

Abhijith Nair

📍Hyderabad, Telangana ✉️ Abhijith799@gmail.com ☎️9676071925 🌐in/abhijithnair1

SUMMARY

A python programmer and machine learning practitioner with **5+ years of experience** in building machine learning and deep learning models.

EXPERIENCE

Verisk Analytics

Hyderabad, TS

Senior Data Scientist (prior-roles Data Scientist I, II)

Jan, 2022 – Present

- **Designed**, implemented, and deployed multiple *RAG* architectures using *LangChain* with multimodal *retrievers* (meetings, documents, presentations, images) for sales insights and customer feedback. Integrated *NER*, normalized filters for *Qdrant* and optimized chat memory with (<4s) streaming latency. Built **UI & backend DB** to store **enriched sales data** for fast, insight-driven responses.
- **Mentored** interns, while collaborating across business units on *GenAI* and *data science* initiatives to drive innovation and cross-functional alignment.
- Developed and **productionized** core document AI pipelines for Underwriting Assistant platform, integrating multimodal *vision-LLMs* to segment, classify, and extract structured data from complex unstructured insurance documents (SoVs, loss runs, Acord), reducing underwriter manual workload by **30-40%**.
- Developed and **deployed** multimodal AI (*text analytics*, *Llama*, *Blip2* Image models) for Lightspeed Small Commercial, enabling automated extraction of business operations and risk factors from unstructured text and image data sources, improved image analysis precision up by **15%**.
- Implemented *crawling* pipelines for sourcing data; successfully integrated with BIDB dataset and *Senzing* search UI application for rapid POC iterations; built ML models for classification; engineered **end-to-end OCR** document processing workflows; conducted comprehensive **data analysis** on large-scale datasets to drive actionable insights.
- Optimized on-premise *LLM* inference by deploying *quantized Llama* model with *TGI* container, reducing latency by **60%** while maintaining the original accuracy.

IntellectFaces, Inc

Chennai, TN

Data Scientist

Nov, 2020 – Jan, 2022

Project Name: AI-powered Organization Intelligence Platform - rytfit.ai

- Built and deployed an **end-to-end**, low-latency, and scalable **recommendation engine** that matches suitable candidates to job opportunities.
- Implemented deep **neural network**-based model to classify resume segments and use finetuned **entity extraction** pipeline to detect entities like skills, education, and experience from various resume formats (*PDF*, *DOCX*, *DOC*, *JPEG*).
- Developed data pipelines using *Kafka*, *ksqlDB*, and *Elasticsearch* to process candidate records and streaming data topics, reducing processing time by **20%**.
- Deployed an unsupervised **GPU-clustering model** to infer relevant skills and certifications for job roles, enabling AI-assisted job creation.
- Derived actionable insights from large datasets to identify feature relationships and improve recommendation engine performance.
- Deployed multiple containerized **microservice APIs** and **quantized models** onto production with low-latency.

Data Scientist Intern
IntellectFaces, Inc

May, 2020 – Nov, 2020
Chennai, TN

- Implemented resume parsing, search, and ranking workflows using custom **finetuned neural networks**, improving candidate matching accuracy.
- Developed and fine-tuned *BERT*-based word **embedding** models on a large volume of resumes, this helped with better text representation which improved the candidate-job match accuracy by **14%**.
- Automated data extraction/transformation from web sources with an improved **ETL** script, reducing processing time.

Deep Learning Intern - NLP
Enthire, Inc

Jan, 2020 – Mar, 2020
Bangalore, KA

- Developed and **implemented** a career-path prediction model for candidates using an **encoder-decoder sequential model**, based on the **research paper** [Next Career Move Prediction with Contextual Embedding](#).
- Built data pipelines for model using *TensorFlow* input pipeline, *SQL* and data wrangling.
- Analyzed candidate data from various sources to identify trends and relationships, informing model development.

EDUCATION

Bachelor of Technology in Information Technology

Vardhaman College of Engineering • Hyderabad, TS • 2019

ACCOMPLISHMENTS

- Awarded Verisk "Way to Go" Award for exceeding expectations in **Q1 and Q4 of 2023**.
- Recognized for **exceptional overall performance** with 'Verisk - Set a New Standard' ratings in **2022** and **2023**.
- Awarded IntellectFaces - **Best Employee** of the Year in **2020** for outstanding contributions and commitment.
- Recognized as a IntellectFaces - Star Performer in **Q1'21, Q3'20** for consistently exceeding performance expectations.
- Ranked in the top 5% of Stack Overflow answerers for Python and Pandas tags in 2021.

SKILLS

Programming:

- Python, SQL, Bash

Machine Learning / Deep Learning:

- Classification & Regression, Clustering (K-Means, DBSCAN, Agglomerative)
- Dimensionality Reduction (PCA, t-SNE, Autoencoders, Guided-LDA)
- Natural Language Processing (LLMs, Transformers)
- Computer Vision (Segmentation, CNNs)

Databases:

- MySQL, PostgreSQL, Elasticsearch, MongoDB

Cloud Computing:

- AWS (EC2, S3, Boto3, SageMaker, Bedrock), GCP, Google Vision API, Gemini API

Tools & Frameworks:

- Scikit-Learn, TensorFlow, PyTorch, Transformers, CuML, OpenCV, Spacy, Faiss, Pandas, Matplotlib
- PyTorch Lightning, Autogen, Langchain, Ragas Framework, TGI, Deepspeed, Senzing, Streamlit
- Docker, FastAPI, Redis, Kafka, ksqldb, Git

PERSONAL PROJECTS

- Fine-tuned multiple LLM models using **LoRA** approach on a publicly available prompt-completion dataset and deployed these using **LoRAX** for efficient multiple adapter-based deployment.
- Built a question-answering application powered by LLMs on health guidelines documents, leveraging the Microsoft **AutoGen** framework to facilitate complex interactions between agents.
- Utilized open-source instruct LLMs to create high-quality synthetic fine-tuning datasets from unstructured text, improving data availability and LLM performance.

CERTIFICATIONS

Deep Learning Specialization – Coursera, 2019

Microsoft Technology Associate Certification (Certiport – A Pearson VUE Business, 2015)