

# JOSHUA MIRANDA

DATA SCIENTIST

✉ joshuamiranda2110@gmail.com 🌐 <https://joshhmiranda.github.io/portfolio/> ☎ +91 7738839044 📍 Mumbai, Maharastra, India  
in /in/joshua-miranda21/ 🐙 JoshhMiranda

## SKILLS

**LANGUAGES:** Python, SQL, Java

**PYTHON FRAMEWORKS AND PACKAGES:** Tensorflow, Keras, Langchain, MLflow, Sckit-learn, Numpy, Pandas

**CLOUD PLATFORMS:** GCP, AWS

**DATA SCIENCE, MACHINE LEARNING AND DEEP LEARNING:** Classical ML, Computer Vision, NLP, Sequence-to-Sequence models, Transformers, RAG, Data Analysis and visualization

**TOOLS:** Docker, SQL Workbench, Tableau, PowerBI, Neo4J

**SOFT SKILLS:** Critical Thinking, Problem Solving, Stakeholder Management, Communication

## PROFESSIONAL EXPERIENCE

**Quantiphi Analytics Solution Pvt. Ltd, Machine Learning Engineer, Mumbai**

Aug. 2024 - Current

- Deployed an end to end Deep Learning project on AWS using AWS EC2, ECR, Elastic BeanStalk, Docker for containerization and integrated CI/CD pipelines using GitHub Actions along with MLFlow for model experimentation, versioning and tracking as a part of a 12 week training program under the mentorship of a data scientist
- Created Retrieval-Augmented Generation (RAG) systems for conversational agents that interact with data from user uploaded PDFs and websites using the LangChain framework, OpenAI and HuggingFace embeddings and models, ChromaDB and FAISS as vector databases and Streamlit Cloud for deployment
- Covered classical Machine Learning, Deep Learning, Computer Vision, Natural Language Processing, Sequence-to-sequence models during the training program

**Quantiphi Analytics Solution Pvt. Ltd, Business Analyst, Mumbai**

Aug. 2022 - Aug. 2024

- Responsible for understanding, analyzing and solving complex problems through AI across a variety of industries including BFSI, Energy and Retail & CPG
- Collaborated with client stakeholders to ensure business and technology alignment through meetings, workshops, and workflow analysis
- Proposed solutions meeting defined specifications and needs with alignment from client and internal development teams
- Developed business solution flows and technical architectures basis client requirements and translated requirements into tangible deliverables
- Handled leads from 5 different regions globally, namely US East, Canada, EMEA, APAC, and India
- Successfully took ownership and assisted in closing engagements with a total deal value of over \$1M

**Ernst & Young LLP, Business Consulting Performance Improvement Intern, Mumbai**

Dec. 2021 - June 2022

- Extracted and transformed data from client databases to build dashboards in Power BI which tracked inventory levels of over 2,000 suppliers and 43,000 parts
- Developed a dashboard in PowerBI to identify production bottlenecks in the supply chain
- Worked on SQL query optimization to decrease execution time by over 30%
- Identified Key Performance Indicators and developed storyboards to assess the client's situation and assist in informed decision making
- Facilitated the implementation of a Graph DB solution using Neo 4j

## PROJECTS

**Chat with your PDFs using Retrieval Augmented Generation (RAG)**

Sept. 2024

Developed a RAG-based application using OpenAI models and Instructor Embeddings from Hugging Face, leveraging Faiss for storing PDF embeddings, allowing users to query multiple uploaded PDFs via natural language, deployed on Streamlit Cloud.

**Chat with any website using Retrieval Augmented Generation (RAG)**

Sept. 2024

Developed a RAG-based application using OpenAI models and embeddings, with ChromaDB for storing website embeddings, enabling users to interact with websites via natural language queries, deployed on Streamlit Cloud.

**Brain Tumor Classification from MRI Scans using CNN**

Aug. 2024

Deployed an end to end Deep Learning project on AWS using AWS EC2, ECR, Elastic BeanStalk, Docker for containerization and integrated CI/CD pipelines using GitHub Actions along with MLFlow for model experimentation, versioning and tracking

**Segmentation of Cars from images using UNET**

Aug. 2024

Trained a UNET model for image segmentation to detect and extract cars from horizontal and vertical orientations using 20 annotated images with image masks.

## EDUCATION

**St. Xavier's College**  
M.Sc Big Data Analytics  
CGPA: 9.81/10

Aug. 2020 - July 2022

**St. Xavier's College**  
B.Sc Information Technology  
CGPA: 3.33/4

June 2017 - Apr. 2020

## CERTIFICATIONS

**Google Associate Cloud Engineer**

Dec. 2023

Credential Link: <https://google.accreditable.com/d6a12469-791e-4126-8c3e-2ecf759d1d90>

**Google Cloud Digital Leader**

Dec. 2022

Credential Link: <https://www.credential.net/1047c45a-0895-4238-8e41-79b945f24d6e?key=b1c09bec95fefe3e07e4ab24acf3a48f5bb8da0ce16ea9f92c36644d47a8ed75>