

Vedant Patel

555 Guava Lane, Davis, CA, 95616

530-204-8133 | vsxpatel@ucdavis.edu | [linkedin.com/in/Vedant1033](https://www.linkedin.com/in/Vedant1033) | github.com/Programmer7129 | vedantpatel.vercel.app

EDUCATION

University of California Davis

Expected Jun 2025

Bachelor of Science (B.S.) in Computer Science, Minor in Technology Management

Davis, CA

- Data Structures & Algorithms, AI/ML, Computer Architecture, Operating System, Concurrency and Threads, Object-Oriented Programming, Kernel Abstraction, Process Scheduling, File Systems, Virtual Memory, Circuits

EXPERIENCE

Software Engineering Intern

Apr 2025 – Present

Microsoft (through Drevol)

Seattle (Hybrid), CA

- Designed and implemented **privacy-preserving, distributed** log-analysis workflows using retrieval-augmented (RAG) LLMs, cutting error-triage time by **65%** and boosting resolution accuracy by **40%**
- Designed and optimized **scalable, load-balanced** AI-driven test orchestration pipelines leveraging **Kubernetes, Docker, and Azure cloud services**, increasing throughput by **50%** and accelerating release cycles by **25%**
- Engineered and refined modular automation solutions in **C#, .NET, and SQL**, leveraging **containerized, microservice-friendly architecture** to increase task efficiency by **20%**
- Developed and deployed **containerized AI solutions** using **image recognition and NLP (Natural Language Processing)**, reducing manual effort by **25%** while ensuring **high availability and horizontal scaling**

Machine Learning Research Assistant

Feb 2024 – Present

UC Davis

Davis, CA

Project 2: Health LLM for Personalized Health Insights | Reinforcement learning, LLMs

- Spearheaded development of a **Health LLM** that aggregates temporal, **multi-dimensional** health data across **distributed sources**—improving prediction precision by **92%** and reducing response latency by **30%**
- Engineered secure, distributed and scalable data pipelines using **Apache NiFi** to convert raw health data into **FHIR JSON schema**, increasing processing throughput by **25%** and enabling seamless integration from diverse sources with **fault-tolerant design**

Project 1: TinyML for ECG Classification & Anomaly Detection

- Engineered energy-efficient **edge-computing quantized TinyML** Random Forest ECG Classifier (92.8% accuracy) with an **event-driven, adaptive burst-mode** data collection architecture—extending **wearable battery life** from 14 days to over a month via **hybrid cloud offloading**
- Enhanced **anomaly detection to 93.6% accuracy** via advanced feature engineering and strategic hyperparameter tuning with **GridSearchCV** on resource-constrained devices, ensuring on-device inference and optimized latency
- Optimized deep learning for **time-series data** using **SHAP and LIME**, achieving a **7% accuracy boost**, architected **private-inferencing, server-side services with autoscaling**, and enabling hybrid offloading of complex multilabel classification to **server-side CNN and XGBoost models** for comprehensive health monitoring

PROJECTS

CorpCred corpcred.vercel.app | NextJS, TailwindCSS, Vercel, Render

Oct 2024 – Jan 2025

- Engineered a full-stack, microservice-friendly web app with a Random Forest **corporate credit rating** model and **RESTful APIs**, boosting accuracy by **15%** and cutting latency by **30%**

InvestIQ AI investiqai.vercel.app | NextJS, TailwindCSS, Vercel, LLMs

Jun 2024 – Dec 2024

- Developed a **containerized personal finance platform** featuring an **LLM-powered chatbot and event-driven architecture**, achieving 90% predictive accuracy, 95% query resolution, and reducing decision-making time by 30%

Dog Breed Classification ecs170.onrender.com | Python, Render, MobileNetV2

Sep 2024 – Dec 2024

LEADERSHIP

AI/ML Project Manager | CodeLab | Davis, CA

Jan 2024 – Apr 2024

- Led the Summarizit project team to develop an innovative, high-availability web application that **boosted user productivity by 40%** through efficient video content summarization while engineering scalable solutions with **React, Node.js, Express, and MongoDB**, increasing data retrieval speed by 40% and overall system performance by 25%

PUBLICATIONS

[1] **Extended Operational Life for Wearable Health Devices: A Hybrid TinyML and Server-Side ML Approach** 

[2] **Corporate Credit Rating Prediction** 

TECHNICAL SKILLS

Languages: Python, C/C++, Java, JavaScript, Typescript, C#, SQL HTML5/CSS3, R, MATLAB

Frameworks: Next.js, Postgres, PyTorch, Prisma, Vercel, .NET, React, Node.js, Flask, MongoDB, AWS, JIRA

Developer Tools: Git, Docker, CI/CD, Google Cloud Platform, VS Code, PyCharm, Eclipse, TensorFlow, Jupyter