

PRAJYOT KORE

Senior AI Engineer | Machine Learning Researcher

Phone: +91-9579545875 | Email: pakore1999@gmail.com | LinkedIn: linkedin.com/in/prajyot-kore | GitHub: github.com/PrajyotKore

Senior AI Engineer specializing in **Generative AI, Agentic Systems, and Meta-Reinforcement Learning**. Proven track record in building production-scale ML systems with expertise in **LLM Agents, RLHF, and Financial Engineering**. Strong foundation in **Statistical Learning and Deep Learning** from IIT Bombay.

EDUCATION

Degree	Specialization	Institute	Year	GPA
M.Sc.	Applied Statistics & Informatics	IIT Bombay	2022–2024	7.03/10
B.Sc.	Statistics	Shivaji University	2018–2021	9.18/10

PROFESSIONAL EXPERIENCE

Assistant Manager - AI Product Team

April 2026 – Present

Anaptyss | Leading Technical Product Team of 8

- Leading the technical development of **3 end-to-end AI products** in the BFSI sector, overseeing architecture, stakeholder management, and cross-functional team execution of 8 engineers
- Architecting **ANA (Adaptive Neuro Intelligent Assistant)**, a conversational AI platform for **ERM Control Testing**, utilizing a multi-agent system (14+ agents) to automate RCSA analysis and walkthrough scheduling
- Engineered complex agentic workflows for **Test of Design (ToD)** and **Test of Effectiveness (ToE)**, integrating **ADK, tool calling, and TTS/STT** for autonomous evidence validation and audit report generation

AI Engineer – Multi-Agent Systems

Feb 2025 – March 2026

Anaptyss | CovenAce | Index-AI | *Financial Express Award Winner*

- Architected end-to-end **multi-agent platform** using **Google Agent Development Kit (ADK)** and LLMs for autonomous covenant extraction from legal documents, processing 1000+ credit agreements
- Engineered **self-correcting validation system** with parallel agent architecture, achieving **94.47% accuracy** and **0.93 recall** in entity classification while maintaining 98% consistency across document types
- Optimized inference pipeline with **batch processing and caching strategies**, reducing latency by **97% (5 hours → 10 minutes)** and enabling real-time compliance monitoring for financial institutions
- Designed **prompt optimization framework** with aspect-based feedback mechanisms, improving extraction F1-score by 12% across diverse covenant types

ML Engineer – Computer Vision & Optimization

Jun 2024 – Jan 2025

Ripik.AI | Industrial AI Solutions

- Compressed **Segment Anything Model (SAM)** for edge deployment using **ONNX quantization and structured pruning**, reducing VRAM by **43%** and inference latency by **32%** while maintaining 95% accuracy
- Deployed production **YOLOv8 fire detection system** on AWS EC2 with **real-time video processing**, achieving **85% mAP@50** through synthetic data augmentation and temporal consistency filtering
- Built **hybrid optimization engine** combining Genetic Algorithms and Linear Programming for manufacturing scheduling, increasing client operational efficiency by **18%** and net profit by **5%**
- Implemented **MLOps pipeline** with Docker and AWS for model versioning, A/B testing, and automated deployment

Research Intern – Multimodal Emotion AI

May 2023 – Jun 2023

I'mBesideYou, Japan | *Awarded Letter of Recommendation*

- Developed **multimodal late-fusion architecture** combining **BERT (text)** and **Wav2Vec2.0 (audio)** for inner emotion recognition, achieving **93% accuracy** on Japanese emotional speech dataset
- Implemented **ROI-based facial landmark extraction** pipeline for video analysis with temporal modeling, achieving **0.91 micro-F1** and reducing false positives by 23%
- Conducted ablation studies on fusion strategies (early vs. late) and published findings in internal research reports

TECHNICAL SKILLS

AI/ML & Agents: LLM Agents (14+), Tool Calling, TTS/STT, Large Context, RLHF (PPO/DPO), Meta-Learning, RAG, MCP, A2A

Frameworks & Tools: PyTorch, TensorFlow, JAX, HuggingFace, ONNX, LangChain, LangGraph, Google ADK, OpenAI Gymnasium

Engineering & Ops: GCP, Vertex AI, Kubernetes (K8s), Docker, AWS (EC2/S3), MLOps, System Design, ChromaDB, SQL, Git

Domain Expertise: Computer Vision (YOLO/SAM), Financial Engineering (VaR/Sharpe), Optimization (Genetic/Linear/Convex)

RESEARCH & KEY PROJECTS

UNIFIED: Meta-RL Trading Architecture

Independent Project

- Architected **UNIFIED**, a Meta-RL framework achieving **10-20x sample efficiency** and slashing regime adaptation latency from **6 months to 2 weeks** via **Transformer World Models** and MAML
- Engineered high-fidelity simulators using multi-task **BiLSTM-Transformer encoders** and **Domain Randomization** to model slippage and volume dynamics, bridging the **Sim-to-Reality gap**
- Optimized for **Risk-Adjusted Returns (Sharpe 1.5-2.0)** through **Composite Reward Engineering**, achieving a **50% reduction in max drawdown** via multi-objective optimization for P&L and stability
- Integrated **walk-forward validation** on NIFTY 50 data, surfacing auxiliary risk signals (VaR/DD) for **95%+ policy interpretability** and risk-aware trading execution

MAP-Boosting: Multi-Agent Prompt Optimization

Under Review

- Proposed novel framework treating prompt optimization as **pseudo-gradient descent over discrete instruction space** using ensemble of specialized agents with weighted aspect-based feedback
- Designed **boosting-inspired failure mode prioritization** mechanism (frequency/complexity-weighted) to systematically address domain-specific weaknesses in LLM reasoning
- Achieved **SOTA performance: +11.7% on AIME 2024, +10.6% on SWE-bench Verified, +8.5% on LiveCodeBench**, outperforming DSPy and OPRO baselines by significant margins
- Conducted comprehensive ablation studies on temperature sensitivity, agent architecture, and feedback mechanisms, demonstrating robustness across diverse reasoning tasks

End-to-End RLHF Pipeline for LLM Alignment

Independent Research

- Implemented complete **RLHF pipeline from scratch** using PyTorch and HuggingFace trl: Supervised Fine-Tuning (SFT), Reward Modeling, and Proximal Policy Optimization (PPO)
- Trained reward model on **Anthropic HH-RLHF dataset** using **Bradley-Terry loss**, achieving 87% agreement with human preferences and 0.82 rank correlation
- Optimized policy with **PPO + adaptive KL-divergence penalties** to prevent reward hacking and mode collapse, improving helpfulness scores by 34% while maintaining safety

RAG-Based Code Visualization System

Jan 2025

- Built **Retrieval-Augmented Generation system** with **ChromaDB** and **Gemini 2.0** to generate complex Manim animation code from natural language mathematical descriptions
- Designed **hierarchical chunking strategy** with semantic overlap to handle long contexts, enabling generation of mathematically accurate visualizations for calculus, linear algebra, and probability theory

Tumor Detection on Whole Slide Images (WSI)

IIT Bombay & Tata Memorial Centre, Mumbai

- Fine-tuned **KimiaNet** (histopathology-pretrained CNN) with DenseNet-121 backbone for multi-organ tumor detection on gigapixel WSIs, achieving 89% AUC on test set
- Engineered preprocessing pipeline with **Reinhard color normalization** and **adaptive patch extraction**, enabling robust cross-organ generalization across kidney, breast, and colon tissues

SCHOLASTIC ACHIEVEMENTS

- **All India Rank 22** in GATE Statistics 2023 among 5,000+ candidates
- **All India Rank 69** in ISI M.Stat Entrance Examination 2022
- **University Topper** in B.Sc. Statistics with **9.18/10 CGPA** at Shivaji University
- Secured **97.8 percentile** in IIT JAM Statistics 2022
- Honored with multiple internal accolades at **Anaptyss**, including the **Extra Miles** and **Exceptional Work** awards, for driving high-impact AI innovation and delivering production-ready agentic systems