

Bilal Momin

Data Scientist | ML & LLM

bm.works00@gmail.com

+918097068954

Mumbai, India

<https://www.linkedin.com/in/bilal-momin>

<https://github.com/bilalmomin>

<https://bilalmomin.com>

About Me

LLM-focused Data Scientist with 4+ years of experience in Python, cloud platforms (GCP, AWS), and building scalable ML pipelines. I specialize in GenAI applications using LangChain, LangGraph, and fine-tuned LLMs for NLP tasks like sentiment analysis, classification, and document Q&A. I've built multi-agent systems, RAG-based tools, and end-to-end AI products. Google Cloud certified and passionate about turning real-world problems into intelligent, user-centric solutions through production-grade ML systems.

Skills

Python

Machine Learning

Large Language Models

Natural Language Processing

NoSQL/SQL

RAG

GenAI

Doccker

Git

ETL

AWS/GCP

Prompt Engineering

Fine-tuning LLM

Flask/FastAPI

Work Experience

Data Science Consultant

Freelance | Jan 2025 – Present

- * Designed and deployed a multi-agent LLM system using LangGraph for personalized resume generation, automating content adaptation and reducing manual effort by 80%.
- * Built a natural language QA assistant enabling users to query large Excel/CSV files, increasing accessibility of structured data for non-technical stakeholders.
- * Developed a RAG-based GenAI solution for an electronics company to automate product-related inquiries, improving response accuracy and reducing customer support load.

Data Scientist

Admazes Limited | Dec 2021 – Oct 2024

- * Led ML development for a trend prediction system using 100M+ records/month
- * Built ETL pipelines and trained clustering/regression models
- * Developed a TensorFlow-based topic classification engine
- * Engineered web crawlers (Quora, Reddit, Twitter) collecting 10M+ records
- * Managed GCP-hosted ML APIs for real-time demographic insights

Python Developer

Codemarket | Dec 2020 – Feb 2021

- * Built Python/ReactJS services with GraphQL APIs over MongoDB, reducing query latency by 40%
- * Automated backend workflows using AWS Lambda, Fargate, and ECR, improving processing speed by 25%
- * Managed 10+ projects with Git-based CI/CD pipelines, ensuring on-time delivery across cross-functional teams

Projects

LangChain-Based Document Chat API

- * Built a LangChain-powered Document Chat API using agent-based execution to handle natural language queries across PDF, CSV, and XLSX files with 92%+ response accuracy and <200ms latency.
- * Designed metadata-aware chunking and multi-format parsers, reducing processing time by 35% through cloud-optimized logic.
- * Architected the system to handle 1M+ daily queries using parallel task execution and asynchronous document agents.

Query Trend Forecasting Platform

- * Designed an ML-based keyword trend forecasting platform using 100M+ monthly Google SERP records, achieving 95% model accuracy.
- * Applied clustering and regression models to detect demographic patterns and keyword growth signals.
- * Scaled model training and inference workflows over 10TB+ data in BigQuery; served real-time insights via low-latency APIs on GCP.

Automated Resume Creation using LLM

- * Designed and deployed a multi-agent resume generation service using LangGraph, coordinating task-specific LLM agents for parsing user profiles, analyzing job descriptions, and enriching content with domain-relevant keywords.
- * Leveraged contextual reasoning and feedback loops to optimize content structure, tone, and ATS compatibility.
- * The system reduced manual resume creation time by 80% and significantly improved the relevance and personalization of generated resumes for diverse roles and industries.

Education

Bachelor of Engineering

Mumbai University

GPA: 7.4 | Jul 2016 – Sept 2020

Certification

Google Cloud Professional Data Engineer

Issued by Google Cloud Platform | Jul 2024