

Mansi Kulshrestha

+91 6395756963 — mansikulshrestha87@gmail.com — linkedin.com/in/mansi-kulshrestha-38388121a — github.com/mansikulshrestha

Objective

Data Analyst with strong proficiency in Python, SQL, and AI-powered tools, seeking to contribute to regulatory data transformation through GenAI, NLP, and automation. Experienced in building intelligent document pipelines and extracting insights from large structured and unstructured datasets.

Skills

AI & GenAI	OpenAI API, LangChain, Hugging Face, Prompt Engineering	Workflow & Pipelines	Apache Airflow, Snowflake, dbt
Data Tools	Python, SQL, Pandas, Power BI, Tableau, Scikit-learn	Languages	Python, C++ , SQL
Document Systems	Objective, Veeva Vault (familiarity)	Soft Skills	Attention to Detail, Communication, Critical Thinking

Education

Vellore Institute of Technology (Expected Graduation: 2025) Bachelor of Technology in Computer Science and Engineering Specialization: Artificial Intelligence and Machine Learning CGPA: 8.85 Relevant Coursework: Operating Systems, DBMS, NLP, Computer Networks, Cybersecurity, Data Mining	O
--	---

Experience

Clazar (March 2025 – Present) RevOps Intern Focus: CRM (HubSpot), Data Extraction, and Analysis <ul style="list-style-type: none">Extracted and analyzed structured data from HubSpot for performance metrics and reporting.Supported CRM hygiene and automated key RevOps workflows.Built dashboards to visualize pipeline velocity and deal health.	
--	--

Projects

GenAI Regulatory Document Classifier Tech Stack: Python, OpenAI API, LangChain, Veeva Vault API <ul style="list-style-type: none">Built a custom GPT-powered tool to classify, summarize, and search regulatory documents.Integrated with Veeva Vault-like mock storage to process and extract insights from unstructured content.Achieved 85% accuracy in document tagging and 40% time savings in manual review.	Mar 2025 – Apr 2025
LLM Evaluation Framework for Regulatory Content Tech Stack: GPT-4, Python, ROUGE, BLEU, Prompt Engineering <ul style="list-style-type: none">Created an evaluation system for comparing LLM outputs on regulatory content summarization.Included human-in-the-loop feedback, hallucination detection, and precision scoring.Helped improve response quality and factual accuracy through iterative prompt tuning.	Feb 2025 – Mar 2025
Movie Recommender System Tech Stack: Python, Pandas, NumPy, Scikit-learn, Streamlit, Render, Pickle <ul style="list-style-type: none">Built a content-based filtering recommendation system using cosine similarity.Engineered features from metadata and deployed as a web app.Link: CineSuggest	Jul 2024 – Aug 2024
Email Spam Classifier Tech Stack: Python, NLTK, Scikit-learn, XGBoost, Seaborn <ul style="list-style-type: none">Built multiple spam detection models with optimized TF-IDF text vectorization.Evaluated models using precision, recall, and confusion matrix metrics.	Jan 2024 – Mar 2024

Sales Insights Dashboard (Tableau)

Jul 2024 – Aug 2024

Tech Stack: Tableau, Excel, SQL

- Created interactive dashboards for sales trends and forecasting using mock datasets.
- Automated reporting and visual storytelling with KPI filters.

Co-Curricular Activities

- Organized technical events and industry speaker sessions on AI and data topics.
- Active participant in hackathons and data challenges.

Hobbies

- Gym, Doodling, Reading