

# SALEH YAHYA

(714) 673-5706 | salehyahya10.20@gmail.com | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

## EDUCATION

### University of Illinois Urbana-Champaign

Master of Science, Computer Science

Dec 2027

- Relevant Courses: Parallel Programming, Distributed Systems, Scientific Visualization, Numerical Methods, Data Cleaning, etc

### Lewis University

Bachelor of Science, Computer Science

Dec 2026

Bachelor of Science, Information Technology

- Concentration: Systems Programming

## PROFESSIONAL EXPERIENCE

### Paidwork

#### Machine Learning Engineer Intern

Remote

Sep 2025 – Nov 2025

- Improved identity verification reliability by building real time liveness detection features that use blink analysis, head movement tracking, and automatic best frame selection, reducing unusable frames by more than 30 percent.
- Increased document verification accuracy by retraining object detection and optical character recognition models across multiple international identification formats, improving text extraction and layout detection by about 20 percent.
- Strengthened fraud prevention by developing anti spoofing systems that use depth estimation and texture based models, lowering false accept rates for printed images and replay attacks by 18 percent.

### Lewis University

Chicago, Illinois

#### Undergraduate Teaching Assistant

Aug 2025 – Dec 2025

- Elected from 30+ students To assist for Data Structures and Algorithms course, providing tutoring, coaching and mentorship.
- Organized 2 weekly study sessions, for collaborative problem solving and stronger understanding of core algorithms.
- Improved my student's performance, resulting in a 25% higher average grade compared to the previous semester.

### OveoAI

Irvine, California

#### Software Engineer Intern

May 2025 – Aug 2025

- Developed claim denial prediction models; Extracted and transformed data from multiple vendors using Python, PostGreSQL
- Built models that detect insurance claim denials, automatically classify the denial reason, and generate suggestions for corrected resubmission; improved accuracy by 8% by incorporating payer rules and coding compliance checks.

## PROJECTS

### Chromium (Google) - Embedded Open Source Contributor | [Github](#)

C++, Python, Git

- Not yet completed
- 
- 

### Market Data Simulator | [Github](#)

C++, gRPC, MySQL, Git

- Not yet completed
- 
- 

### Automated Financial Analytics Engine | [Github](#)

Python, Scikit-learn, Gemini 2.0 Pro, FastAPI, Redis, MySQL, Plaid, Git

- Built a full data pipeline ingesting real financial data from a Plaid API into a FastAPI backend, Python, MySQL, including concurrent batch processing and Redis streaming for low latency updates to improve ingestion throughput by 4x.
- Implemented regression and time series forecasting models alongside volatility, Sharpe ratio, beta, drawdown, VaR/ES, and Monte Carlo simulations, to automate my financial data, achieving under 8 percent prediction error on test data.
- Integrated LLM layer that converts raw transaction and portfolio data into interpretable insights, enabling natural social queries for anomaly detection, risk exposure, and allocation optimization, reducing manual financial analysis by over 60%.

## SKILLS

**Languages:** Python, C/C++, Java, JavaScript, TypeScript, Assembly, SQL

**Frameworks & Tools:** Kubernetes, Git, FastAPI, Flask, Linux, PostgreSQL, MySQL, MongoDB

**Machine Learning:** PyTorch, Tensorflow, Keras, Scikit-Learn