

# Mohammad Ishaan Hasan Ansari

+91-9634038353 | [mihansari.jh@gmail.com](mailto:mihansari.jh@gmail.com) | [ishaan-ansari.github.io](https://ishaan-ansari.github.io)

[in](#) ishaanansari | [gh](#) ishaan-ansari | [tw](#) iamihansari | [code\\_breaker007](#)

New Delhi, India

## SUMMARY

Machine Learning Engineer with over 2 years of experience working at the intersection of computer vision and large language models.

## EXPERIENCE

### • Think Future Technologies Pvt. Ltd

February 2024 - Present

Software Engineer L1 (Machine Learning)

Gurugram, India

#### ◦ yHealth - GenAI health coach

- \* **Product:** Designed and piloted an AI-driven health coach specifically for diabetic patients to automate routine doctor-patient interactions.
- \* **Retrieval:** Designed a multi-step retrieval pipeline with 15+ data readers and 13+ LLM calls, integrating filtering, recursion, and intent classification.
- \* **Tools:** Built 4+ components, which includes food logging, next meal suggestion, digital twin, RAG pipeline, & chatbot for AI-driven action plans.
- \* **Optimization:** Explored and benchmarked various techniques, optimized the inference pipeline and cut latency by 65%.
- \* **HIPPA Compliance:** Implemented end-to-end PII anonymization and secured third-party integrations, resulting in zero compliance incidents across external audits.

#### ◦ BSES - BRPL/YPL

- \* **Document-Parsing Search Optimization:** Redesigning search algorithms to index and retrieve document sections 3× faster, cutting average parse time from 6 s to 2 s %.
- \* **API-Call Reduction Logic:** Implementing relevance-based chunking to limit OpenAI API calls to critical text segments—reducing total monthly calls by 60% (from 10,000 to 4,000), while maintaining 95% answer accuracy.

#### ◦ CAIR - Ministry of Defence

- \* **Layout Detection Model:** Developed a layout detection model to classify and segment newspaper content into articles, images, titles, and subtitles. Fine-tuned the model to achieve an accuracy exceeding 97%.
- \* **Optical Character Recognition (OCR):** Implemented OCR to extract textual data from perso-arabic images and documents, achieving overall accuracy of 83%.
- \* **Language Detection:** Built and fine-tuned a language detection model using FastText, optimized for multiple Perso-Arabic languages.
- \* **Language Translation:** Conducted R&D on translation models like NLLB and M2M100. Developed training, fine-tuning, and evaluation pipelines for multilingual translation tasks.

#### ◦ Recruitment system

- \* **Proctoring System:** To ensure fairness during remote assessments developed a proctoring system was developed, improved accuracy by 22%, and reduced latency by 40%.
- \* **Recruitment system:** Trained Machine Learning Ranking Models for ranking developers for jobs using Gradient Boosting Decision Tree (GBDT) and Logistic Regression models using pandas, numpy, scipy, scikit-learn and lightGBM frameworks.

### • Think Future Technologies Pvt. Ltd

August 2023 - January 2024

Software Trainee (Machine Learning)

Gurugram, India

- Designed CNN architecture which reduces HEVC video compression artifacts by 6% and leveraged GoogleNet's Inception blocks to reduce the number of learnable parameters by 50%
- Experimented with GAN architecture using 3 different loss functions (perceptual loss, smooth loss, and MSE loss) and performed hyperparameter tuning.

## EDUCATION

### • Jamia Hamdard University

August 2019 - July 2023

Bachelor of Technology in Computer Science - 1st division with distinction

New Delhi, India

## SKILLS

- **Technologies:** Python, C, C++, PyTorch, Tensorflow, LLMs, Agentic & Custom frameworks, Docker, Git/GitHub, FastAPI, HuggingFace.
- **Product development:** AI Product Management, Low and Hi-Fi Prototyping, Proof of Concepts, Risk Validation & Deployment, and Iterative MVP Development.

Last updated: 2025-09-07