Prem Dommalapati

dvnpremchowdary@gmail.com • https://www.linkedin.com/in/prem-chowdary/ •https://github.com/dvnpremoutlook

PROFESSIONAL SUMMARY

Customer-focused AI Engineer with 3+ years of Python, C++,OOD, modern NLP algorithms, transformer models, deep representations, generative models. Adept at TensorFlow, Pytorch, Linux , chromadb, langchain and AWS tooling to process and analyze large volumes of data, extract meaningful insights through the form of visualization to make data-driven business decisions. Passionate about studying and developing state of the art techniques in deep learning, machine learning, and perform in-depth analysis and optimization to ensure the best possible performance on current- and next-generation GPU architectures and seeking an opportunity as AI/ML engineer.

EDUCATION/CERTIFICATIONS

San Jose State University   **May 2026**
Master's Degree in Artificial Intelligence (AI)

San Jose State University   **May 2023**
Bachelor of Science, Computer Engineering

Relevant Coursework: Machine Learning, Programming Concepts (C, C++, Python, Java), Algorithms & Data Structures (C++), Microprocessor Design, Assembly Programming (MIPS), Digital Design I & II, Circuit Design, Electronics for Comp Systems, Engineering Stats, Database Systems, Systems Verilog, and Compiler Design

Coursera   **August 2021**
Deeplearning.AI TensorFlow Developer

WORK EXPERIENCE

Microsoft   **August 2023 – Present**
AI Support Engineer
- Worked with Azure to create multiple Azure Machine Learning base pipelines, various types of datasets
- Made multiple types of OpenAI deployments for GPT4 and GPT 3.5 models
- Worked with Contentive Search and other OpenAI and Azure Machine learning services

Uniquify   **October 2021 – August 2023**
AI Engineer
- Build and scale existing data annotation pipelines while evaluating Neural Network models across several domains to select a proper deep learning model for object detection within images.
- Improved object detection by 50% through research, development model selection, and implementing models for automatic labeling; Selected pose estimation models using UNET and Blazepose, leveraging datasets provided by the company.
- Analyzed training results, wrote neural network model, and orchestrated training runs, which used resnet50 as the base, that increased overall data accuracy by 25% for Auto labeling product
- Improved the accuracy of *deep learning* models such as yolov3, blazepose, Unet using various frameworks including pytorch, TensorFlow, and a private architecture for datasets that mainly revolve around image processing.
- Partnered with Research and Engineering teams to implement AI algorithms and neural networks, reducing initial training time by 95%
- Built and assessed performance of image processing and deep learning models written in ChainerCV and TensorFlow for standard CNN architectures, which involve the basics of fetching training data and preprocessing data to orchestrate training runs.

Partnership HealthPlan of California   **June 2017 – July 2018**
Intern
- Utilized JavaScript to enhance applications and improve performance through testing and validations against requirements and performance standards.

- Developed, designed, tested, and debugged applications from foundational levels to full production levels.
- Created web front-end user interface for new/existing databases using combination of HTML, SQL, and JavaScript to make accessible business web applications.

INTERNSHIP/PROJECT EXPERIENCE

Personal Project   **December 2021 - Present**
Web Portfolio
- Built a Portfolio using ReactJS as Frontend and Python as Backend Utilizing Python to run a neural network in the background uses ReactJS, Python, and Flask

SJSU Class Project   **January 2023 - May 2023**
Neural Network Visualization Tool
- Built a C++ GUI application that visualizes the inner workings of a neural network.
- Built this application as a prototype, to help engineers visually understand the model that they are building.
- Used C++ with OOD for the GUI aspect of the application and python for the backend of running the pretrained model of choice.

SJSU Senior Project   **February 2022 - December 2022**
LFADS with Pose Estimation
- Researched a new method in machine learning (ML) called Latent Factor Analysis via Dynamical Systems (LFADS) to create a pose estimation model.
- Built this new model to understand and improve upon the current methods of machine learning such as CNN's and Convolutional Variational Auto Encoder (CVAE) models
- Used python and COCO dataset in an attempt to produce a better model for pose estimation.

SJSU Class Project   **January 2021 – May 2022**
Pascal Compiler
- Built a compiler for a new programming language that I created called SWING - a Java-based pascal compiler that is geared toward simplifying the process creating GUIs.
- Utilized a plug-in called ANTLR to create the compiler in eclipse Used C++, JVM, Pascal, Java

Personal Project   **August 2019 – November 2020**
YouTube to MP3/MP4 Converter
- Built an application that converts a YouTube link to an MP3 file within a specified location on a computer and used various data structures to optimize for speed.
- Utilized QT Designer to create the GUI; Used C++ within the GUI and used Python for backend conversion.
- The C++ for this project was used in the GUI and the backend where the conversion happens python is used.

SKILLS

- Programming Languages: C, C++, Python, Shell Scripting
- Software & Tools: Git, Linux , OOD , AUTO Encoders,Transformers, MATLAB, TensorFlow, Machine Learning (ML), Chromadb,Pinecone, Langchain,LLM's,Terraform,AWS,Computer Vision,NLP,NodeJS
- Personal Interests: Machine Learning (ML), Linux Scripting, SOC integration techniques, Robotics