Miguel Contreras

(831) 225-9841 | miguel.contreras@berkeley.edu | LinkedIn

EDUCATION

University of California, Berkeley **Expected Graduation: May 2028**

B.A. in Computer Science and B.A. in Data Science

Courses: Optimization Models in Engineering; Computer Architecture; Discrete Mathematics and Probability; Linear Algebra; Data Structures and Algorithms; Principles and Techniques of Data Science; Structure and Interpretation of Computer Programs

TECHNICAL SKILLS

Languages: Python (Pandas, NumPy, Matplotlib, Plotly), Java, JavaScript, TypeScript, C, C++, Scheme, SQL

Tools and Frameworks: React.js, Git, LaTex, Cloud Computing, CSS, Data visualization, Databases, Exploratory Data Analysis (EDA), Machine Learning, MCP, Natural Language Processing (NLP), Object-Oriented Programming

AWS: Glue, Athena, IAM Roles, Bedrock, S3, Lambda, API Gateway

RELEVANT EXPERIENCE

CourseWise (Equivalence Systems, LLC) **Aug 2025 - Present**

Machine Learning Engineer *Berkeley, CA*

- Engineered a pipeline and ingestion workflow transforming unstructured student support emails into a searchable, structured Q&A knowledge base, enabling automated assistance at scale.
- Developed reproducible parsing and cleaning systems (de-threading, signature removal, Q/A boundary identification) ensuring auditability and privacy via PII redaction.
- Implemented a retrieval-augmented generation (RAG) chatbot MVP using the cleaned Q&A dataset, including similarity-based retrieval and citation of source IDs for transparency and trust.
- Built a front-end chat interface that allowed stakeholders to interact with the AI system and evaluate the complete end-to-end product prior to platform integration.

Amazon **May 2025 - Aug. 2025**

Software Engineering Intern *Redmond, WA*

- Architected and delivered a full-stack data retention query management system that transformed manual 48–72 hour investigative workflows into automated processes completing in 5–10 minutes, serving global law enforcement agencies.
- Deployed and fine-tuned LLM-powered applications on AWS Bedrock + Claude 3.5 Sonnet to translate natural language into optimized SQL queries, demonstrating hands-on experience with LLM deployment and retrieval-augmented generation (RAG) for search and data access.
- Integrated 4+ AWS services (Athena, S3, Bedrock, GLUE) into a secure, production-ready architecture with real-time query monitoring, automated file handling, cross-account authentication, and compliance with international data protection regulations

BayFC (Professional Soccer Team) **March 2025 - May. 2025**

Machine Learning Analyst *San Jose, CA*

- Designed and deployed personalization and user modeling algorithms by applying KNN, PCA, and KMeans clustering to classify BayFC players into tactical archetypes, analogous to recommendation system personalization and behavioral modeling in large-scale platforms.
- Applied dimensionality reduction (PCA) and KMeans clustering to identify and visualize player roles, with a focus on midfielder archetypes
- Engineered player-level features (off-ball movement, sprint patterns, line-breaking actions) to predict user engagement-like metrics, aligning with CTR/CVR-style prediction and ranking tasks
- Cleaned and engineered player-level features from raw tracking data to support clustering and performance analysis

PROJECTS

Scheme Interpreter **Nov. 2024**

- Implemented a Scheme interpreter with recursive parsing and syntactic analysis, developing deep understanding of language models and tokenization processes foundational to LLM architecture and search query understanding.
- Applied recursion, lexical scoping, and tail call optimization to support key interpreter features while deepening understanding of programming language fundamentals.
- Generated a random map based on a 2D array with multiple pages for the user to traverse on their west

Auto-Corrected Typing Software **Dec. 2024**

- Programmed Python software to measure a user's typing speed and an autocorrect feature correcting user typos
- Engineered scoring feature tracking players' quickest typing speed per presented word in a multiplayer game setting

LEADERSHIP EXPERIENCE

Berkeley Anova **Jan. 2025 - Present**

Project Development Officer

- Developed and taught individualized curricula in computer science to middle and high school students, enhancing equitable access to technology education .Focused on cultivating an inclusive environment within STEM fields for underrepresented minorities.