Shubham Pandey

Pune,Maharashtra, India •shubhampandey45.sp@gmail.com • (+91) 6392376261 • LinkedIn

SUMMARY

Data Engineer experienced in ETL pipelines, SQL, and optimizing large-scale production data systems for performance and reliability. Passionate about AI, with hands-on GenAI projects (RAG) and aspiring to transition into AI-driven solutions.

PROFESSIONAL EXPERIENCE

TATA CONSULTANCY SERVICES (TCS)   Pune, India
Python Developer | ETL & Data Engineering   Jan 2024–Present

- Developing a production support **AI assistant** using **RAG**-based knowledge base to reduce human effort, streamline issue resolution, and improve operational efficiency.
- Designed and optimized ETL pipelines supporting business-critical data workflows, reducing execution time by 50% through **Python** and **SQL** optimizations.
- Improved **PostgreSQL** performance by tuning queries and adding indexing strategies, cutting average response times by over **60**%.
- Implemented automated data validation and quality checks to detect anomalies early and ensure end-to-end data integrity.
- Resolved real-time production issues by performing **root cause analysis,** implementing long-term fixes, and improving system reliability.

EDUCATION

United College of Engineering & Research,   Prayagraj, India
Bachelor of Technology*; Major in Computer Science*   2019-2023

- Secured Global rank of **3370** and College rank of **3rd** in TCS Codevita Season 10.

PROJECT   VIEW SITE , GITHUB

Movie Recommendation

- Developed a content-based recommendation system using **machine learning** and **vectorization** techniques to deliver accurate similarity-driven suggestions
- Built a user-friendly **Streamlit** interface for seamless exploration and improved user engagement.

Tech Involved: Python, Scikit-learn, Streamlit, Pickle, Pandas, NumPy.

Parth – AI Voice Assistant   **View Site , Github**

- Developed a **real-time voice assistant** that delivers **contextual answers** from **100+** curated **Bhagavad Gita verses** in **Hindi**, understanding questions in both Hindi and English, helping users gain spiritual and practical guidance effortlessly.
- Built a **RAG-based semantic search pipeline** integrated with **Google Gemini (`gemini-1.5-flash`)** for **LLM and TTS**, enabling **smooth, natural conversations** with follow-up interactions.
- Deployed with **Node.js, Docker,** and **cloud infrastructure**, creating a **scalable, production-ready platform** accessible across devices, making spiritual knowledge easily available to anyone.

ADDITIONAL INFORMATION

- **Technical Skills**: Python, PostgreSQL, ETL Processes, AWS (S3, RDS, Redshift, Lambda, Glue, Athena), Django REST Framework (DRF), Docker, Control-M, PySpark, Generative AI, RAG-based Systems, LangChain, LangGraph
- **Tools**: Git, GitHub, Pandas, NumPy, Streamlit, FAISS (Vector Database), Apache Airflow, Jira
- **Interests:** AI-driven automation, building GenAI (RAG)-based applications, and Stock market analysis.
- **Languages**: Fluent in Hindi(native), English.