# SYED MUSHARAF

## Data Scientist — GenAI, LLMs Workflow Automations

+91 9666610052  Hyderabad, India  mushumushu656@gmail.com

LinkedIn Profile  My Portfolio

---

## OBJECTIVE

Data Scientist with practical experience and specializing in Generative AI and Machine Learning. Possess solid knowledge of data. developing intelligent data pipelines from preprocessing to model deployment.

---

## EDUCATION

**Bachelor's in Computer Applications**, Telangana University                    **2020 − 2023**
CGPA: 80%

---

## AI Data Science Toolkit

- **Programming Language:** Python, MySQL

- **Technical Skills:** Machine Learning, Deep Learning, Natural Language Processing Generative AI(LLMs)

- **Automation Agents** N8N (Workflows), MCP(Basics), Crew AI

- **Libraries:** LangChain, Scikit-learn, TensorFlow, Keras,Prompt Engineering

- **Dashboard Tools:** Power BI, Data Analysis, Statistical Modeling

- **Deployment:** Google Cloud Platform(basics),Git Hub,Pinecone

---

## EXPERIENCE

Nov 2024 – Present Associate Data Scientist,Mentored by Senior Data Scientist

### RAG Intelligent Chatbot

- **Action:** Developed and deployed an advanced **Retrieval-Augmented Generation (RAG)** chatbot powered by **OpenAI models**, integrating over **100+ PDF documents** for real-time, context-aware Q&A. Included **multimodal features** such as chat history retention, voice interaction, and image-based queries to improve accessibility.

- **Scope:** Engineered a scalable backend architecture for **data ingestion, embedding generation, and vector database storage**, enabling efficient retrieval and accurate contextual responses. Deployed the solution on **Google Cloud Platform (GCP)** and **GitHub** to ensure global accessibility, reliability, and security.

- **Impact:** Improved retrieval accuracy by **30%**, enhanced user engagement by **45%**, and reduced manual document lookup time by **70%**. Achieved a consistent **99.9% uptime**, driving measurable gains in efficiency and workflow automation.

### Procurement Agent Automation Platform for Finance Operations

- **Action:** Designed and implemented an end-to-end workflow automation platform using **n8n**, streamlining the procurement process including invoice intake via email, PDF data extraction, invoice-to-PO validation against Google Sheets, ERP posting, daily bank reconciliation, exception handling, and audit trail management.

- **Scope:** Automated the complete procurement lifecycle—invoice processing, validation, ERP integration, and reconciliation—by integrating **Gmail, Google Sheets, and ERP systems** for real-time data synchronization and operational efficiency.

- **Impact:** Reduced manual effort and validation time for invoice and payment processing by **60%**, increased data accuracy through automatic validation and exception flagging, and enhanced financial control and compliance with automated logs and real-time alerts.

### AI-powered Resume Analyzer

- Developed and deployed an AI-powered **Resume Analyzer** using **Streamlit**, **TF-IDF**, and a pre-trained ML model to automatically predict resume categories across multiple job domains.

- Built an NLP-driven data preprocessing pipeline with regex-based text cleaning and integrated multi-format resume parsing using **PyPDF2** and **python-docx**.

- Improved resume classification speed by **80%**, enabling recruiters to efficiently screen and categorize resumes in real time, reducing manual workload significantly.

## PROJECTS

### AI Job Researcher & Assistant (n8n)

- Automated job search by integrating form inputs with AI and SERP API, filtering results within 24 hours.

- Delivered structured job alerts via Gmail, including job title, company, location, and apply link.

- *Workflow:* Form Submission → AI Agent (Groq Model) → SERP API → Gmail Job Alerts.

### HR Data Analysis Dashboard (Power BI)

- Developed an interactive HR Analytics Dashboard visualizing workforce metrics.

- Analyzed employee demographics, salary distribution, and role-based KPIs for HR decision-making.

## CERTIFICATIONS

- **Advanced Data Science Certification** – Full Stack Academy, Feb 2024

- **Generative AI and LLM Applications** – Udemy (Instructor: Krish Naik), Jan 2025

- **Generative AI Professional Certification** – Oracle University, 2025