Job Title: AI / ML Intern (The "Builder")

Focus: OCR, LLMs, Python Scripting, Deployment.

Goal: Build the engine that processes data.

Vikhroli (west), Mumbai (Work from Office, Monday to Friday)

Duration: 6 Months (Full-Time In-Office)

Stipend: 15,000/- month

About WONDRx
WONDRx is a Mumbai-based health-tech startup transforming how healthcare data is created and connected.
Our core innovation, the **Smart Rx Kit**, digitizes handwritten prescriptions without changing how doctors write. We connect **Doctors, Clinics, Labs, Pharmacies, and Patients** into one digital ecosystem enabling healthcare access, continuity, and analytics at scale.

We're now building a strong marketing foundation to amplify our story, drive awareness, and enable growth across healthcare stakeholders.

Visit our Website: www.wondrx.com
Doctor Product Video: https://youtu.be/i4JR-lajhDQ?si=MWsegqukXi_ZBqy_
Provider Video:https://youtu.be/iQ-uFWA9fgc?si=Yw71FKhOX4Bj8HDm
WONDRx Demo Overview:https://youtu.be/s6owD0sjx6M?si=QRUP0uINdRRfxwZM

Office Location and Expansion plans

We are headquartered in Mumbai (Vikhroli-W – very conveniently located for ease of office access). We plan to expand to 50 targeted cities in next 24 months. We need passionate and energetic team players to help us achieve the aggressive target.

Position details

Job Description

Non-Negotiable Requirements (The "knock out criteria")

The "Knock out criteria" JD

    Mandatory Requirements:

1. **The "Beyond model.fit()" Rule:** Must have trained or fine-tuned a model (e.g., YOLO, Tesseract, or a HuggingFace Transformer) on a **Custom Dataset**. If they only used pre-trained models without modification, they are users, not engineers.

2. **OCR & Vision:** Must know **OpenCV** basics (Grayscaling, Thresholding) to clean images before sending them to OCR. (Critical for your doctor prescriptions).
3. **LLM "Plumbing":** Must have used **LangChain** or **OpenAI API** to build a small app (e.g., a Chatbot or Summarizer). They must understand "Prompts" and "Embeddings."
4. **Deployment Basics:** Must know how to wrap a model in a **FastAPI/Flask** wrapper so the frontend can use it.

Key Responsibilities:

- **OCR Pipeline Optimization:** You will not just "use" Tesseract. You will use **OpenCV** to preprocess images (noise removal, binarization, perspective correction) to improve the accuracy of reading handwritten doctor prescriptions.
- **LLM Integration (RAG):** Build and maintain the "Context Engine" for our Chatbots. You will convert medical documents into **Vector Embeddings** (using tools like LangChain/Pinecone) so our AI answers doctors accurately without hallucinating.
- **API Wrapper Development:** An AI model on your laptop is useless. You must wrap your models in **FastAPI** or **Flask** endpoints so our Backend Team can consume them in the real app.
- **Dataset "Grunt Work":** You will spend 20% of your time manually annotating or correcting data (labeling images or fixing OCR text) to create "Ground Truth" datasets for fine-tuning.
- **Model Benchmarking:** Regularly test new open-source models (e.g., from HuggingFace) to see if they are faster or lighter than our current models