Aman Parasher

AI Graduate | Data & ML Engineer | International Researcher

+91 9667509302 amanparasher.ai@gmail.com linkedin.com/aman-parasher github.com/amanparasher

Summary

B.Tech graduate in Artificial Intelligence with research experience in Taiwan on computer vision and machine learning for road safety. Contributed to peer reviewed research publications and gained experience in building production grade data pipelines, AI models, and data workflows through industry internships, national hackathons, and technical certifications.

Education

Bachelor of Technology (B.Tech) in Artificial Intelligence (AI) **Sep. 2021 – Jun. 2025**

CINTEL (Department of Computational Intelligence), SRM University, India *CGPA 8.2/10*

Experiences

Data Engineer Intern **Jul. 2025 – Present**

Mactores *Remote (India)*

- Trained extensively on Apache Spark, PySpark, Spark SQL, and Databricks through guided learning resources, hands-on labs, and an independent end-to-end ETL project executed on the Databricks platform.
- Earned Databricks Data Engineer Associate and Professional certifications, and successfully passed the Snowflake SnowPro Core certification.

International Research Intern **Dec. 2024 – Mar. 2025**

National Chung Cheng University,Taiwan (Taiwan India Joint Research Centre on AI) *Chiayi, Taiwan*

- Engineered a real time vehicle turn signal violation detection system using computer vision and ML algorithm.
- Applied YOLO-based models for fast and accurate object detection on video frames.
- Integrated SegNet and CNN-LSTM architectures for precise light masking and temporal signal analysis, achieving an F1-score of 0.88.
- Developed a full pipeline to analyze 1,000+ vehicle videos, generating detailed violation reports with labeled images and timestamps to enhance road safety monitoring.

Projects

AI Document Theme Chatbot | *RAG, LangChain, LLM, FastAPI*

- Engineered an end-to-end Retrieval-Augmented Generation (RAG) system to query 75+ PDF/scans for theme-based summarization.
- Utilized LangChain, OCR, and FAISS for vector search and extraction across 1000+ paragraphs with citation capability.
- Achieved 85% relevance in multi-document answers, deployed via Streamlit/FastAPI with <2.5 sec average query speed.

Ride Hailing Lakehouse Analytics | *Databricks, PySpark, SQL, MLflow*

- Built a Medallion Lakehouse pipeline processing 3.06M NYC taxi records using PySpark/Spark SQL.
- Cleaned ∼ 90K bad records and engineered features incl. trip_duration + Haversine distance.
- Trained Prophet (12.5% MAPE) and Isolation Forest to predict demand + flag top 1% anomalous trips.
- Tracked 15+ experiments with MLflow and delivered dashboard of 5,000+ pickup hotspots.

Technical Skills

Cloud & Data Platforms: Azure, AWS, Databricks, Snowflake, Apache Airflow, Kafka

Programming & ML: Python (PySpark, Pandas), SQL, TensorFlow, PyTorch, Computer Vision, LaTeX

Databases & BI Tools: PostgreSQL, MongoDB, PowerBI, Git

Research Publications

Risk Assessment of High Growth AI Stocks Using Time Series Analysis *25th Annual Convention of Strategic Management Forum, Indian Institute of Management (IIM) Tiruchirappalli* (Accepted 2025)

Multi class Ship Classification of Commercial and Naval Vessels Using Convolutional Neural Network *International Conference on Smart Trends for Information Technology and Computer Communications (Springer Nature Singapore)*, pp. 439–455, 2025