

# Dipesh Maindolia

Machine Learning Engineer

Ramnagar, Uttarakhand, India | +91 7017982390 | [dipeshmaindolia000@gmail.com](mailto:dipeshmaindolia000@gmail.com) | [LinkedIn](#) | [GitHub](#)

## Professional Summary

Machine Learning Engineer with hands-on experience building end-to-end ML and GenAI pipelines – from data ingestion and feature engineering to model deployment. Skilled in Python, XGBoost/scikit-learn, LLM-based application development, and cloud data pipelines on AWS, with working knowledge of Docker and MCP servers for building and deploying model-serving applications. GATE CSE and GATE DA qualified.

## Technical Skills

- **Programming & Query Languages:** Python, SQL, C++
- **Machine Learning:** scikit-learn, XGBoost, Feature Engineering, Hyperparameter Tuning (Bayesian Optimization), Model Evaluation
- **Generative AI:** LLM Application Development, Prompt Engineering, Few-Shot Learning, LangChain, Groq API (Llama 3.2), MCP Servers
- **MLOps & Deployment:** Docker, Streamlit, AWS (S3, Lambda, Glue, Step Functions, SageMaker), Model Serving, CI-Ready Pipelines
- **Data Engineering:** ETL / Data Pipelines, Medallion Architecture, Data Cleaning, Data Validation, Pipeline Monitoring
- **Core Competencies:** Statistical Analysis, Data Visualization, Problem Solving, Documentation

## Experience

### Product Analyst Intern

*Interviewkit.ai (Remote)*

*April 2026 – June 2026*

- Collaborated directly with the founder on product strategy, feature planning, and execution, translating usage data and user feedback into prioritized improvements to platform workflows and engagement systems.
- Built a REST API editor and an intelligent job recommendation module using keyword extraction, structured filtering, and profile-matching techniques to surface the most relevant listings for each user.
- Developed a resume-based system to auto-generate personalized technical interview questions and detailed, role-specific model answers for candidates preparing across multiple domains.

### Analyst Fellow

*CrackNonTech*

*May 2026 – June 2026*

- Studied profitability analysis, market entry frameworks, and cost-benefit evaluation through case studies and mentor-led sessions on real-world, ambiguous business problems and scenarios.
- Sharpened SQL querying, Python scripting, and guesstimate skills to model business problems, estimate market sizes, and present structured, data-driven strategic recommendations to peers and mentors.

## Projects

### Flight Price Prediction using AWS SageMaker – Python, XGBoost, scikit-learn, Streamlit

[GitHub](#)

- **End-to-End ML Pipeline:** Built a complete flight price prediction system on AWS SageMaker, covering preprocessing, feature engineering, and hyperparameter tuning using XGBoost with Bayesian optimization.
- **Feature Engineering:** Engineered features from raw flight data (airline, route, timestamps, duration) using scikit-learn Pipelines, FeatureUnion, and custom transformers.
- **Deployment:** Deployed a Streamlit web application to Streamlit Cloud, enabling users to input flight details and receive real-time price predictions.

### LinkedIn Post Generator – Gen AI (Llama 3.2) – Python, LangChain, Groq API, Streamlit

[GitHub](#)

- **GenAI Pipeline:** Built a two-stage Generative AI pipeline that extracts topic, language, and length metadata from a user's past LinkedIn posts, then uses Llama 3.2 (via the Groq API) to generate new posts that match their personal writing style.
- **Prompt & Retrieval Design:** Implemented few-shot prompting by retrieving stylistically similar past posts for the selected topic, language, and length, and used them to guide the LLM's tone, structure, and output quality.
- **Application Layer:** Deployed an interactive Streamlit application allowing users to select topic, length, and language, and generate on-brand posts on demand with minimal manual editing.

### YouTube Trending Data Pipeline on AWS – AWS S3, Lambda, Glue, Step Functions

[GitHub](#)

- **Pipeline Design:** Built a fully automated end-to-end data pipeline using AWS S3, Lambda, and Glue with the YouTube Data API to ingest, clean, transform, and store data.
- **Architecture & Orchestration:** Implemented Medallion Architecture (Bronze/Silver/Gold) with AWS Step Functions for reliable orchestration, scheduling, and retry logic.
- **Monitoring & Reliability:** Configured automated alerts across pipeline stages to ensure consistent, analysis-ready data with minimal manual intervention.

## Education

### B.Tech in Information Technology

*College of Technology, Pantnagar*

**CGPA: 7.56**

**2022 – 2026**

### Senior Secondary (Class XII)

*St. Joseph Convent School*

**92%**

**2021**

### High School (Class X)

*USR Indu Inter College*

**92%**

**2019**

## Position of Responsibility

### Convocation Digital Card Team Lead

- Led the end-to-end digital invitation and QR-based security verification system for the university convocation, coordinating a cross-functional team and eliminating paper-based workflows to contribute Rs. 1.5 lakh in cost savings via real-time entry validation and audit logs.

## Achievements

- **GATE:** Cleared **GATE CSE 2026** (94.3 percentile) and **GATE DA 2025** (86 percentile), qualifying among the top rankers nationally.
- **Certification:** Completed the **Foundations of Data Science and Applications** certification from the **Indian Institute of Technology (IIT) Madras**.
- **EdTech Initiative:** Built and scaled [Pantnagar Pathshala](#) with an associated [YouTube channel](#) that crossed **1,000+ subscribers**.