

Yosha Mundhra

San Francisco Bay Area | mundhra.yosha09@gmail.com | [LinkedIn.com/in/yoshamundhra](https://www.linkedin.com/in/yoshamundhra) | [GitHub.com/YoshaM09](https://github.com/YoshaM09)

SKILLS

Languages *Proficient* Python(5yrs) · SQL(5yrs) · Java(3yrs) *Intermediate* JavaScript(2yrs) *Beginner* Typescript(1yr) · React(1yr)
Software Azure · GCP · Langchain · TensorFlow · JupyterNotebook · Pandas · PySpark · Numpy · Apache Spark · MLFlow

EXPERIENCE

Data Scientist

Kaiser Permanente via Spotline Inc

Jan 2024 - Present

San Jose, CA

- Designing an ETL pipeline with Databricks & Spark to process identity data improving speed, accuracy of IAM analytics
- Improved response quality of an LLM chatbot by 20% with advanced prompt engineering and instruction fine-tuning
- Reduced hallucinations by 20% by implementing RAG using LangChain, OpenAI text embedding & chroma db retriever
- Analyzed cybersecurity data using SQL & Python to detect & resolve abnormalities, enhancing system integrity by 70%
- Automated and tested SaaS connector upgrades using Python and Selenium, reducing deployment time by over 85%
- Enhanced IAM CyberArk SaaS components like PVWA, PAM, and P-Cloud decreasing cybersecurity risk by 40%

Software Engineering Resident

Headstarter

Apr 2025 - Present

San Jose, CA

- Building 12+ machine learning, ai-engineering and full-stack projects in fast-paced software team environments
- Engineered an AI voice assistant with React, Groq API, LLaMA 3, and Deepgram STT/TTS, for human-like interactions
- Saved users time by designing a smart chat feature to answer natural language queries based on content of a video
- Designed MCP server for 100% automated trip planning integrating Gemini API, Exa Search, and Cursor as MCP client

Data Science Intern

CDK Global

Jun 2022 - Aug 2022

San Jose, CA

- Identified top 10 profitable products and customers, enhancing decision-making for new service pitches by 50%
- Analyzed and evaluated profit/loss margins post-migration using SQL & Python, leading to improved financial strategies
- Generated insights with (CSV, Excel) to build Power BI dashboard improving data accessibility for stakeholders by 60%

Software Engineer

Tata Consultancy Services

Dec 2018 - Aug 2021

India

- Developed 55+ functionalities with Python, Java for Humana enhancing home treatment and Rx delivery by 80%
- Delivered all user stories within 2 weeks timeline every sprint to ensure 100% agile compliance with project timelines
- Resolved 30+ defects by root cause analysis, code debugging and error-handling to reduce production bugs by 70%
- Mentored 15 new team members, enhancing code quality through code reviews and comprehensive documentation

Software Engineer

ITC Infotech

Jun 2018 - Nov 2018

India

- Integrated Rest API calls with Spring, for ITH TravelSMART, a B2B product for 20+ travel services, flight, hotels and visa
- Connected travel management application to 10+ APIs like Amadeus via GDS to ease corporate travel booking by 50%

AI/ML PROJECTS

MCP Server for Travel Planning | Open-Source, LLM, MCP, NLP (~40 hours) - [GitHub Link](#)

Jun 2025 - Jul 2025

- Developed an AI-powered MCP server to automate trip planning by integrating Gemini for natural language queries, Exa Search for travel data, and Cursor as the client, significantly improving user experience and planning efficiency.

Multimodal Video Analysis | Open-Source, LLM (~40 hours) - [GitHub Link](#)

May 2025 - Jun 2025

- Built a React and TypeScript frontend to enable natural language chat with YouTube videos using Gemini, generating timestamped answers to redirect to relevant sections, saving users time to find specific content in long videos.

Toxic Comment Classification | Open-source, NLP, Keras, TensorFlow (~30 hours) - [GitHub Link](#)

May 2023 - Jun 2023

- Designed pipeline for data pre-processing using Textthero to clean text data, form tokens and apply padding
- Built text classification model with cell state LSTM, dense layers, batch normalization layers and dropout layers to avoid overfitting. Trained 20 epochs with 35 batch size to attain a test accuracy of 0.90.

EDUCATION

College of Science, San Jose State University (SJSU)

Dec 2023

Master of Science in Data Science (Computer Science and Mathematics) | 3.5 GPA

College of Engineering, SRM University (SRMIST)

May 2018

Bachelor of Science in Information Technology | 3.8 GPA

ACTIVITIES

AUTHOR: Mundhra, Yosha, "An Incentivization Mechanism towards Network-Slice as a Service". DOI: doi.org/10.31979/etd.y8sf-pd42