Joseph Le

Palo Alto, CA – josephle@stanford.edu – github.com/JosephLeKH – josephle.dev
Incoming Software Engineering Intern, Apple (Summer 2026)

EDUCATION

Stanford University Palo Alto, CA
Bachelor of Science in Computer Science (Systems and AI Track), GPA: 3.95/4.0 *June 2028*

- Relevant Coursework: Data Structures & Algorithms, Operating Systems, Computer Systems, Artificial Intelligence, Machine Learning, Databases, Natural Language Processing, Linear Algebra
- Leadership & Activities: VP, CS + Social Good; Teaching Assistant, CS 106B; Stanford AI Club

WORK EXPERIENCE

Software Engineering Intern Palo Alto, CA
Vimes *Feb 2026 - Present*

- Architected and deployed a registry database with optimized indexing for fast search and built native law enforcement system integrations enabling officers to access, search, and auto-fill forms with live records.
- Engineered an AI-powered form completion system with context-aware LLM optimization and prompt engineering, reducing token usage while improving suggestion accuracy and maintaining strict data security requirements.
- Built a modular notification orchestration system from scratch enabling administrators to configure dynamic alert routing across channels and recipient types with role-based access control over information flow.

Machine Learning Research Intern Palo Alto, CA
Neurotrack *Oct 2025 – Present*

- Developed and evaluated machine learning models to detect and predict Mild Cognitive Impairment (MCI) and dementia from large-scale cognitive assessment datasets.
- Designed scalable data preprocessing, feature engineering, and validation pipelines to support experimentation and clinical research.

Software Engineer Palo Alto, CA
Stanford Carta Lab *Dec 2025 – Present*

- Built core planning features for Stanford's four-year course planner, implementing interactive scheduling workflows, prerequisite and unit constraint validation, and robust client–server state synchronization across complex multi-quarter plans.
- Identified and resolved 20+ frontend state and edge-case bugs, improving UI consistency and reducing student-reported planning errors in a high-traffic production environment.

Founding Engineer Palo Alto, CA
Bases Insight *Aug 2025 - Mar 2026*

- Architected and built a production financial analysis platform powering automated valuation workflows (DCF, comps, three-statement modeling) with an asynchronous FastAPI backend, React frontend, and PostgreSQL database.
- Designed distributed async processing pipelines for document extraction and validation, improving financial model generation throughput and system scalability.
- Implemented CI/CD pipelines with GitHub Actions and automated testing, reducing friction and increasing release reliability.

Data Science Intern Austin, TX
Bonterra *Jun 2025 - Aug 2025*

- Developed predictive models using scikit-learn, XGBoost, and CausalML to estimate treatment effects across 2.7M+ records.
- Designed and optimized large ETL pipelines with SQL, Pandas, and NumPy for feature engineering, model training, and evaluation.

TECHNICAL PROJECTS

GrantMate: Grant Writing Assistant — FastAPI, React, TypeScript, RAG, Gemini, NumPy

- Developed a full-stack retrieval-augmented generation (RAG) system for Project Homeless Connect, implementing vector search, semantic retrieval, and source attribution to improve generation accuracy and reduce hallucinated outputs.

SemEval-2025: Multilingual Emotion Prediction — aclanthology.org/2025.semeval-1.244/

- Fine-tuned multilingual BERT and XLM-RoBERTa models for emotion intensity regression across 11 languages, implementing custom training pipelines, cross-lingual evaluation, and metric optimization under benchmark constraints.

Wildfire Insurance Analysis — Python, Pandas, SQL, Regression

- Built end-to-end data pipelines and regression models to analyze California wildfire severity (2016–2023), applying time-series analysis and statistical modeling to quantify insurance non-renewal trends and FAIR Plan dependence.

TECHNICAL SKILLS

Languages: Python, Java, C++, C, JavaScript, TypeScript, SQL, Swift, Go, Bash
Frameworks & APIs: React, React Native, Node.js, Express, Spring, FastAPI, Flask, REST APIs, GraphQL
Databases & Storage: PostgreSQL, MySQL, MongoDB, Redis, Elasticsearch, Database Design, Query Optimization
Systems & Cloud: AWS (EC2, S3, Lambda, IAM), Docker, Kubernetes, Microservices, Distributed Systems, Scalability, Load Balancing, Caching, CI/CD, GitHub Actions, Unix/Linux, Networking, Containerization
Machine Learning: PyTorch, TensorFlow, scikit-learn, XGBoost, HuggingFace Transformers, Pandas, NumPy, NLP, LLM Integration