AMIT MANJARLY

amitmanjarly99@gmail.com | +1-623-275-0326 | LinkedIn | My Website

SUMMARY

Machine Learning Engineer with **3.5 years** of professional industry experience in data engineering and algorithm optimization. Proven track record in developing **LLM fine-tuning strategies**, **RAG pipelines**, and **hybrid recommendation engines**. Proficient in Python, PyTorch, and deploying scalable solutions using Docker and Kubernetes.

SKILLS

ML Techniques: LLMs (Llama-3.2), RAGs, Prompt Engineering, CNN, RNN, Reinforcement Learning, Bayesian Networks
ML Models: Logistic Regression, Gradient Boosting Trees (XGBoost/LightGBM), Deep Neural Networks (DNN), Wide & Deep Models.
Recommendation & Embedding Techniques: Collaborative Filtering, Matrix Factorization (SVD), Factorization Machines, Word2vec.
Frameworks & Libraries: PyTorch, TensorFlow, Scikit-learn, NumPy, Pandas
Infrastructure & DevOps: Docker, Kubernetes, Git, Apache Kafka
Performance & Efficiency: CUDA, Profiling, Convex Optimization, Algorithmic Efficiency, 4-bit Quantization
Languages: Python (Advanced), SQL, C++, Java
Databases & Big Data: Hive, Impala, Oracle, Neo4j, Sparse User-Item Matrices
Web & Visualization: JavaScript, HTML, CSS, Tableau
Platforms: Jupyter Notebook, Google Colab, Kaggle

EXPERIENCE

Accenture PLC Hyderabad, India
Data Engineer / ML Infrastructure Analyst March 2021 – July 2024

- Engineered and refined **Python** scripts for core algorithms, achieving a **15% reduction** in total processing time.
- Optimized complex **SQL workflows** to reduce data processing latency by **33%** for international banking systems.
- Performed **data analysis** using **Tableau** and generated reports to effectively communicate work strategies.
- Debugged critical software issues using **qTest** and enhanced the system stability on international banking project.
- Co-ordinated with other teams via **JIRA** and **KANBAN** boards to consistently meet deliverable deadlines.
- Participated in **User Acceptance Testing (UAT)** to ensure client expectations were fulfilled in the deliverables.
- Developed an **automated data-testing framework** using Python and SQL, reducing testing overhead by **40%**.
- Administered large datasets on Hive and Impala, utilized Toad for Oracle to optimize query execution & result extraction.

Terminal Trend Ahmedabad, India
Software Engineer - Intern January 2021 – February 2021

- Progressed from foundational to advanced level for front-end development resulting in better proficiency in **JavaScript**.
- Developed **front-end UI** games using HTML, CSS and JavaScript gaining in-depth knowledge of core concepts.
- Contributed to the deployment of a **live web application**, ensuring seamless client collaboration and integration.

PROJECTS

LLM Fine-Tuning for Enhanced Mathematical Reasoning (STaR) (Github) September 2025 - October 2025

- Enhanced **Llama-3.2-3B LLM mathematical reasoning** via strategic fine-tuning and **GSM8K** benchmarking.
- Built Self-Taught Reasoner(STaR) pipeline using prompt engineering to generate and refine bootstrapped training data.
- Implemented **4-bit quantization** and **float16** to enable memory-efficient fine-tuning of a 3B parameter model.
- Achieved **47.56%** accuracy with STaR-SFT, outperforming Zero-Shot-CoT and Vanilla-SFT **(44.00%)** baselines.

End-to-End Hybrid Movie Recommendation Engine (Github) May 2025 – June 2025

- Built a hybrid recommender using Collaborative Filtering and content-based methods, increasing Precision@10 by 16x.
- Leveraged **Word2vec embeddings** to resolve the cold-start problem through semantic metadata analysis.
- Leveraged a **Wide & Deep** model to integrate **Logistic Regression memorization** and DNN generalization.
- Resolved the **cold-start problem** for new users via content based suggestion engine which enhanced scalability.
- Benchmarked **Gradient Boosting** and **Factorization Machines** to optimize ranking and computational efficiency.

Image Classification with CNNs and SVMs (Github) September 2024 - November 2024

- Developed a **custom CNN architecture** for image classification on CIFAR-10 dataset using **PyTorch**.
- Optimized model training performance using **Stochastic Gradient Descent (SGD)** and **ADAM** optimizers.
- Engineered an end-to-end **Scikit-learn pipeline** for automated preprocessing, training, and evaluation.
- Improved feature space efficiently by applying **Principal Component Analysis (PCA)** for dimensionality reduction.

EDUCATION

Arizona State University Tempe, Arizona, USA
Master of Science in Data Science, Analytics and Engineering August 2024 – May 2026

Gujarat Technological University Ahmedabad, India
Bachelor of Engineering in Computer Engineering August 2015 – June 2019