

Sai Vishaal Yalamanchali

(510) 556-7738
vishaal.yalaman@gmail.com
vishaal.org

Machine Learning Engineer experienced in developing Generative AI solutions, with expertise in applied machine learning, feature engineering, and cloud computing. Former AWS Solutions Architect.

EDUCATION

University of California, Irvine – *B.Sc Computer Science & Bioinformatics*

SEPT 2019 - APRIL 2022, GPA: 3.5

EXPERIENCE

Medeloop AI

MACHINE LEARNING ENGINEER

APRIL 2024 - PRESENT

Led development of a grants agent using custom agentic framework and LoRA adapters.

Created custom KnowledgeGraph Agent to link predictors and outcomes, identifying novel connections between pediatric patients' lipase levels and motor tics.

Amazon Web Services, Solutions Architect, L5

AI/ML STARTUP SOLUTIONS ARCHITECT

DEC 2023 - APRIL 2024

Managing 100+ High Priority GenAI startups in Greenfield. Created Hackathon Initiative to increase adoption of GenAI services

STARTUP SOLUTIONS ARCHITECT

JUNE 2022 - DEC 2023

Managed 800+ Pre-seed to Series A Fintech and Generative AI startups, achieving 40M ARR through advisory, POCs, and relationship-building.

Enhanced SMGS Ops Team efficiency using salesparty to produce Salesforce Account Summaries, recognized at AWS All Hands in Q3 2023, reducing seller research time from 1.5 hours to 20 minutes a week.

STARTUP SOLUTIONS ARCHITECT INTERN

JUNE 2021 - AUG 2021

Developed an internal tool for Solutions Architects and Product Managers to track Product Feature Requests and Customer Influences entirely server less on AWS. Increased traction of startup Product Feature Requests by 78%.

Egnyte, Software Engineering Intern

JAN 2021 - APR 2021

Worked to improve and implement optimizations to permissions systems by modifying data structures and moving the tree structure out of the JVM to extend memory usage and increase read performance by 23%.

RELEVANT LINKS

[HuggingFace](#)

[GitHub](#)

[LinkedIn](#)

[Portfolio](#)

SKILLS

PyTorch & Torch-XLA,
Tensorflow, JAX,
AWS Neuron

Python, C++, C, NodeJS,
Java.

SageMaker, DSPy, OpenAI
Vertex AI, LangChain,
LlamaIndex.

AWARDS

Startup Solutions
Architect of H1 - 2023.

SA Builder of 2023.

SF AI Hackathon -
Hosted by MongoDB |
Anthropic, LangChain,
LlamaIndex & Together
AI

Won with Nuvo.
[link.](#)

CERTIFICATIONS

PROJECTS

Nuvo – nuvo.fyi

Created a search engine that lets you ask Anthony Bourdain for restaurant recommendations near you.

Utilized Figma to design the website, created with NextJS and OpenAI, GPT-4 to generate the recommendations.

Salesparty – [3P Dataset Here](#)

Salesparty is a suite of internal tools for AWS sellers utilizing an AWS domain expert created using custom fine-tuned Llama2-13B models trained on data from CommonCrawl and 3rd Party sources.

Piloted salesparty with 20 sellers, reducing research time from 1.5 hours to 20 minutes on avg for outbound demand and sales generation reps. Project was absorbed into another internal organization to generate Account Summaries in Salesforce for AWS sellers globally and featured in the Q3 AWS All-Hands.

Amazon Identify – [Figma Design](#)

A modular, managed KYC solution for AWS customers in a B2B2C model. Acts as a Data Custodian between users and businesses.

Utilized Figma to design user experiences, customer stories, and an MVP, securing buy-in from L8+ leadership and presenting at the AWS Financial Services Summit to 100's of stakeholders.

AWS Solutions Architect Professional

AWS Solutions Architect Associate

AWS Certified Developer Associate

LANGUAGES

English & Telugu

INTERESTS

Machine Learning, Open Source, Recommendation Systems & CUDA.