SUKESH REDDY RONDLA

E-Mail: Contact.sukeshreddyrondla@gmail.com   **Phone: +1 (970)-827-8470**   **www.SukeshReddyRondla.com**

SUMMARY

Machine Learning Engineer with 5+ years of experience building production-grade NLP and computer vision systems. Skilled in PyTorch, TensorFlow, FastAPI, and MLOps with AWS and Azure.

TECHNICAL SKILLS

Programming & Data: Python, SQL, Java, Pandas, NumPy, PySpark
ML Frameworks: PyTorch, TensorFlow, Keras, Scikit-learn, XGBoost, LightGBM
Apple ML Stack: Core ML, Create ML, Vision, Accelerate, Swift for TensorFlow
NLP & LLMs: Hugging Face Transformers, LLaMA 2, BERT, LangChain, Prompt Engineering
Deep Learning: CNNs, RNNs, Transformers, Vision Models (ResNet, EfficientNet)
MLOps & Deployment: MLflow, DVC, Docker, Kubernetes, CI/CD (GitHub Actions, Jenkins), Airflow
Model Serving: TensorFlow Serving, TorchServe, Flask, FastAPI
Monitoring & Explainability: SHAP, LIME, Prometheus, Grafana, Evidently AI
Cloud & Data Platforms: AWS (SageMaker, EC2, S3), GCP (Vertex AI, BigQuery), Azure (ML Studio), Databricks, Snowflake
Version Control & DevOps: Git, GitHub, GitLab CI/CD
Soft Skills: Agile, Cross-functional Collaboration, Communication, Problem Solving, Technical Leadership

WORK EXPERIENCE

Machine Learning Engineer, Alchemy AI - Cognitus   *March 2024 - Present*

- Led customer-facing demos and onboarding sessions for Alchemy's AI agents, translating client use cases into deployable automation workflows and iterating in real-time to address business-specific challenges.
- Collaborated directly with enterprise clients to prototype and deploy intelligent agents tailored to their operational workflows, accelerating time-to-value and driving adoption through personalized solutioning.
- Integrated AI pipelines with external systems (e.g., REST APIs, webhooks, third-party data platforms) to enable end-to-end automation, reducing manual intervention and improving throughput.
- Tuned prompt logic and fallback flows across agents to optimize performance, ensuring high accuracy and resilience in production scenarios involving variable inputs and user behavior.
- Translated customer needs into technical specs, working directly with product and customer teams to define automation rules, business logic, and exception handling in intelligent agents.
- Deployed and versioned agents in containerized environments using Docker and GitHub Actions, enabling CI/CD workflows and ensuring reproducible, scalable deployments.
- Monitored agent telemetry, accuracy, and edge cases, conducting A/B testing and leveraging observability tooling to trigger iterative improvements and model fine-tuning.
- Served as a trusted technical advisor, consulting clients and internal stakeholders on scalable architecture, prompt design, and long-term maintainability of AI-driven systems.
- Mentored junior engineers on applied AI practices, agent deployment patterns, and customer-first iteration using Agile and sprint-based delivery.

Data Scientist, Retinawise AI – RCN Networks   **June 2023 - March 2024**

- Fine-tuned multimodal AI models using PyTorch and Hugging Face Transformers for retinal disease classification, improving diagnostic suggestion alignment by 22%.
- Collaborated with ophthalmologists to define structured outputs for differential diagnoses, enhancing the clinical relevance of generated summaries.
- Developed scalable preprocessing and evaluation pipelines for grayscale retinal image data using Pandas, NumPy, and OpenCV.
- Deployed inference services using FastAPI and Docker, orchestrated on Azure Kubernetes Service (AKS) with auto-scaling for batch predictions.
- Integrated Azure Blob Storage for dataset management and leveraged Azure Key Vault to securely handle authentication tokens.
- Built interactive Streamlit tools to visualize predictions and assist internal review of model outputs by medical analysts.
- Evaluated vision-language models such as BLIP, BioViL, and CLIP for structured ophthalmic diagnosis, focusing on image-text alignment and disease pattern recognition.
- Developed prompt-based inference workflows to generate explainable diagnostic summaries from fundus scans, leveraging pretrained image encoders and clinical language models.

Machine Learning Engineer , Vamstar   **Feb 2020 - Dec 2021**

- Built and productionized NLP pipelines using spaCy, Scikit-learn, and Hugging Face Transformers to extract entities and classify healthcare procurement tenders, improving labeling efficiency by 70%.
- Developed multilingual classifiers with BERT-base-multilingual-cased, achieving 85 %+ F1-score on tender documents across English, German, and French.
- Implemented UNSPSC classification using LightGBM and rule-based feature engineering, increasing product mapping accuracy by 20%.
- Created entity resolution pipelines using RapidFuzz, TF-IDF, and MinHash LSH, deduplicating supplier/buyer records with over 90% precision.
- Engineered a semantic search and recommendation engine using SBERT and FAISS, boosting supplier match coverage by 35%.
- Deployed ML services via Docker and Azure Kubernetes Service (AKS), orchestrated with Azure ML pipelines for scalable inference.
- Built a harmonized procurement data lake using Azure Data Lake Gen2, improving model reliability and team-wide data access.
- Integrated inference APIs into the product using FastAPI and Celery, enabling real-time supplier recommendations.
- Partnered with product and data teams to define standardized taxonomies across 20+ countries, ensuring consistent downstream analytics.

Research and Development Intern, IIIT Hyderabad   **May 2018 - Nov 2018**

- Cleaned and preprocessed raw datasets using Pandas and NumPy, handling missing values, encoding categorical variables, and scaling numerical features to prepare data for ML model training.
- Performed exploratory data analysis (EDA) using Matplotlib and Seaborn to identify outliers, visualize feature distributions, and guide data-driven feature selection.

CERTIFICATION & EVENTS

- **Presented Alchemy AI at SAP Sapphire**
  Showcased intelligent agent automation using Alchemy AI for enterprise workflow transformation.
- **Apple VisionOS Developer Meet**
  Attended to explore augmented reality and multimodal application development.
- **CES 2025 Presentation: PathVisVR**
  Showcased an innovative application in digital pathology and medical imaging with Alafia AI.
- **Google Developer Groups**
  Designed AI-powered tools for smart vertical communities, including an AI Concierge and ESG-focused facility management solutions using Google Cloud AI and Aparavi.
- [AWS Certified Solutions Architect - Associate](#)
- [SQL for Data Science](#)

EDUCATION

- **Master's Degree in Computer Science**   **– Western Illinois University (2022 - 2023)**
- **Bachelor's Degree in Computer Science and Engineering**   **– Ace Engineering College(2016 - 2020)**