



Alex Joseph

+971-56-157-9017 | alex.joseph.2207@gmail.com
LinkedIn: AlexJoseph22 | Github: AlexJoseph22
Nationality: Indian | Passport Number: Y1139241
Visa: Visiting (valid till 18-11-2025)

PROFESSIONAL SUMMARY

Data Scientist with strong research and analytics acumen, expertise in optimizing large language models (LLMs), computer vision models, Generative AI, predictive analytics and creating dashboards using Tableau and Power BI. Passionate about developing innovative AI solutions and working on challenging problems. Has prior experience in education and after-sales operations in the automotive industry.

EXPERIENCE

Trainee (AI Data Scientist)

April 2024 – Dec 2024

Zoho Corporation Ltd c/o Pinakin Business Solutions Pvt Ltd

Kerala, India

- Researched and developed algorithms to enhance the performance of Large Language Models (LLMs) and tested state of the art methods on existing production models for their efficiency.
- Optimized LLM performance, achieving a 1.35x speed improvement of tokenization by implementing three novel techniques.
- Developed an efficient root-word extraction algorithm, enhancing tokenizer performance.
- Explored tokenization strategies to improve fertility rates in language models.
- Streamlined the workflow by creating data pipelines to ensure that the processes are uniform throughout.
- Used Langchain to leverage the power of Retrieval Augmented Generation (RAG).
- Leveraged Agile methodology to manage and finish the projects assigned through the use of sprints.

Consultant (Data Scientist)

Aug 2020 – Dec 2020

Rubix

Remote, Bengaluru, India

- Created a machine learning model that predicts if a customer would default his/her loan using the various data points of the customer with an f1-score of 96%.
- Solve data science problems for Rubix's clients based on the data that they provide and by offering suggestions and other solutions.
- Forecasted the inventory requirements for a Bengaluru-based automobile workshop using ARIMA and LSTM.
- Researched and performed feature selection of a medical project which further improved the accuracy score of the ML models by more than 5%.

Guest Lecturer (Automobile Theory)

Aug 2017 – Dec 2019

Government Polytechnic College Kottayam

Kerala, India

- Made classes more interactive through the use of visual media in classrooms and familiarized students about the state-of-the-art technology in the automotive industry.
- Increased the average student attendance from 86% to 95%
- Orchestrated in increasing the pass percentage of students by 12%, by taking regular feedback and incorporating the necessary items in their studies.

System Coordinator

Sep 2015 – Mar 2016

AVG Motors Ltd

Kerala, India

- Generated periodic performance reports of the company's after-sales department on a daily and monthly basis to act as the management information system (MIS).
- Crafted a manual for using their business intelligence (BI) tool so as to shorten the time required to acclimatize with the software for new MIS employees.

EDUCATION

Indian Institute of Technology, Madras

Diploma in Data Science

Chennai, Tamil Nadu, India

Sept 2021 – August 2023

Nehru College of Engineering and Research Centre

Bachelor of Technology in Automobile Engineering

Thrissur, Kerala, India

Sept 2011 – Mar 2015

PROJECTS

Retrieval Augmented Generation (RAG) using Phi3

Python, Numpy, pandas, ollama, chromadb, Sentence-transformer

Sept 2025 – Oct 2025

- Leveraged Phi3, the small language model (SLM) from Microsoft and created a RAG that runs on premise.

Customer Segmentation and Sales Forecast and Analysis of a Textile Firm

Python, Numpy, pandas, Matplotlib, ARIMA, Scikit-learn

Mar 2023 – Jul 2023

- Segmented customers on the basis of their spending pattern
- Predicted the overall sales (in rupees) for the next 5 months of the textile firm

E-commerce Customer Behaviour Understanding

Python, NumPy, pandas, Scikit-learn

Feb 2023 – Apr 2023

- Predicts if a customer would make the transactions based on their online behaviour and socio-economic features

SKILLS

Programming Languages: Python, MySQL, PostgreSQL, C++, HTML, CSS

Developer Tools: Git, VS Code, Visual Studio, Eclipse, Docker, AWS

Libraries: PyTorch, Tensorflow, Huggingface, Pandas, NumPy, Matplotlib, Scikit-learn, Selenium, BeautifulSoup, Langchain

Analytics Tools: Tableau, MS Power BI, MS Excel, Google Sheets

LANGUAGES

English - Proficient

Malayalam - Native

Japanese - Intermediate

Tamil - Elementary

Hindi - Elementary