Atharv Sabde

Email: atharv.sabde19@gmail.com

Pune, India
Phone: +91-9307287870

LinkedIn: linkedin.com/in/atharv-sabde-ind19
Portfolio: atharvsabdeportfoliov1.vercel

Professional Summary

I'm an AI Product Manager with an engineering background. I've built ML models, developed real-world AI systems, and worked through the challenges of taking them from experimentation to production. Today, I focus on deciding what to build, why it matters, and how to turn complex, real-world problems into reliable AI products that deliver measurable impact.

Experience

AI Product Manager | StampMyVisa *Oct 2026 – Present*

- Own **product management across the AI Ops and core engineering teams**, driving PRDs, prioritization, roadmap planning, and delivery metrics to keep work aligned with business goals.
- **Manage escalations from operations** and work closely with sales to surface recurring issues and new product opportunities, turning them into prioritized roadmap items.
- Lead the AI Ops Engineering team while building internal tools hands-on, including a **code review agent** and a **root cause analysis (RCA) agent** that speed up reviews and debugging.
- Driving the early steps toward AI-driven development and automation-driven operations by bringing AI agents into both engineering and operational workflows.

SDE-1 AIML | StampMyVisa *June 2026 – Oct 2026*

- Promoted to full-time SDE-1 after a successful internship, **leading AI/ML initiatives** across the organization, with a current focus on improving OCR accuracy and visa document quality evaluation pipelines.
- Partnering with engineering, operations, and customer support teams to design and deploy AI-driven workflow automation, accelerating the organization's AI-first transformation.
- **Mentoring the AI Operations team** on AI application development, troubleshooting production issues, and designing workflow automations that improve internal operational efficiency.

Founder's Office AIML Intern | StampMyVisa *Feb 2026 – June 2026*

- Built and deployed Smara, a WhatsApp-based customer support AI agent that reduced customer abandonment from **32.9% to 13.3% within the first month**, enabling the support team to scale operations with only three agents by automating routine conversations and escalating only complex cases.
- Redesigned the image quality assessment pipeline for ID-photo processing, using **computer vision techniques** that reduced false rejection of valid images from **~93% to ~4%** across a 550-image benchmark.
- Developed an **in-house document auto-cropping pipeline** using edge detection to automatically identify passport front/back pages and relevant PDF pages, producing standardized inputs for downstream OCR and quality evaluation systems.

AI Engineer Intern | The Nirmata *August 2024 – December 2024*

- **Built end-to-end document intelligence pipeline for financial automation** using AWS Textract for OCR and OpenAI LLMs for entity extraction, enabling automated financial statement analysis and insight generation from unstructured documents.
- **Engineered serverless ML deployment** on AWS with Lambda functions for model inference, S3 for document storage, API Gateway for REST endpoints, and CloudWatch for monitoring - architecting for scalability and cost efficiency.
- **Created time series forecasting system** using Facebook Prophet combined with custom feature engineering from transactional data to predict financial trends and handle seasonal patterns.
- **Built autonomous AI agent system** using CrewAI framework where specialized agents collaborate on complex accounting tasks, leveraging CrewAI's orchestration capabilities to automate multi-step financial analysis workflows

AI Research Intern | Pixel Stitch *April 2024 – July 2024*

- **Fine-tuned Stable Diffusion models** using Lora (Low-Rank Adaptation) for custom style transfer, experimenting with hyperparameter optimization (rank, alpha, learning rate) to achieve style consistency while preserving generation diversity.
- **Developed production workflows in Comfy UI** integrating ControlNet for compositional control, IP-Adapter for style consistency, and custom samplers - creating reusable pipelines for non-technical team members.
- **Managed training workflows** using **Kohya-ss** interface, configuring batch sizes, learning rate schedules, and checkpoint intervals to balance training time with model quality.

PROJECTS (GITHUB)

- **X_ Automation**: *Python, Agno framework, Groq-hosted LLMs (LLaMA 3, Gemma 2), Selenium*
  Multi-agent AI system that mimics user writing styles for automated Twitter content generation. Built agent orchestration where separate agents handle tweet scraping (Selenium), style analysis (few-shot learning with historical tweets), topic research (web search), and synthesis (maintain in g linguistic markers like sentence structure and tone). Implemented evaluation using semantic similarity metrics to validate style consistency.

- **Cognitive Stress Detection System**: *Python, OpenAI Whisper, Hugging Face transformers, Gemini API, Isolation Forest, PCA*
  End-to-end speech analysis pipeline detecting cognitive stress through linguistic deviation patterns. Engineered audio preprocessing (noise reduction, normalization), ASR transcription (Whisper), and feature extraction of 15 linguistic markers across fluency (speech rate, pauses), prosody (pitch variance), and lexical complexity. Applied unsupervised anomaly detection (Isolation Forest + PCA) to identify speech patterns deviating from baseline cognitive load.

PUBLICATIONS

Nagpure, D., **Sabde, A**., Yadav, T., Dhoble, P., Dhanduke, S. (2025). Prompt Engineering: Techniques, Failures, and Evaluations - A Comprehensive Survey. International Research Journal of Modernization in Engineering Technology and Science (IRJMETS), Vol. 7(11), November 2025

Nagpure, D., **Sabde, A.,** Yadav, T., Dhoble, P., Dhanduke, S. (2026). PromptOptix: A Defect-Aware Multi-Agent Framework for Automated Prompt Optimization in Large Language Models. International Research Journal of Modernization in Engineering Technology and Science (IRJMETS), Vol. 8(6), June 2026.

TECHNICAL SKILLS

- **Core Programming & Data:** Python, JavaScript, TypeScript, SQL, Pandas, NumPy, Matplotlib, Seaborn
- **Machine Learning & Deep Learning:** PyTorch, TensorFlow, Scikit-learn, Computer Vision, Time Series Analysis
- **LLM & Generative AI:** Langchain, LlamaIndex, Hugging Face, RAG Systems, Prompt Engineering, AI Agents, Multi-Agent Systems, AI Evaluation Frameworks.
- **Visualization & BI:** Power BI, Tableau
- **Cloud & DevOps**: AWS, Git, Docker, Model Deployment, REST APIs

EDUCATION

Savitribai Phule Pune University (SPPU) Pune, IN

• *Bachelor of Engineering – Artificial Intelligence and Machine Learning (AI & ML)* *2022 –2026*

- **Relevant Coursework:** Deep Learning, Information Retrieval, Cloud Computing, DevOps, Data Structures & Algorithms, Database Management Systems, Operating Systems, Object-Oriented Programming

Achievements and Certifications

- **Smart India Hackathon 2025 – National Finalist**: Selected as a Top 5 team in the assigned problem statement. Qualified for the national finals as part of 8,160 finalist teams (Top 0.98%) from 8,26,635 total participants across all problem statements nationwide.
- **Kaggle Expert (Discussions & Notebooks):** Discussion Rank: 463 / 358,052 | Notebook Rank: 1,303 / 294,291

  Profile: kaggle.com/atharvsabde. Recognized for strong community contributions and high-quality machine learning notebooks.
- **Machine Learning Specialization –** Stanford University / DeepLearning.AI (via Coursera)
- **Generative AI with Large Language Models –** DeepLearning.AI (via Coursera)