Gaurav Prakash

📞 +1 (510)-980-1988 | ✉ gauravprakash2104@gmail.com | 🔗 gaurav-prakash | ⬛ Gauravp2104

EDUCATION

University of Maryland   College Park, MD
*Master's in **Data Science***   *Sep 2025 – Present*
National Institute of Technology Karnataka (NITK)   Surathkal, India
*Bachelor of Technology in **Information Technology***   *Dec 2021 – May 2025*

EXPERIENCE

Analytics Intern | *GlobalFoundries*   Apr 2025 – Aug 2025
- Developed a **RAG-based chatbot** using GPT-4o to assist new hires with policy and technical documentation, improving onboarding efficiency by 40%.
- Built a **resume parser** using Named Entity Recognition to extract attributes and compute candidate-job fit scores.
- Integrated dashboards for HR analytics, forecasting trends in workforce metrics using regression-based models.

Data and AI Intern | *IBM India*   May 2024 – Jul 2024
- Designed an **AI assistant** using RAG to summarize sales data and technical documents for Parle's distribution operations.
- Developed a **FastAPI + Twilio** dashboard enabling WhatsApp-based natural language queries.
- Performed geospatial analysis via Google Maps API to optimize delivery routes and resource allocation.

PROJECTS

Twitter Bot Detection using Heterogeneous GraphSAINT | *PyTorch, Graph Learning*   Jan 2025
- Implemented scalable GNNs for anomaly detection in user interaction networks using GraphSAINT sampling.
- Improved inductive node classification on TwiBot-22 dataset to 87.3% F1-score through weighted adjacency optimization.
- Benchmarked against GraphSAGE and GAT, demonstrating 1.8x speedup on large graph inference.
- Applied model explainability tools (**GNNExplainer**) to visualize edge-level feature importance for decision transparency.

Refining LLMs with Reinforcement Learning for Human-like Text | *FLAN-T5, PPO, LoRA* Jan 2024 – Apr 2024
- Fine-tuned large language models with PPO for improved human-like summarization of technical documents.
- Reduced AI detectability by 28% while maintaining factual accuracy; published at **IEEE CONECCT 2024**, IISc Bangalore.
- Utilized LoRA-based parameter-efficient tuning to achieve 30% faster convergence on multi-domain datasets.
- Applied reward shaping to optimize text coherence and factual consistency across RL iterations.

Driver Drowsiness Detection using CNNs | *TensorFlow, OpenCV*   Feb 2024
- Developed a **real-time safety monitoring system** to detect driver fatigue using CNN-based eye-state classification.
- Trained on 10K+ labeled frames to reach 91% accuracy on real-world webcam data.
- Deployed lightweight model on edge devices for real-time inference at 25 FPS.
- Prototyped integration into automotive alert systems for predictive safety response.

TECHNICAL SKILLS

Languages: Python, C++, SQL, JavaScript, HTML/CSS
Frameworks & Libraries: TensorFlow, PyTorch, scikit-learn, LangChain, OpenCV, FastAPI, Pandas, NumPy, Matplotlib
Core Areas: Predictive Analytics, NLP, Computer Vision, Reinforcement Learning, Graph Learning, Time Series Forecasting
Tools: Git, VS Code, Jupyter, Google Cloud, MS Excel, Power BI