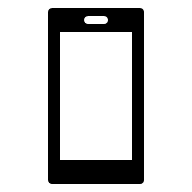
**Shubham Srinivas Karwankar**

📍 Magdeburg, Germany | 📧 [karwankarshubham@gmail.com |](mailto:karwankarshubham@gmail.com%20|) +4915563015462 | 🔗 [LinkedIn](https://www.linkedin.com/in/shubham-karwankar-335094168/)

**PROFESSIONAL SUMMARY**

Results-driven Software Engineer with hands-on experience in Python programming, machine learning, and AI application development. Skilled in building end-to-end AI systems including web scraping, embedding models, and retrieval-augmented generation (RAG) using Hugging Face Transformers and LangChain. Experienced with cloud platforms and data engineering tools such as GCP, BigQuery, Airflow, and containerization with Docker, enabling robust and scalable deployments. Proficient in developing interactive applications using FastAPI and Streamlit, with a strong focus on prompt engineering and real-world noisy data handling. Currently pursuing M.Sc. Digital Engineering with specialization in AI, ML, and DL, eager to contribute innovative AI solutions in dynamic environments.

**CORE COMPETENCIES**

**Programming Languages & ML Frameworks:** Python, SQL, C, Pandas, NumPy, PyTorch, Scikit-learn

**Cloud Platforms & Data Tools:** Google Cloud Platform (GCP), BigQuery, AtScale, SSAS, Data Catalog

**Data Engineering & Workflow Automation:** ETL development, Data Warehousing, Apache Airflow, Jenkins, Docker, Git, CI/CD pipelines

**Visualization & Reporting:** Streamlit, Tableau, Matplotlib, Seaborn, MS Excel (VLOOKUP, PivotTables)

**Databases:** SQL Server, MySQL

**AI Platforms & Libraries:** LangChain, LangGraph, LangSmith

**Languages:** English (C1), German (A2), Marathi (Native), Telugu (Native).

**PROFESSIONAL EXPERIENCE**

**Data Engineer & BI Engineer TCS (Client: The Home Depot) – Hyderabad, India | Aug 2021 – Sep 2024**

* Built and optimized high-performance ETL pipelines, increasing data availability by 40% and reducing processing times by 25%, ensuring timely data delivery for analytics teams.
* Collaborated with stakeholders to redesign and migrate over 30 complex SSAS cube metrics to a cloud architecture using AtScale and BigQuery, improving data granularity and business insight capabilities.
* Automated critical data workflows using Jenkins, Airflow, and GCP BigQuery, resulting in streamlined operations and a reduction of more than 10 TB in daily data processing through optimized SQL query refactoring.

**IoT-Embedded Systems Intern *ADOBE CREATIVE TECHNOLOGY ACADEMY*– India | Jul 2020 – Dec 2020**• Developed IoT prototypes: IR line-follower bot, smart street lighting, and home automation.  
• Integrated sensors & microcontrollers using Python, Arduino IDE, ESP8266, Raspberry Pi.  
• Built mobile apps via MIT App Inventor for real-time device control.

**EDUCATION**

M.Sc. Digital Engineering (Data Science) – Otto von Guericke University, Germany | OCT 24 – Present | CGPA: 2.3  
B.Tech. Electrical & Electronics Engineering – Guru Nanak Institutions Technical Campus, India | CGPA: 8.0

**PROJECTS**

**Taxi Fare Prediction – End-to-End ML Pipeline**

***(****BigQuery, Pandas, FEAST, Scikit-learn, MLflow, Docker, Git, SQL)*

* Designed and implemented a complete machine learning pipeline including data ingestion, cleaning, feature engineering, and feature store integration using Pandas and FEAST to ensure high-quality, reusable features.
* Built and trained regression models with Scikit-learn on large-scale real-world taxi fare data, optimizing hyperparameters for improved prediction accuracy based on trip and fare attributes.
* Deployed the trained model using MLflow and Docker for scalable, containerized serving, integrating with cloud infrastructure to support monitoring and continuous evaluation.
* Developed an interactive user interface to deliver fare predictions accessible to non-technical users, enhancing usability and stakeholder engagement through an intuitive **UI.**

**RAG-Powered Product Q&A Chatbot with Web Scraping**

*(Python, LangChain, Hugging Face Transformers, BeautifulSoup, FAISS, FastAPI, Streamlit, Git)*

* Developed a product Q&A chatbot using Retrieval-Augmented Generation (RAG) leveraging Hugging Face Transformers models to generate context-aware answers from scraped e-commerce product data.
* Built web scrapers using Python and BeautifulSoup to collect and preprocess product info from multiple online stores, chunked text for effective retrieval.
* Generated semantic embeddings with Hugging Face sentence-transformers and stored them in Pinecone/FAISS vector database for fast similarity search.
* Integrated Hugging Face language models through LangChain to perform retrieval and generate coherent natural language responses.
* Deployed the solution end-to-end using FastAPI for API serving and Streamlit for an interactive user interface supporting multi-turn conversations.
* Applied prompt engineering, fallback handling, and optimized vector search to improve response accuracy on real-world, noisy product datasets.