# Abhinav Balla

LinkedIn: www.linkedin.com/in/abhinav-balla
Mobile: (669) 212-0057
Expected Graduation: May 2026
Email: abhinavballa28@berkeley.edu
Open to Relocation
Citizenship: U.S.

## Education

- **University of California, Berkeley** — Berkeley, CA
  *Double Major in Computer Science & Data Science, Business Analytics D.E.*
  - **Related Coursework**: Operating Systems, Machine Learning, Artificial Intelligence, Computer Architecture, Data Structures & Algorithms, Databases, Computer Security, Networking, Responsible AI Innovation & Management
  - **Programming Languages**: Python, Java, JavaScript, C++, C, SQL, R, HTML, CSS, Typescript, Golang
  - **Technologies**: MySQL, PostgreSQL, Pandas, Numpy, Tensorflow, PyTorch, Spark, CI/CD, Scikit Learn, Linux, Unix, Databricks, Tableau, REST APIs, FastAPI, React, Flask, Node.js, Terraform, Git, Docker, Figma, AWS EC2, S3, Microsoft Azure, Google Cloud Platform, Zapier, n8n, Ansible, Kubernetes, Jenkins

## Experience

- **Ciroos** — Pleasanton, CA
  *Software Engineer Intern* — *Aug. 2025 – Present*
  - **Fleet Management**: Developed an internal reinforcement learning (RL) playground used to train and evaluate the Ciroos SRE agent, simulating a fleet management platform for electric vehicles.
  - **Full-Stack System on Google Cloud**: Built a full-stack system with a React frontend and a Python backend deployed on Google Cloud Platform (GCP), integrating cloud services to support scalable agent simulation.
  - **Multi-Agent Simulation**: Designed a system capable of managing and orchestrating 30+ virtual car agents concurrently, enabling large-scale experimentation and policy evaluation for autonomous agent behavior.
- **Wells Fargo** — San Francisco, CA
  *Software Engineer Intern* — *Jun. 2025 - Aug. 2025*
  - **AKS Deployment Automation**: Created an Ansible playbook to automate microservice deployment to Azure Kubernetes Service (AKS) using Ansible Automation Platform, replacing Harness CD.
  - **Scalability Optimization**: Enabled deployment scalability to increase by 2x, allowing significantly more jobs to run concurrently on the infrastructure.
  - **Kubernetes Service Validator**: Developed a Python-based validator to ensure Kubernetes service YAML files created by the Cloud Operations team included all required parameters, improving deployment reliability and consistency.
- **Alkira** — San Jose, CA
  *Software Engineer Intern on Infrastructure/Platform Team* — *Jun. 2024 - Aug. 2024*
  - **Full-Stack Application**: Developed a full-stack application for margin analysis, used for analyzing opportunity cost, revenue, and total utility of money spent on cloud accounts, using Flask, Python and PostgreSQL.
  - **Cloud Monitoring Dashboard**: Created a dashboard for network vulnerability scans across cloud providers like AWS and Azure using Python and PostgreSQL.
  - **Cost Savings**: Projected to save 40% in cloud provider costs through automation of searching and deleting unused cloud accounts or leaks.

## Projects

**500 Global Hackathon Grand Prize Winner: RegWatch**: Built a compliance monitoring platform that autonomously scans regulatory sources to detect changes in healthcare/enterprise regulations and automatically generates risk assessments and remediation suggestions, including opening Git pull requests or issues when violations are detected.
**RecipeAgent**: Made a voice AI agent with a Gordon Ramsay persona using LiveKit, integrated LlamaIndex with OpenAI GPT-4 to implement RAG for answering cookbook questions from a PDF; developed a React frontend for live transcription and a FastAPI backend for JWT token authentication.
**FedEx Router Assistant**: Created a supervised model predicting delivery times and optimizing route assignments using Sci-Kit Learn, Pandas, Tensorflow, and Dijkstra's algorithm. Reduced route assignment times by 80%.
**Smart Environmental Monitoring System**: Created a device for analyzing air quality, radiation levels, light intensity, and temperature with real-time data relay to an app interface using Flask, Plotly.js, Pandas, and Numpy.
**PinPoint**: Built a web-based 20-questions-style guessing game in TypeScript where players identify cities or countries based on yes/no questions; integrated OpenAI API to generate dynamic responses and Supabase for daily-resetting leaderboards and multiplayer tracking.