PRAVEEN Kumar Chittem

PROFILE

Research-oriented **ML/Deep Learning Engineer** with 5+ years in state-of-the-art AI techniques including **multi-agent systems, LLMs and vector databases**. Expert in end-to-end ML pipelines from problem framing to production deployment. Proven track record in **intelligent sales support systems**, large-scale data processing, and translating cutting-edge AI research into business applications. Strong foundation in both theoretical AI and practical implementation with **distributed computing** experience.

CONTACT DETAILS

@ impravin22@gmail.com
📞 +886 958 605 515
✉ Taipei, Taiwan.

CORE EXPERTISE

AI Systems: Multi-Agent Systems, LLMs, Vector Databases
ML/DL Frameworks: PyTorch, TensorFlow, Keras
Programming: Python, C++, SQL
Backend: FastAPI, PostgreSQL, Redis, APIs
Infrastructure: Docker, Kubernetes, AWS, Multi-GPU
Data Engineering: ETL Pipelines, Qdrant, Data Processing

RESEARCH PUBLICATIONS

Conference Papers: 2 publications in AI/ML conferences
Thesis Research: GPU-accelerated HPC simulation, Real-time object detection with deep learning

EXPERIENCE

SR. AI SOFTWARE DEVELOPER at *ViewSonic (Taiwan).*   **2022.03–pres.**

◇ **AI Sales Assistant System - Lead Developer**
Designed, and led the comprehensive multi-agent AI system to assist ViewSonic sales teams with product information, competitor analysis, and sales support. Designed end-to-end ML pipeline handling diverse data sources and providing intelligent responses through conversational AI interface. *Skills: Project Leadership, Multi-Agent Systems, LLMs, System Architecture*

◇ **Multi-Agent Architecture**
Developed intelligent agent system with specialized agents for product information, software features, competitor comparison, and general inquiries. Implemented agent coordination and routing to provide contextually relevant responses for sales scenarios. *Skills: Agent Design, Coordination Systems, Natural Language Processing*

◇ **Data Pipeline & Integration**
Built robust data collection pipeline from PostgreSQL, Qdrant databases, Dropbox, and SharePoint APIs. Preprocessed and organized product documentation, specifications, and sales materials into searchable knowledge base using Qdrant vector database for semantic search capabilities. *Skills: Data Engineering, Vector Databases, API Integration*

◇ **Backend Development**
Implemented Python backend with FastAPI framework supporting real-time query processing. Integrated Google Search API for real-time market intelligence and competitor information. Deployed system using openWebUI for intuitive sales team interaction. *Skills: FastAPI, Real-time Processing, Web APIs*

◇ **Graph Neural Networks for Stroke Recognition**
Developed stroke-based sketch recognition achieving 92% accuracy using GNN architectures optimized for real-time inference. *Skills: Graph Neural Networks, Real-time Processing*

◇ **Large-Scale Object Detection**
Organized 500K+ image dataset and trained object detection models achieving mAP of 0.82 using PyTorch. *Skills: Computer Vision, PyTorch*

◇ **Model Distillation Pipeline**
Fine-tuned 600M parameter model achieving 2x speed improvement while maintaining accuracy. *Skills: Model Optimization, Knowledge Distillation*

◇ **Distributed Computing Infrastructure**
Set up Kubernetes clusters for multi-GPU training and deployed distributed ML workloads. *Skills: Kubernetes, Multi-GPU Computing*

◇ **Reinforcement Learning - Pose Classification**
Developed pose classification system achieving 80% real-time accuracy using reinforcement learning on edge devices. *Skills: Reinforcement Learning, Edge Deployment*

◇ **Computer Vision with Heat Map Regression**
Implemented heat map-based detection achieving 85% accuracy using OpenVino C++. *Skills: Computer Vision, C++*

◇ **Speech-to-Sign Language GAN**
Built speech-to-sign translation system integrating Whisper ASR with generative models. *Skills: Whisper ASR, Generative Models*

AI SOFTWARE DEVELOPER at *Neurelli (Taiwan)*   **2020.08–2022.03**

◇ **Edge AI Optimization:** Deployed license plate recognition using TensorRT optimization and CUDA acceleration.
◇ **Anomaly Detection:** Implemented Variational Autoencoder for LCD panel defect detection using PyTorch.
◇ **Industrial ML Pipeline:** Built end-to-end ML solution for casting process optimization achieving 85% prediction accuracy.

EDUCATION

MASTER IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
National Taipei University of Technology, Taiwan.   **2018–2020**
◇ Thesis: *"GPU Acceleration of Flooding Simulation Using High-Performance Computing"* - Published in IEEE conference, achieved 15x speedup using CUDA optimization.

KEY ACHIEVEMENTS

◇ **Sales AI System:** Led development of intelligent sales support platform for ViewSonic
◇ **Technical Leadership:** Led cross-functional AI teams, guided peer engineers across departments
◇ **Research Impact:** 2 peer-reviewed publications in AI/ML conferences
◇ **Performance:** Achieved best results in 3 computer vision projects