

Deep

✉ kumardeep2003@gmail.com | 📞 8278279630 | 🔗 linkedin | 🎧 Deep841 | 🌐 Portfolio

Education

Thapar Institute of Engineering & Technology, Patiala Master of Computer Applications (MCA)	<i>Aug 2024 – May 2026</i> CGPA: 8.18
Kurukshetra University, Kurukshetra Bachelor of Science (B.Sc.)	<i>Aug 2021 – July 2024</i> Percentage: 64.2%

Experience

Machine Learning Intern | Cognifyz Technology (Remote) — [GitHub Link](#)

- Executed three ML projects in the restaurant domain covering regression, recommendation, and classification with complete data preprocessing and model evaluation.
- Built a rating prediction model using Linear Regression and Decision Tree Regressor ($R^2 = 0.98$) and a cuisine-based recommender using one-hot encoding and Jaccard similarity.
- Implemented a cuisine classification model via Decision Tree Classifier, optimizing accuracy through encoding, normalization, and handling class imbalance.

Projects

NLP Text Summarization with MLOps | Python, Transformers, Pegasus, FastAPI, Docker, DVC, MLflow, GitHub Actions — [GitHub Link](#)

- Fine-tuned Pegasus Transformer for abstractive text summarization in a configuration-driven MLOps pipeline.
- Automated data flow, training, and experiment tracking with MLflow & DVC for reproducible results.
- Deployed a real-time summarization API using FastAPI and Docker.

E-commerce Sales Analytics Dashboard | Power BI, MS SQL Server, DAX, Power Query, Data Modelling — [GitHub Link](#)

- Designed an interactive Power BI dashboard to analyze YTD Sales, Profit, Quantity, and Profit Margin across categories, segments, and regions.
- Integrated data from MS SQL Server and flat files, performed ETL in Power Query, and built a star schema model with a Date dimension.
- Created dynamic KPIs, trend lines, and slicers using advanced DAX for data-driven business insights.

End-to-End Medical AI Chatbot | Python, LangChain, Hugging Face, Pinecone, Flask, FLAN-T5 — [GitHub Link](#)

- Built an AI medical chatbot using The Gale Encyclopedia of Medicine for accurate, context-based health Q&A.
- Implemented a RAG pipeline with LangChain, Hugging Face embeddings, and Pinecone for semantic retrieval.
- Developed a Flask web app for real-time user interaction with secure API key handling.

Technical Skills

Programming: C++, Python, SQL

Tools & Platforms: VS Code, Jupyter, Git/GitHub, MySQL Workbench, Docker, FastAPI

Data & Visualization: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Power BI

ML & NLP: Scikit-learn, Random Forest, Logistic/Linear Regression, KNN, NLTK

Core CS: OOP, Software Engineering, DBMS

Certificates

- **Kaggle** — Intro to Machine Learning — [Certificate Link](#)
- **Deloitte Australia (Forage)** — Data Analytics Job Simulation — [Certificate Link](#)
- **Tata (Forage)** — GenAI Data Analytics Job Simulation — [Certificate Link](#)