University of Maryland — College Park, USA
Master of Science - Data Science - CGPA:4.0/4.0 — Aug 2024 - May 2026*

Jawaharlal Nehru Technological University Hyderabad — Telangana, India
Bachelor of Technology - CGPA:7.15/10.0 — Jul 2018 - Aug 2022

Skills

- **Programming Languages:** Python, C, SQL, R
- **Machine Learning:** Clustering, Feature Extraction, Product Analytics, NLP, Generative AI
- **Frameworks:** Pandas, Numpy, Scikit-learn, Pyspark, Pytorch, Seaborn, Spark SQL, Tensorflow, Matplotlib, Keras, NLTK
- **Tools:** Git, VS Code, Spyder, MySQL, JIRA, Tableau, Microsoft Excel
- **Platforms:** Jupyter Notebook, Apache Spark, Docker, Kubernetes, Linux, Unix, Windows, GitHub, AWS
- **Modeling and Analysis:** Causal Inference, Hypothesis Testing, Statistical Modeling, Regression, Classification, Segmentation, Business Analysis, Cross-Validation, Random Forest, Business Intelligence, Mathematics,Statistics, Relational Databases, Data Mining, Data Analytics, Decision Trees
- **Technical Knowledge:** Design and Development, Testing and Debugging, Big Data Technologies, AI Model Development

Professional Experience

Infor — Hyderabad, India
Associate Machine Learning Engineer — Apr 2022 - Aug 2024

- Designed and deployed end-to-end machine learning algorithms on the Coleman AI platform, streamlining model development and improving performance for 15+ enterprise clients across finance and supply chain domains
- Implemented and fine-tuned forecasting models for intermittent demand data, reducing error rates by 20%, and supporting more accurate demand planning decisions. Engineered optimization algorithms using Python and PySpark, achieving a 25% improvement in resource allocation strategies across client operations.

Projects

Facial Expression Recognition; ***Tech:*** *Python, CNN, Google FER dataset*;

- Developed a deep learning model using Convolutional Neural Networks (CNNs) to classify facial expressions with an accuracy improvement of 10% on the Google FER dataset. Demonstrated real-time emotion recognition applications, enabling scalability across industries such as healthcare and customer engagement.
- Employed image preprocessing techniques, including data augmentation and hyperparameter tuning, to reduce overfitting by 15% and enhance the model's generalization capabilities.

Credit Card Fraud Detection; ***Tech:*** *Python, SVM, Decision Tree, Logistic Regression, KNN*;

- Built a predictive model for detecting fraudulent transactions, improving precision by 30% and recall by 25% on an imbalanced credit card dataset. Utilized resampling techniques like SMOTE and feature engineering to enhance class balance and model performance.
- Conducted comprehensive evaluations across 5 different machine learning algorithms for scalable fraud detection, leading to optimized model performance that safeguarded transactions from potential fraudulent activities.

Vision-Language Model (VLM) Safety Evaluation; ***Tech:*** *Python, VLLM Safety Benchmark, GPT-4o, Claude 3, Gemini 2.0*;

- Evaluated leading multimodels including GPT-4o, GPT-4o-mini, Claude 3 Haiku, and Gemini 2.0 using the VLLM Safety Benchmark, achieving 96.52% accuracy on OODCV-VQA, Sketchy-VQA, and weather-dependent datasets to assess robustness in Out Of Distribution and counterfactual scenarios
- Developed strategies to mitigate hallucination issues, enhancing model safety and reliability in vision-language tasks for real-world AI deployment

Achievements and Volunteer Experience

- Won **silver medal** in Programming Data Structures and Algorithms using Python from NPTEL.
- Certified as a **Top Performer** in Programming with Python and Machine Learning from Internshala.
- **Core Member x NSS JNTUHCEJ**, Spearheaded 10+ NSS projects, positively impacting over 1,000 beneficiaries.
- **Event Organizer** for technical events at the University's EMBLAZON FEST - Department of Electronics and communication engineering.