Prudhvi Kumar Daruvuri

Singapore 139951 • prudhvi.daruvuri@gmail.com • 825-526-86 • linkedin.com/in/prudhvi-daruvuri
github.com/Prudhvi-Daruvuri • prudhvi-daruvuri.github.io

Innovative AI/ML Developer with a Master's degree in Artificial Intelligence and 6+ years of experience delivering state-of-the-art solutions in Computer Vision, NLP, and deep learning. Proficient in designing cutting-edge multimodal systems, leveraging LLMs, and building high-performance models to solve complex challenges. Adept at integrating emerging technologies like vector databases and embedding systems for scalable applications.

WORK EXPERIENCE

Artificial Intelligence Software Developer — **Apr 2023 - Present**
Simulation Software & Technology (S2T) Pte Ltd — **Singapore**

- **Implemented a high-performance facial recognition system** capable of processing large-scale image and video databases, significantly improving identification accuracy and scalability.
- **Designed and deployed a reverse multi-modal search application** using CLIP and ImageBind embeddings, enabling instant retrieval from over **5M images** with seamless user interaction.
- **Developed and trained token classification models** for tasks such as named entity recognition and part-of-speech tagging, enhancing NLP capabilities for client applications.
- **Built scalable AI applications leveraging open-source LLMs** and optimized vector storage with Milvus, achieving **sub-second query performance** for embeddings.
- **Performed multimedia analysis** for object detection, OCR, SpeechToText, and keyframe identification, integrating insights into a unified graph database with Neo4j for improved relational analysis.
- **Delivered speech-to-text and multilingual translation solutions**, enhancing the accessibility and usability of processed content for diverse audiences.
- **Containerized all services with Docker for seamless deployment** and orchestration in the Azure Cloud, ensuring scalability and reliability.

Machine Learning Developer (Computer Vision) — **Apr 2021 - Apr 2023**
Cynapse.ai — **Singapore**

- **Developed and optimized video analytics pipelines** using C++, enhancing processing speed and ensuring seamless integration with client applications.
- **Converted deep learning models to ONNX, UFF, and TensorRT** for efficient deployment in production pipelines, reducing model inference latency by **35%**.
- **Trained and fine-tuned custom models** for tasks such as object detection, classification, semantic segmentation, key-point detection, and action recognition.
- **Implemented semi-automatic hyperparameter tuning pipelines**, improving model performance by up to **10%**, and streamlined workflows with regularization techniques.
- **Designed and automated data processing pipelines** using Python, optimizing multi-core utilization to boost speed by **40%**.
- **Researched and developed synthetic image data generation** pipelines using GANs and Variational Autoencoders, increasing the diversity and volume of training datasets.
- **Enhanced multi-class, multi-label classification** models for facial attribute recognition (e.g., age, gender, hair type, and color) by leveraging advanced architectures and regularization techniques.
- **Explored and implemented Neural Architecture Search (NAS)** frameworks like AutoKeras and KerasTuner to design efficient deep learning models with minimal computational overhead.

Software Engineer — **Jun 2018 - Mar 2021**
Tata Consultancy Services — **Hyderabad**

- **Designed and developed ETL pipelines** using Informatica PowerCenter, automating data ingestion and transformation for Bank of America projects, ensuring **100% compliance** with data processing standards.
- **Managed loading** by designing and implementing**multi-source data files** from systems like Netezza, Oracle, and MySQL into Exadata staging areas, streamlining the data consolidation process.
- **Automated repetitive tasks** with Python scripts (e.g., Openpyxl for Excel and Regex for data validation), achieving up to **60% reduction in task completion time**.
- **Developed and tested Spark-based scripts** in Scala, leading proof-of-concept initiatives to assess Apache Spark's applicability for large-scale data processing.
- **Collaborated with onshore and offshore development leads**, conducting unit tests, maintaining detailed test logs, and ensuring seamless integration of workflows into production.
- **Improved data validation processes** by designing and implementing reusable workflows and mappings, enhancing data accuracy and consistency.

Software Developer Intern — **May 2017 - Jul 2017**
Third Wave Power — **Singapore**

- **Assisted the R&D team** in prototyping and setting up a LoRaWAN network, enabling long-range, low-power wireless communication for IoT devices.
- **Researched and evaluated open-source network servers**, identifying feasible solutions for integration with RisingHF gateways.
- **Designed and implemented a network** connecting two gateways and twenty nodes, demonstrating the scalability of LoRaWAN technology in real-world scenarios.

- **Conducted feasibility analysis** for porting network server software to RisingHF gateways, ensuring compatibility and efficiency.
- **Documented research and implementation processes**, creating detailed technical reports to support future development and deployment efforts.

EDUCATION

Master of Computing: Artificial Intelligence	**Jan 2021 - Jan 2023**
National University of Singapore	
Bachelor of Technology: Computer Science	**Apr 2014 - Apr 2018**
Gitam University	

PROJECTS

SmartAnnotator ⧉
- Built an end-to-end data annotation platform to streamline dataset preparation for various Machine Learning (ML) models, including token classification, fine-tuning Large Language Models (LLMs), and image understanding.
- Designed and implemented both the frontend and backend, fully containerized with Docker for seamless deployment and scalability.
- Integrated LLM-based auto-annotation features through external API calls (e.g., ChatGPT) as well as self-hosted solutions using Ollama.
- Developed an automated training pipeline that triggers model training after every configurable batch of new samples (default: 1,000).
- Enabled model-assisted annotation by incorporating an option to auto-annotate new data with the latest trained model.

NewsBriefer ⧉
- Developed a mobile and web-based news aggregator application that retrieves the latest articles from multiple channels via RSS feeds.
- Implemented content parsing to extract key details such as titles, images, and article text.
- Leveraged Large Language Models (LLMs) to automatically summarize news in under 100 words for quick consumption.
- Integrated text-to-speech functionality, allowing users to listen to summaries without reading.
- Introduced a "History" feature enabling users to track the evolution of news stories for better context.
- Fully containerized the application with Docker, ensuring seamless deployment and scalability.

SKILLS

Tech: Python, C++, SQL, PyTorch, Scikit-learn, Data Wrangling, Visualization, Statistical Analysis, Spark, Hadoop, Kafka, Hugging Face, OpenCV, Spacy, NLTK, AWS, Docker, CI/CD Pipelines, SQL (PostgreSQL, MySQL), NoSQL (MongoDB, Elastic), Vector Database (Milvus)

Soft Skills: Communication skills, Teamwork, project management, Presentation

Sports: Solo Paragliding, Wakeboarding, Skateboarding, Ice Skating, Trekking

AWARDS

Kaggle Competition Top 3 by **Kaggle**
Successfully developed a 'Video to Text' sequence model using CNNs and LSTM. My team ranked top 3 in Kaggle competition.

C Certificate by **NCC (National Cadet Corps)**
Completed C Certificate in NCC (National Cadet Corps).

CERTIFICATIONS

Tensorflow Developer Certificate	**Dec 2021**
Spark Developer Certificate (CCA 175)	**Jul 2020**
The Complete Regular Expressions Course by **Udemy**	**Apr 2020**
Python for Data Science Professional by **Edureka**	**Feb 2020**
Programming in C++ by **IIT Kharagpur**	**Feb 2017**

LANGUAGES

English, Upper intermediate • Hindi, Upper intermediate • Telugu, Upper intermediate