

Pratik Mehta

☎ (+91) 908-215-0282 | ✉ pratikmehta1494@gmail.com | 📱 pratikmehta14 | 💻 pratikmehta14 | 🐦 @pratikmehta94

Experience

Deloitte India (Offices of the US)

Mumbai, India

TECH CONSULTANT

Nov 2020 - Present

- As back-end engineer on a new engagement - conducted POCs, helped shape engg. practices, and shipped code to production
- Set up a micro-service based application from scratch for performing RAG over structured enterprise data for a client
- Previously worked on a microservice-based application for document classification and information extraction

Information Sciences Institute - University of Southern California

Waltham, MA

VISITING RESEARCH ASSISTANT

Jun 2018 - Oct 2018

- Studied correlation between machine translation (MT) and downstream Cross-lingual Information Retrieval (CLIR) system performance

Center for Indian Language Technology - Indian Institute of Technology Bombay

Mumbai, India

RESEARCH ENGINEER

Jun 2015 - Jul 2016

- Trained and deployed English-Marathi and English-Indonesian machine translation systems for translation of medical domain text
- Developed an error analysis engine to identify and classify syntactic errors in translation, and a greedy algorithm for automatically post-editing translations
- Co-compiled the (then) largest of its kind parallel text corpus called the IIT Bombay English-Hindi parallel corpus

Education

University of Massachusetts Amherst

Amherst, MA

M.S. IN COMPUTER SCIENCE (GPA: 3.21)

2016 - 2019

Reinforcement learning, Data visualization, Deep Learning, Ethical considerations in CS

D. J. Sanghvi College of Engineering, University of Mumbai

Mumbai, India

B.E. IN COMPUTER ENGINEERING (FIRST CLASS)

2012 - 2015

Publications

The challenges of optimizing machine translation for low resource cross-language information retrieval

EMNLP 2019

CO-AUTHOR

2018

The IIT Bombay English-Hindi parallel corpus

ICON 2015

CO-AUTHOR

2014-2015

Projects

Alfred: An LLM system to trigger APIs in stateful systems

Mumbai, India

OPENAI API

Dec 2023

- Developed a proof-of-concept for a stateful system that is able to batch-serve notifications to users based on their highly personalised natural language preferences

Legal-ease

Mumbai, India

COHERE, LANGCHAIN, QDRANT, GRADIO

March 2023

- Co-developed Legal-ease: a very early RAG system for legal document QnA
- Also includes a cross-lingual search over a multi-lingual document collection
- Legal-ease made it to finale of the multilingual semantic search hackathon organised by lablab.ai

Skills

Machine learning	numpy, scikit-learn, pytorch, pandas
REST/ORM/DBMS	FastAPI, SQLAlchemy, PostgreSQL, vector-DB (qdrant)
Front-end/Web	gradio, streamlit
Languages	python, rust