

# Rihaz Raheem

Alappuzha, Kerala, India | +918547570698 | [rihasraheem@gmail.com](mailto:rihasraheem@gmail.com) | [GitHub](#) | [LinkedIn](#)

---

Results-driven data professional with 4+ years of experience in analytics, machine learning, and data science. Skilled in Python and end-to-end ML development, including preprocessing, feature engineering, model training, and evaluation. Hands-on with Sentence Transformers, FAISS, TF-IDF, YOLOv8, and TensorFlow. Strong background delivering scalable, data-driven solutions across structured and unstructured datasets.

## KEY SKILLS

---

Programming & Machine Learning: Python, Scikit-learn, TensorFlow, Flask, OpenCV, TF-IDF, Text Classification, SBERT, NLP, Semantic Search

Analytical Skills: Advanced Excel, Feature Engineering, Data Preprocessing, EDA (Exploratory Data Analysis), Model Evaluation, NumPy, Pandas

Other Skills: Basic HTML & CSS, Basic SQL, Google Colab, MS Office

Soft Skills: Collaboration, Communication, Mentoring, Problem Solving

## WORK EXPERIENCE

---

### Scipy Technologies • Remote • 05/2025 - Present

#### Data Scientist

- Built and optimized end-to-end ML and deep learning models using Python, Scikit-learn, TensorFlow, and Keras, improving model performance and training stability across multiple classification and prediction tasks.
- Developed NLP systems for text classification and semantic similarity leveraging TF-IDF and Sentence-BERT (SBERT) embeddings, improving semantic relevance over traditional keyword-based approaches.
- Designed a semantic search and document retrieval system involving text extraction, chunking strategies, embedding generation, and fast similarity search using FAISS, enabling fast and scalable similarity-based retrieval.
- Performed comprehensive EDA, feature engineering, and data preprocessing on structured and unstructured datasets to improve model performance and interpretability.
- Evaluated models using business-relevant metrics (Precision, Recall, F1-score) and developed basic Flask applications for model integration and gained utilized SQLAlchemy for managing structured datasets.

### District Industries Centre • Alappuzha • 04/2022 - 03/2025

#### Enterprise Development Executive

- Led enterprise development initiatives and conducted data-driven project evaluations that influenced funding and other decisions.
- Designed Excel dashboards and reports for performance tracking, budgeting, and strategy optimization.
- Achieved 100% target fulfillment across multiple development schemes through data-driven planning.

### SMEG Edulabs • 10/2020 - 03/2022

#### Assistant Manager

- Resolved student and parent concerns efficiently via email, phone, and in-person communication, boosting satisfaction and contributing to improved student retention and engagement rates.
- Built Excel-based reporting workflows to streamline administrative performance evaluation.

## **EduBridge India • 01/2020 - 06/2020**

### **Trainer**

- Trained 100+ students in data analysis and Excel-based numerical reasoning
- Delivered hands-on sessions using VLOOKUP, PivotTables, and conditional formulas.

## **EDUCATION**

---

### **Master's in Signal Processing**

Government Engineering College Barton Hill, Thiruvananthapuram (2017 – 2019)

### **Bachelor's in Electronics and Communication Engineering**

College of Engineering and Management Punnapra, Alappuzha (2013 – 2017)

## **PROJECTS**

---

**Customer Churn Prediction:** Built end-to-end ML models (Logistic Regression, Random Forest, XGBoost) with preprocessing pipelines, improving churn recall to 0.78 and ROC-AUC to 0.81.

**Semantic Search Engine:** Developed a vector-based search engine using Sentence Transformer embeddings (SBERT) and FAISS for fast similarity retrieval.

**News Article Classification:** Implemented news classification using TF-IDF and SBERT embeddings; compared traditional vs modern NLP approaches, achieving 98% accuracy.

**Age & Gender Detection:** Designed a custom CNN and enhanced it with MobileNetV2 transfer learning, reducing age MAE from 5.1 to 4.7 years and increasing gender prediction accuracy from 69% to 82%.

**AI-Based Person Detection:** Tracking & ID Tagging: Developed a real-time computer vision system for multi-person detection and tracking using YOLOv8 and DeepSORT.

## **CERTIFICATIONS**

---

- PG Certificate in Entrepreneurship: Small Business Management – Digital University Kerala
- Business Analytics with Excel – Simplilearn
- Data Science and real-time project training on Learn to Build a Real Time Application for Emotional Detector– Null Class
- HTML5 & CSS3 – Pirple

## **LANGUAGE**

---

- English (Excellent)
- Malayalam (Mother Tongue)