

K X Vinay Jose

ML Engineer

Updated on 10th April, 2025.

+91 8547 574 174

vinayjose1510@gmail.com

in /kxvinayjose

/vinay-jose

Education

National Institute of Technology, Karnataka

M.Tech in Design & Precision Engineering

2018 – 2020

CGPA: 8.47/10

Experience vinayjose.com

Cognizant Technology Solutions

Data Scientist

Sep '24 – Present

Bangalore

- > Developed ML apps at Mercedes-Benz RDI - Image Classification, Clustering & 3D CAD Regression
- > Automation Scripting using Javascript and Python to streamline Crash Simulation Process

Skills: Python, JavaScript, Pytorch, Fastai, 3D Deep Learning, LLMs, FastHTML

Dbiz.Ai

Software Engineer

Feb '24 – Aug '24

Remote

- > Research on Advanced RAG Systems & LLM Inference Optimization for portfolio product
- > First ML Engineer at Dbiz Research Labs

Skills: Python, Pytorch, Ollama, RAG, Langchain, LlamaIndex, CrewAI, LLMs, FastAPI, MongoDB, QdrantDB, AWS

Hyreo Technologies

Software Engineer

Apr '22 – Sep '23

Trivandrum

- > Led the team that developed SaaS Virtual Assistants - 65% queries automated, 5 MNCs onboarded
- > Implemented new chatbot architecture for faster deployments - reduced client onboarding times by 3x
- > Initiated performance monitoring process for all client chatbots and improved accuracy from 61% up to 89%

Skills: Python, NLP, Flask API, RASA, Shell, MongoDB, SQL, Git, Docker, Microsoft Azure, K8s, Jenkins CI, Jira

Tata Motors Ltd.

Senior Manager

Jan '21 - Apr '22

Pune

- > Worked in Product Development - coordinated digital design, validation & testing, prepared DFMEA reports
- > Analysis of supplier issues and customer feedback to detect defects & create dashboards

Skills: Product development, Customer centricity, Stakeholder management, VBScript, MS Excel

Research Intern

July '19 - June '20

- > Completed the project: **Accelerated Topology Optimization using ResNet-based Modified U-Net**
- > Developed an ML architecture from scratch to predict the solution of 2D Topology Optimization problem
- > Posed as an image segmentation problem, a U-Net model with skip connections like resnet was used.
- > Achieved results with accuracy (93.7 % in test) higher than previously published models

Skills: Python, Computer Vision, SQL, Pandas, NumPy, Pytorch, Scikit-learn

Intelligent Virtual Assistant with NLU for CRM - SaaS product

[[developed at Hyreo Inc.](#)]

- > Developed chatbots using RASA framework for users to raise concerns & answer queries using NLU
- > Built REST API layer using Flask for ticketing and operations - handles 100k requests per week
- > The new bot architecture enabled faster feature shipments - enabling 3x faster client onboarding

Advanced RAG system to chat with Open Source Github Repositories

- > Used hybrid retrieval (BM25 search + ColBERT reranking) to create a fastai expert for the learning community.
- > Powered the LLM using Gemma API endpoint from Google.

Atliq Technologies Sales Data Analysis & Visualization using MySQL, PowerBI

[[»demo«](#)]

- > The real sales data of [Atliq Technologies](#) was cloned from [CodeBasics](#) (made public for learning & research)
- > MySQL was used for data wrangling and analysis, PowerBI used to create a visualisation dashboard
- > Insights were drawn on zone-wise sales all over India and identify top products, sales/profit margins, etc.

Other Projects

- > Automation tool to test and report chatbot model accuracy for real data provided by users
- > Descriptive & Predictive Analyses of COVID-19 Student Learning Loss using R
- > Mutual Fund Portfolio Performance Prediction Analysis - Case Study
- > Data Visualization & Descriptive Analytics for Gender Gap Study in Employees DB using SQL and Tableau
- > Fake Job Postings Classifier using Random Forest Classifier and other models

Skills

Programming Python, Computer Vision, NLP, LLMs, JavaScript, Shell, SQL, MongoDB, LanceDB

Technologies & Tools Git, AWS, Docker, Kubernetes, Jenkins, Jira

Frameworks Pytorch, LangChain, LlamaIndex, Flask, Pandas, NumPy, Scikit-learn, RASA, ReactJS

Achievements & Certifications

- > Solve It With Code by Jeremy Howard & Jonathan Whitaker of Answer.ai in February 2024
- > Practical Deep Learning for Coders by Jeremy Howard of fast.ai in February 2024
- > Recipient of the annual excellence award, "Making a Difference", at Hyreo Technologies in December 2022
- > Qualified in 97th percentile for GATE (Mechanical Engineering) 2018
- > Ranked in the top 1% of students nationwide in the CBSE 12th board examinations in 2013