Pratik **Mehta**

Profile Summary

Experienced *ML engineer* with over 4 years of expertise. Proficient in deep learning foundations and building GenAI-based backend systems. Strong communicator with a track record of working across disciplines with diverse teams.

Contact details

✉ mehtapratik@outlook.com

📞 +91 9082150282

Personal information

Citizenship: **Indian**

Languages: **English**, **Hindi**, **Marathi**, **French (A1)**

Skills

- Prog. languages: python, rust
- Data science libraries: pandas, NumPy, seaborn, scikit-learn, pytorch
- LLMs: GPT-x, llama
- Backend: Postgres-SQL, FastAPI, langchain, pandas-ai
- CLI: git, vim, std. unix programs
- Docker, Github actions

Experience

Consultant, Machine Learning & GenAI at *Deloitte Digital USI - Mumbai* — **Nov 2020 – Jun 2024**

◇ Successfully designed & developed REST APIs using FastAPI, PostgreSQL, and Llama served over AWS to answer user questions about live quantitative enterprise KPI data for Beta testing by a few thousand users

◇ Initiated driving deep changes to the in-house black-box testing practice for LLM-in-the-loop applications for use across clients

◇ Developed use-cases and AI-adoption strategies for AI-aspirant enterprise clients, and conducted tutorial sessions for improving ML literacy

Research Engineer at *IIT Bombay - Mumbai* — **Aug 2015 – July 2016**

◇ Trained and deployed phrase-based statistical machine translation systems for English-Marathi and English-Indonesian using medical-domain textbooks

◇ Designed an error analysis engine to identify and classify syntactic errors in translations, and developed a greedy algorithm for automatic post-editing, evaluated via BLEU scores and human evaluators

Research Assistant at *Information Sciences Institute, University of Southern California - Waltham, MA* — **June 2018 – Oct 2018**

◇ Studied correlation between different evaluation metrics for machine-translation systems and their downstream impact on cross-lingual search applications

Education

Master of Science in Computer Science (MS)
University of Massachusetts Amherst — **2016–2019**

Bachelor of Engineering in Computer Engineering (BE)
DJ Sanghvi College of Engineering, University of Mumbai — **2011–2015**

Projects

Sony Persona-Grounded Commonsense Reasoning challenge - API track | *AIcrowd* — **May 2025**
◇ Developed agents and finely-tuned 'system prompts' to improve 'API and parameter choice' and function-calling accuracy for NPC agents interacting with a videogame player

Alfred: LLM-powered dispatch of hyper-personalized batch notifications | *llama* — **Jan 2024**
◇ Conceptualized and developed a proof-of-concept system to track users' plain-language notification preferences and notify them based on live-data events in the application

LegalEase: Cross-lingual semantic search engine | *cohere, qdrant* — **Mar 2023**
◇ Developed an application that answers questions factually based on lengthy legal documents (now called Retrieval Augmented Generation) ◇ Also developed a multi-lingual semantic search engine for any-language-query to any-language-document retrieval