Sai Kiran Reddy Dyavadi

Austin, TX | +1 (979) 985-1632 | sdyavadi@tamu.edu | linkedin.com/in/saikiran | github.com/dyavadi8769

Azure Certified Data Scientist with over 2.5+ years of experience in Machine Learning, Deep Learning, Gen AI and Data Analytics. Proficient in Python, SQL, Cloud Computing, develop AI solutions enhancing data processing efficiency and predictive accuracy

Education

Texas A&M University, Mays Business School — College Station, TX
Master of Science in Management Information Systems, GPA:4.0/4.0 — *Aug 2023 – May 2025*

Birla Institute of Technology and Science, Pilani — Pilani, RJ
Bachelor of Engineering in Electronics & Communication Engineering, GPA:4.0/4.0 — *Aug 2018 – Jun 2022*

Technical Skills

Programming Language: Python, SQL, PL/SQL, NoSQL, Java, C, C++, R, Scala, HTML, CSS, Data Structures
Libraries: Scikit-Learn, Pandas, Numpy, Plotly, Matplotlib, Scipy, Seaborn, NLTK, Spacy, fastText, OpenCV,Transformers
Tools & Frameworks: Tensorflow, Keras, PyTorch, Tableau, Hive, Spark, Hadoop, Docker, Git, Alteryx, Excel, SAS, MLFlow
Mathematics for ML & DL: Probability & Statistics, Calculus, Algebra, Hypothesis Testing, Statistical Modeling, PCA
Databases & Web Apps: MySQL, PostgreSQL, Redis, MariaDB, MSSQL, MongoDB, Flask, Streamlit, FastAPI
Cloud: AWS(S3, EC2, Lambda, ECR, Beanstalk, BedRock, IAM), Azure(ML, Databricks, Synapse, Data Lake, Cosmos)
AI Skills: Machine Learning, Deep Learning, Natural Language Processing,Computer Vision, Artificial Intelligence
Gen AI: Llamma3, GPT-4o, Gemma2, Mistral, Gemini Pro LLMs, Hugging Face, Langchain, RAG, Lamini AI, Groq API
Certifications: Azure Certified Data Scientist, AWS Certified Data Engineer, Alteryx Certified Designer Core, PSM-I
Soft Skills: Communication, Written, Technical, Self-Starter, Innovative, Problem-Solving, Leadership, Coordination

Professional Experience

Graduate Research Assistant — College Station, TX
Texas A&M University — *Aug 2023 – Present*

- Developed 95% accurate Machine Learning model for analyzing student performance, incorporated exploratory data analysis on datasets, containerized with Docker, automated deployment with AWS EC2, ECR, CI/CD, ensured Git version control
- Orchestrated an 84% effective Flask-based web application for predicting university admissions using linear regression algorithms, used Excel for requirements gathering, interpreted results through data visualizations using Matplotlib, Seaborn

Data Scientist — Bangalore, KA
American Express, Credit and Fraud Risk, Finance — *Jul 2022 – Jun 2023*

- Enhanced Machine Learning models by analyzing 962 variables using Python, Excel, SQL, increased model accuracy by 25%
- Optimized data processing efficiency by 23% through cleansing techniques, statistical data analysis and feature selection
- Collaborated with cross functional teams to optimize data models, communicated AI driven recommendations using Tableau
- Accelerated Pandas-based correlation matrix computation for 300 variables using PySpark, reduced from 40 min to 15 min
- Implemented Agile methods, leveraged ETL, used neural networks for data mining, got key metrics, KPIs with 96% accuracy
- Orchestrated fraud, risk models using Scikit-Learn, Spark, Relational Database, improved risk analysis, fraud detection 44%

Data Scientist Intern — Pune, MH
Vodafone, Shared Services — *May 2020 – Jun 2020*

- Developed a Deep Learning based predictive model using hyperparameter tuning and Tensorflow, achieved optimization in accuracy, forecasting a 10% increase in customer demand, and facilitated decision making through interactive dashboards
- Enhanced customer inquiry processing efficiency by 15% through use of unsupervised learning, statistical modeling, data analysis, data visualization, data exploration, effectively addressed operational inefficiencies, improved overall performance

Projects

Speech To Text Conversion | GitHub | *Natural Language Processing, Tensorflow, AWS Cloud, GitHub*

- Automated data pipeline, model deployment using Docker, AWS EC2, S3, and GitHub Actions, improved efficiency by 40%
- Delivered 89% accuracy in text conversion using robust CNN, Transformers, NLP techniques and data processing workflows

Automatic Number Plate Detection | GitHub | *Computer Vision, Tensorflow, YOLOv8, OpenCV, EasyOCR*

- Engineered an automatic number plate recognition system using YOLOv8 and OpenCV, achieved 98% detection accuracy
- Optimized processing speed by 30% using EasyOCR and TensorFlow for real-time text extraction from license plates

Intel Image Classification | GitHub | *CV, Deep Learning, PyTorch, ResNet, Docker, AWS, Data Augmentation*

- Spearheaded the project, achieved 94% accuracy using PyTorch, ResNet, Transfer Learning, CUDA for GPU acceleration
- Enhanced model deployment with Docker and AWS, reduced inference time by 27%, processed 1000 images per minute

Youtube Video Summarization App | GitHub | *Gen AI, LLMs, Llama 2, Haystack, Whisper*

- Leveraged Haystack and Llama 2 LLM for efficient text summarization, reduced video summary generation time by 50%
- Integrated Whisper for audio transcription, achieved 70% accuracy in topic extraction, used Streamlit for app deployment

Gen AI Powered Q&A Chatbot | GitHub | *Gen AI, LLMs, GPT-4, Open AI, Mistral, Ollama, Langchain*

- Developed a Q&A chatbot using GPT-4, 4o, 4 Turbo, Mistral LLMs within LangChain, increased user engagement by 40%
- Enabled user selection among 4 LLMs, integrating OpenAI GPTs and Ollama LLMs, reduced query response time by 25%