Ronit Amar Bhatia

roncy.bhatia@gmail.com | +1 310-956-7419 | LinkedIn | GitHub | Personal Website

Education

Cornell University Aug 2024 – May 2025
Master of Engineering in Engineering Management
University of California, Davis Sept 2020 – June 2024
Bachelor of Science in Computer Science
Minor in Technology Management

Experience

Machine Learning Engineer Intern, QAlienAI (Remote) Oct 2025 – Feb 2026
- Developing an LLM-driven compliance system using semantic similarity and rule-based checks for regulatory review.
- Building multimodal pipelines with AWS Bedrock and AssemblyAI, improving OCR and ASR accuracy by 15%.
- Implementing pgvector semantic search with optimized embedding generation and hybrid retrieval for regulatory content.

Software Engineer Intern, Gallox Semiconductors (Remote) Nov 2024 – Jan 2025
- Developed automation solutions for power device characterization, improving testing efficiency.
- Engineered a real-time data visualization system in Python, accelerating debugging and analysis during lab evaluations.
- Developed Python APIs to interface with lab instruments, automating data acquisition and control during device testing.

Student Research Assistant, Cornell CALS Aug 2024 – Dec 2024
- Utilized machine learning techniques (k-means clustering, regression models) to analyze crop-based GHG emissions.
- Built Python data pipelines to preprocess geospatial crop emissions data, enhancing feature consistency for ML models.
- Implemented a clustering algorithm to group crops by emission intensity, enabling more targeted mitigation strategies.

Software Developer Intern, ColentAI (Remote) Jan 2024 – Mar 2024
- Enhanced generative AI model performance through targeted fine-tuning and hyperparameter optimization.
- Integrated external APIs to streamline data acquisition and improve training efficiency for AI models.
- Developed an automated skill taxonomy generator leveraging NLP and keyword extraction techniques.

Projects

UnClogAI: AI-Powered Workflow Bottleneck Detector [GitHub] Aug 2025 – Sept 2025
- Developed an AI workflow automation tool using graph analytics to detect bottlenecks and improve process efficiency.
- Designed a multi-agent system with LangGraph and Hugging Face LLMs for predictive insights and risk assessment.
- Optimized data pipelines and graph computations to enable efficient, real-time workflow analysis at scale.

Startup Planner Agent: AI-Powered Strategic Planning Tool [GitHub] June 2025 – June 2025
- Designed a modular multi-agent architecture using Python to orchestrate local LLM tasks with LLaMA 3 via Ollama.
- Built agent-based workflows to automate startup analysis, including market research, SWOT, and MVP planning.
- Delivered an interactive web app with exportable outputs, enabling offline strategic planning and stakeholder readiness.

Excellensight: AI-Powered Feedback Analyzer, Cornell University [GitHub] Mar 2025 – May 2025
- Built a custom CNN-BiLSTM-Attention model to classify ChatGPT user reviews and generate insight summaries without using large language models.
- Built a full NLP pipeline covering preprocessing, training, evaluation, and automated Markdown/HTML reporting.
- Achieved over 98% validation accuracy on 10K+ reviews; reports include key trend summaries and visualizations.

Taskify: AI-Powered Task-to-Team Member Matching [GitHub] Mar 2025 – May 2025
- Built a Transformer model fusing embeddings and structured data to predict task-member fit with 89.6% accuracy.
- Generated and labeled a synthetic dataset (40K+ samples) using sentence-transformers and feature engineering.
- Deployed a real-time Streamlit app to recommend top-ranked teammates with confidence scores and reasoning logic.

Technical Skills

Programming Languages: Python, C/C++, SQL, Go, JavaScript, HTML, CSS, Lisp, Prolog, MicroPython

Machine Learning & AI Tools: TensorFlow, PyTorch, Scikit-learn, LangChain, Ollama, LangFuse, LangGraph, Strands Agents, AWS Bedrock, AssemblyAI, Resend

Tools & DevOps: Next.js, React, Node.js, Supabase, Clerk, TanStack Query, Docker, GitHub, GitHub Actions, VSCode, MATLAB, JIRA, Power BI

Cloud Platforms: Vercel, AWS (S3, EC2, Lambda, SageMaker, RDS, DynamoDB), GCP (BigQuery, Vertex AI, Cloud Run)