

Arnav Khinvasara

Los Altos, CA 94024 | (650) 867-2862 | akhinvasara@ucsd.edu | [Linkedin](#) | [Github](#) | [Portfolio](#) | US Citizen

Education

M.S. Computer Science (Artificial Intelligence & Computer Systems)

University of California, San Diego

Sep 2024 - Dec 2025

Relevant Coursework: Algorithm Design and Analysis, Statistical Natural Language Processing, Web Mining and Recommender Systems, Data Systems for ML, Principles of ML: Learning Algorithms, Fair and Transparent ML, Computer Security, Networked Services, Music for ML

B.S. Computer Science

University of California, Santa Cruz

Sep 2020 - June 2024

Relevant Coursework: Data Structures & Algorithms, Database Management Systems, Deep Learning

Skills & Familiarities

Programming Languages: Python, C++, Java, SQL, Go

ML/AI/Data: TensorFlow, PyTorch, NumPy, Pandas, scikit-learn, Hugging Face, OpenCV, NLTK, spaCy, Apache Kafka, Apache Spark, Jupyter, MLflow, AI Agents, LLMs, NLP, Deep Learning

Cloud & Infrastructure: AWS (*Certified Cloud Practitioner*), Docker, Terraform, GitHub Actions, Jenkins CI/CD

Development Tools: Git, REST APIs, Gradle, Postman, Cursor, Windsurf

Projects

Classifying Fitness/Nutrition User Reviews Sentiment

- Developed production-ready NLP pipeline using TextCNN and DistilBERT models for real-time sentiment analysis of fitness app reviews
- Implemented advanced preprocessing with GloVe embeddings and custom tokenization, achieving 94.2% accuracy on 50K+ review dataset

Steam Game Recommender System

- Engineered a personalized recommendation engine processing 7.79M user ratings using collaborative filtering, matrix factorization (SVD), and cosine similarity algorithms
 - Achieved 89% hit rate and 0.92 precision@10 through hybrid recommendation approach combining content-based and collaborative filtering
-

Work Experience

Founder & AI Engineer | [clipshot.ai](#)

San Diego, CA | June 2025 - Present

- Developing the first domain-specific sports LLM to deliver insights for professional sports teams by integrating an end-to-end decision engine leveraging private team databases/analytics (e.g., Catapult, Second Spectrum, Synergy), to power coaching strategy, fatigue prediction, scouting, automated film analysis, as well as other applications.
- Initiated and led strategic discussions with coaches, trainers, and data analysts from the NBA Corporate Office, Los Angeles Clippers & Golden State Warriors; actively collaborating with the NBA for official data licensing.

Software Engineer Intern | [Glooko](#)

Palo Alto, CA | June 2024 - August 2024

- Utilized Terraform modularization practices to develop an ACM certificate expiration monitoring system with email notifications via SNS, Lambda functions, and EventBridge rules, and also to develop an automated ECR image management system by implementing lifecycle policies to delete unused Docker images, which optimized AWS cost efficiency
- Automated BI platform, Metabase, with Terraform, by building a Proof of Concept to display functionality of modifying dashboards/collections to enable version control and deployment across environments

Software Engineer Intern | [Malwarebytes](#)

Santa Clara, CA | July 2023 - September 2023

- Helped deploy a complex AWS migration & Jenkins upgrade for multiple accounts, managing Jenkins AMIs and scripting for deployment; ensured secure AMI access for Malwarebytes Build & Release Eng. Team
- Optimized Jenkins for performance and cost, enhancing security with custom security groups, and developed Terraform scripts for secure, scalable Jenkins setups in staging, with encrypted load balancing and automated DNS, which accelerated deployment by 20%
- Automated staging/production publishing with Jenkins pipelines; reduced duplication to improve efficiency

Software Engineer Intern | [Malwarebytes](#)

Santa Clara, CA | June 2022 - September 2022

- Engineered the Telemetry Ingestion Web Service using Java Spring Framework, boosting data ingestion efficiency and reliability by 30%, leading to a reduction in latency
- Streamlined data streaming, ETL pipeline, and Apache Kafka streams integration, fueling improved customer data analytics of the approximately 3.5 million paying consumer subscribers