

Libin B Jose

+91 8281511850, +91 8700784947

libinjose100@gmail.com

<https://www.linkedin.com/in/libin-jose-41464ba6>

Machine Learning Engineer

Professional Attributes

- NLP Engineer with **8+ years** of experience in IT, specializing in Machine Learning and Artificial Intelligence.
- Proficient in building robust chatbots, chat agents, Python-based applications, and data-driven platforms.
- Experienced with various Machine Learning and Neural Network models, including **Universal Sentence Encoder, WMD and LSI, Logistic Regression, and SVM**.
- Applied word embeddings for **NLP tasks** such as document scoring, classification, and question answering from the FAQ corpora.
- Set up **data labeling pipelines** to create training data for ML models.
- **Information Retrieval & data extraction** from documents and audio files.
- Contributed to AWS infrastructure by setting up **AWS Step Functions, Lambda, and ECS** for cloud deployment and workflow management.
- Exceptional communication, collaboration, and team-building skills, with a quick ability to grasp and effectively utilize new technical concepts.

Skills

- Python
- Natural Language Processing, Machine Learning, Deep Learning
- Language Toolkits: Spacy, NLTK
- Tensorflow, Pytorch
- Packages: Gensim, scikit-learn
- Apache Beam
- GIT Version Control
- GIT Actions - CI/ CD
- Web Scraping
- Databases: MongoDB, MySQL, Redis, DynamoDB, PostgreSQL
- Search Platform - Apache Solr, Elasticsearch
- OS: Linux, Windows
- Docker, AWS ECS, AWS Sagemaker, MLflow, AWS Lambda, AWS Step Function
- Prompt Engineering, RAG & Agentic AI, Langchain, Langgraph, n8n

Work History

THUMB TECHNOLOGIES – Remote Work

Machine Learning Engineer

Dec 2021 – Jun 2025

- Creating a real-time meeting assistant.
- Information extraction from call transcripts, such as **NER, Intent, and Topic Identification**, using models & **prompt engineering**.
- Insight generation through **text clustering & summarization**.

-
- Set up the **MLflow** and **CI/CD pipelines**.
 - Implemented **AWS Step Functions** and **Lambda** for user flow management.
 - Configured **AWS ECS** for running and managing **Docker containers**.
 - Built **chat agents** and **RAGs**.

CLEAREYE.AI – Trivandrum, Kerala

Machine Learning Engineer

Dec 2020 – Dec 2021

- Contributed to the development of a product to streamline Trade Finance sanctioning and screening.
- **Data extraction** and **Data Cleaning** from **invoices**.
- Worked on entity extraction, including **Named Entity Recognition (NER)** and **goods description**.
- **Developed Seal/Signature detection capabilities**.
- Information validation across different documents.

LIGHT INFORMATION SYSTEMS – Pune, Maharashtra

Team Lead - AI/NLP Framework

Jun 2018 – Dec 2020

- The primary responsibility of the team was to develop an AI-NLP framework that enables enterprises to create customizable and scalable bots.
- Managed and contributed to 2 live projects and 15+ PoCs on the framework.
- Implemented search techniques using word embedding methods such as **Word2Vec, GloVe, Doc2Vec, Universal Sentence Encoder, and BERT**.
- Developed and integrated various modules into the platform, including a **Finite State Machine (FSM)** and document feature extraction.
- Set benchmarks and performance targets for team members.
- Conducted one-on-one sessions to identify areas for improvement and enhance team performance.
- Coordinated with internal departments to maximize operational efficiency across production and administrative areas.

LIGHT INFORMATION SYSTEMS – Pune, Maharashtra

Software Engineer

Feb 2017 – May 2018

- Worked on NLP tasks for projects and PoC applications for demonstration and evaluation.
 - Performed content crawling, data segregation, and data storage.
 - Handled data search and entity extraction tasks.
 - Developed and integrated APIs.
-

Projects

TIA

- Architected a **real-time AI assistant** capable of joining live meetings, recording conversations, and generating structured insights from speech data.
- Built end-to-end NLP pipelines including **transcription post-processing, abstractive summarization, sentiment analysis, keyword extraction**, and named entity recognition (**NER**).
- Integrated GPT-based models to enhance summarization quality, contextual understanding, and intelligent post-processing of meeting transcripts.
- Designed clustering algorithms to group similar conversations and power dashboard-level insights and analytics.
- Transformed unstructured meeting data into actionable business intelligence through scalable processing pipelines.
- Implemented ML lifecycle management using **MLflow** for experiment tracking, model versioning, and reproducibility.
- Deployed production-grade ML services on **AWS**.
- Collaborated with cross-functional teams to integrate real-time pipelines with dashboards and downstream analytics systems.

TRADE FINANCE AUTOMATION PLATFORM

- Data Extraction from various Trade Finance-related documents.
- Developed and implemented **seal and signature detection**.

NLPBOTS - AI & NLP Framework

- Worked as a Team Lead for framework development, contributing to architecture design and implementation planning.
- Led the development and management of go-live projects and Proof of Concepts (PoCs) within the framework.
- Developed data extraction techniques for documents, including feature extraction for **acronyms, disambiguation, and query attribute** extraction.
- Implemented the **Finite State Machine (FSM)** module for the framework.
- Gained hands-on experience with various search algorithms using word embedding methods such as **Word2Vec, GloVe, Universal Sentence Encoder, and BERT**.

HIRING ASSIST

- Developed a bot to assist HR in candidate hiring, reducing pre-screening time by validating candidate information and conducting pre-screening interviews.
 - The bot retrieved candidate data from the RMS and validated it against extracted resume data.
 - Knowledge base training for clarifying the candidate's job and company-related queries.
 - Implemented **resume ranking** based on job description features.
-

-
- Explored and utilized **TensorFlow, TFRecords and FAISS** for efficient search and retrieval.

MARKETING INSIGHTS

- Performed dynamic content crawling across various categories, including **SWOT analysis, company overviews, and geographical presence.**
- Extracted data from unstructured sources, such as company annual reports.
- Integrated APIs for data collection.
- Developed methods for data segregation and insight generation, such as SWOT analysis and strategic insights, using word embeddings like Word2Vec and GloVe.
- Implemented **SpaCy NER tagger** and **dependency parsing** to extract attributes from user queries.

EMPLOYEE ASSIST

- Built a bot to assist employees with daily tasks and HR-related queries, including leave applications, reimbursements, and company policies.
- Facilitated new employee onboarding by guiding recruits through company processes.
- Reduced HR workload by resolving common employee queries and escalating only essential ones.
- Document retrieval techniques using Centroids of Word Embeddings and Word Mover's Distance (WMD).

Education

CDAC – Knowledge Park, Bangalore
PG Diploma in Big Data Analytics

Feb 2017

COLLEGE OF ENGINEERING– Kottarakkara, Kerala

B.Tech in Electronics And Communication Engineering

May 2015

Awards & Recognition

- Recognized and awarded for contributions to the development of the NLP Bot Framework.
- Received recognition and an award for building the Finite State Machine (FSM) module in the framework.
- Acknowledged and awarded for work on Marketing Insights.

Languages

- English
 - Hindi
 - Malayalam
-