Pranit Manda

Dallas, TX | +1 (682) 557-7999 | pranitmanda2000@gmail.com | LinkedIn | GitHub

Professional Summary

Data & Machine Learning professional with hands-on experience designing end-to-end pipelines, deploying scalable AI/ML models, and optimizing data workflows. Skilled in Python, SQL, TensorFlow, PyTorch, Spark, cloud platforms (AWS/GCP) and BI tools (Tableau, Power BI, Excel). Proven ability to deliver production-ready solutions for real-world impact across research and enterprise projects.

Education

M.S in Computer Science - University of Texas at Arlington *Aug 2023 - May 2025*

B.Tech in Electronics and Communication Engineering - Vellore Institute of Technology *Jul 2018 - May 2022*

Experience

UNIVERSITY OF TEXAS AT ARLINGTON Arlington, TX

Research Assistant *Jun 2025 - Present*

- Designed and deployed ML/GenAI models by integrating preprocessing, feature engineering, and algorithm selection, improving accuracy and efficiency.
- Automated Python + Spark pipelines for preprocessing, training, and evaluation, reducing experiment time by 30%.
- Published research findings and presented benchmarks, strengthening visibility of AI in simulation and NLP projects.
- Partnered with faculty and cross-functional teams to align research workflows with project goals, ensuring timely deliverables.

Parking Enforcement Data Analyst *Aug 2024 – May 2025*

- Analyzed large parking datasets to identify demand patterns, compliance trends, and revenue impact, supporting campus policy refinement.
- Built SQL + Power BI dashboards to visualize occupancy, citation frequency, and peak usage, enhancing decision-making.
- Conducted root-cause analysis of violations, driving actionable recommendations that improved compliance strategy.

COGNIZANT TECHNOLOGY SOLUTIONS Hyderabad, India

Data Engineer *Jun 2022 – Aug 2023*

- Engineered end-to-end ETL pipelines and reusable ingestion frameworks with Spark, SQL, and Informatica, boosting scalability.
- Optimized infrastructure and refactored code, improving processing speed by 20% for enterprise ML analytics workloads.
- Deployed ML models via MLOps pipelines, enabling automated monitoring, drift detection, and anomaly detection at scale.
- Partnered with BI teams to integrate Tableau dashboards with ML-driven pipelines, improving business insight delivery.

Big Data Intern *Jan 2022 – May 2022*

- Built Spark + Hive workflows on 4K+ educational datapoints, enabling robust data analysis and batch pipeline automation.
- Optimized Spark queries and Hive tables, reducing query execution time by 25% and improving pipeline efficiency for analytics teams.
- Applied classification models with hyperparameter tuning, improving predictive accuracy and data consistency.

TECHVEDIKA SOFTWARE PVT. LTD. Hyderabad, India

Software Intern *Dec 2020 – Dec 2020*

- Developed TensorFlow + OpenCV CNN models achieving 96% accuracy for medical diagnostics with real-time inference.
- Designed preprocessing and augmentation pipelines, improving image data quality and boosting diagnostic reliability.

Skills

Languages: Python, R, SQL, Java, C++ | **ML/GenAI Frameworks:** TensorFlow, PyTorch, Hugging Face Transformers, Scikit-learn, OpenCV | **Tools & DevOps:** Anaconda, Jupyter, Git, Visual Studio, Docker, CI/CD, Airflow, Kafka, dbt, Kubernetes | **Cloud:** AWS (SageMaker, EC2, S3), GCP (Vertex AI, BigQuery) | **Data Engineering:** SQL Server, Hadoop, Hive, Apache Spark, ETL Pipelines, DataStage, Informatica | **Data Analytics:** Tableau, Power BI, Excel (Pivot, PowerQuery), Google Data Studio, Matplotlib, Seaborn | **Specialties:** Prompt Engineering, LLM Fine-tuning, Transfer Learning, Computer Vision, MLOps, Statistics, Data Preprocessing

Projects

Smart Traffic Optimization using Google Maps API + Deep RL (2025)

- Built a real-time traffic optimization engine using Deep Reinforcement Learning and live Google Maps data.
- Deployed solution on AWS with traffic dashboards; integrated spatiotemporal forecasting for congestion.
- Applied Q-learning and DQN variants to optimize traffic signal timing strategies.

GenAI Chatbot with RAG and LLMs (2024)

- Developed a Retrieval-Augmented Generation system using LangChain and GPT-4 with Pinecone vector storage.
- Integrated document search with LLM for contextual answers; deployed with Streamlit interface.
- Fine-tuned LLMs on custom Q&A datasets for improved domain-specific accuracy.

Plant Disease Detection (2021)

- Built a CNN-based classifier for 15+ crop diseases with high accuracy, leveraging TensorFlow and extensive image preprocessing.
- Optimized and deployed the model with TensorFlow Lite for mobile/edge inference, enabling real-time predictions in field conditions.

Automatic License Plate Recognition (2020)

- Developed a YOLO-based detection system achieving ~96% accuracy for license plate localization and recognition.
- Integrated OCR with the detection pipeline for real-time inference, reducing manual processing and enabling scalable deployment.

Certifications

- Generative AI with Large Language Models: Deeplearning.ai (2024)
- ChatGPT Prompt Engineering for Developers: OpenAI X DeepLearning.ai (2024)
- AWS Machine Learning Specialization: Coursera (2020)
- IBM Data Science Specialization: Coursera (2022)
- Machine Learning with Python: IBM CognitiveClass (2020)
- Google IT Automation Specialization: Coursera (2020)

Leadership/ Involvements

- Digit (9.9 Group): Campus Ambassador
- DigitSQUAD: Founder and Former Chairperson
- Flipkart Grid Hackathon 2.0: Top 20 Finalist
- JP Morgan Chase: SWE Virtual Experience
- Microsoft: Engineering Virtual Experience