SUCHIR MATHUR

Machine Learning Engineer

A results-driven and innovative Machine Learning Engineer with **over 4 years of experience** specializing in **Large Language Models (LLMs), Computer Vision, and Predictive Analytics**. Seeking to leverage expertise in **developing end-to-end AI solutions, fine-tuning LLMs, and applying advanced techniques** like **Retrieval-Augmented Generation (RAG)** to solve real-world business challenges. Eager to contribute technical skills in **deep learning, automation & data analysis** to drive impactful, data-driven decisions in a dynamic organization.

(+91) 836-971-3175

suchir.mathur@gmail.com

English & Hindi

Mumbai, India

March 04, 1995

Profile Summary

- **Experienced Machine Learning Engineer** with expertise in Large Language Models (LLMs), Computer Vision, and Predictive Analytics.
- **Proven track record in developing AI solutions** for document processing, synthetic data generation, and business intelligence, driving automation and efficiency across industries.
- **Skilled in fine-tuning LLMs and implementing advanced techniques** like Retrieval-Augmented Generation (RAG) to provide real-time insights & improve model performance.
- **Experienced in end-to-end project management**, from scoping and estimation to the finalization of technical specifications and quality management for AI projects.
- **Expertise in leveraging OCR, Named Entity Recognition (NER), and computer vision** to extract valuable insights from unstructured data, optimizing business workflows.
- **Successfully developed predictive analytics models** to target customers for specific product offerings, enhancing business decision-making with data-driven insights.
- **Adept at designing automation pipelines** that streamline complex processes, including background verification, document classification, and data validation.
- **Strong background in deep learning technologies**, applying neural networks and machine learning algorithms to solve real-world challenges in a variety of industries.
- **Passionate about innovation,** with a deep understanding of generative AI, LLMs, NLP, and computer vision techniques to drive breakthroughs in AI application across domains.

Work Experience

➤ Jul 2024 - Present | Unify Technologies Pvt Ltd (Contingent Worker with Meta/Facebook) | Machine Learning Engineer

- Orchestrating the generation of synthetic data for audio transcription models utilizing LLMs, curating diverse conversational datasets to significantly enhance transcription accuracy.
- Engineering an LLM agent-driven evaluation system, "LLM-as-a-judge," to meticulously assess the quality, relevance, and safety of synthetic data, thereby fortifying data integrity through advanced evaluative mechanisms.
- Designing and deploying a multimodal agent-based architecture for the evaluation of outputs from an image generation service, resulting in measurable improvements in conversion rates, click-through rates, and overall user retention.
- Refining toxicity and document classification models through the precise fine-tuning of LLMs on toxic comment datasets, thereby optimizing classification precision for critical organizational use cases.
- Developing and piloting a Proof-of-Concept (PoC) utilizing LLM agents to generate synthetic data for wearable technology applications, enhancing user experience and accelerating adoption during preliminary user trials.

➤ Sept 2022 - Jul 2024 | AuthBridge Research Services Pvt Ltd | Machine Learning Engineer

- Spearheaded the development of LLM-powered document entity extraction solutions, successfully automating background verification and student record validation, resulting in substantial reductions in manual processing efforts.
- Engineered automated email verification systems using meticulously crafted LLM prompts, achieving a tenfold reduction in processing time while maintaining exceptional accuracy in credential validation.
- Formulated an advanced RAG-based business intelligence model, facilitating real-time analysis of company earnings calls and providing decision-makers with actionable, automated insights.
- Developed a high-accuracy object detection & masking solution for Personal Identification Number masking in documents, streamlining KYC verification processes for commercial banks.
- Constructed an OCR and NER-based National ID extraction system, deploying a scalable Flask API that standardized client verification protocols across multiple industries.
- Devised image classification pipelines to identify document sources and languages, optimizing preprocessing workflows and reducing computational overhead for LLM applications.

➤ Apr 2021 - Sept 2022 | Balajikripa Agro-Processors Pvt Ltd | Machine Learning Engineer

- Created an energy price forecasting model that optimized onsite energy production, successfully reducing energy imbalance costs and enhancing renewable energy utilization.
- Architected and deployed a deep learning-based plant and crop disease classification system, facilitating automated health assessments and driving improvements in agricultural productivity.

Core Competencies

Machine Learning Model Development & Deployment

Large Language Model (LLM) Applications & Optimization

Generative AI Solutions & Synthetic Data Creation

Predictive Analytics for Business & Operational Insights

Natural Language Processing & Information Extraction

Computer Vision for Image & Document Analysis

OCR & Automated Document Processing

AI-Based Background Verification & Identity Checks

Deep Learning Model Building for Varied Use Cases

Process Automation Using AI-Driven Tools

API Development & System Integration for AI Services

Real-Time Business Reporting & Decision Support

Compliance Automation for KYC & Credential Validation

Data-Driven forecasting for Energy, Agriculture & Other Sectors

Project Planning, Stakeholder Collaboration & Solution Delivery

Technical Skills

- **Programming & Databases**: Python, MySQL
- **ML & AI Frameworks**: PyTorch, OpenCV, Pandas, Hugging Face Transformers, LangChain, Scikit-learn, Matplotlib, Flask, SpaCy
- **Cloud & Platforms**: Google Colab, Google Cloud Platform (GCP), Vertex AI
- **Software & System Design**: Data Structures & Algorithms, API System Design
- **Specialized AI Techniques**: Large Language Models (LLMs), Generative AI, Deep Learning, Retrieval-Augmented Generation (RAG), Computer Vision, Natural Language Processing (NLP), Named Entity Recognition (NER), Text Classification, Sentiment Analysis, Time Series Forecasting, Language Translation

Education & Credentials

- **B.Tech. in Computer Science and Engineering** | VIT Vellore | 2014-18
- **Deep Learning Specialization certification** | Coursera

Soft Skills

Collaborative Teamwork | Advanced Time Management | Exceptional Interpersonal Skills | Leadership Competence | Effective Decision-Making | Analytical Problem-Solving | Critical Thinking