

# Aman Tripathi

M.Sc. Data Science  
Manipal, Karnataka

+91-9005330740

amantripathi3273@gmail.com

linkedin.com/in/aman-tripathi27

## EDUCATION

---

### •Master of Science in Data Science

2024–2026

*Manipal Academy of Higher Education, Karnataka*

### •Bachelor of Science in Computer Science, Statistics, and Mathematics

2021–2024

*Chandigarh University, Mohali*

## TECHNICAL SKILLS

---

**Languages:** Python, SQL

**Machine Learning & AI:** PyTorch, TensorFlow, scikit-learn, model building & evaluation, predictive analytics, algorithmic solutions

**NLP / GenAI:** spaCy, Hugging Face Transformers, LLM APIs (OpenAI/Gemini)

**Data Engineering & Viz:** ETL pipelines, data transformation, EDA, Power BI, Excel, data storytelling

**Cloud / Dev:** AWS (S3, Athena, Redshift, SageMaker), Git, Streamlit deployment, Jupyter Notebook

## EXPERIENCE

---

### •Data Science Intern

Oct 2023 – Nov 2023

*CodeClause*

Remote

- Built reusable Python workflows applying cutting-edge data processing techniques; demonstrated hands-on, growth mindset in tackling real-world data challenges.
- Developed algorithmic solutions for data transformation and EDA; took ownership of technical implementations delivering stakeholder-ready insights.
- Collaborated with cross-functional teams; communicated technical findings effectively bridging data science and business impact.

## DATA & AI PROJECTS

---

### •Public Perceptions of Government Health Policies | Python, spaCy, scikit-learn

2025

- Built end-to-end NLP pipeline on noisy Reddit data: processed **735 comments** (South Karnataka) expanding to **1,196** for analysis; demonstrated ability to work with complex, real-world datasets.
- Applied cutting-edge machine learning techniques benchmarking **5 ML models** (SVM, DT, NB, KNN, RF) vs. **DistilBERT**; SVM achieved **41.5%** accuracy, DistilBERT **40.4%** showing strong analytical and technical skills.
- Developed algorithmic solutions for semantic clustering using **TF-IDF + Zipf filters**; compared **K-Means (94.6%)** and **BERTopic (52.5%)** topic consistency demonstrating innovation in problem-solving.
- Communicated findings through peer-reviewed publication (IEEE CIES 2025); translated theoretical knowledge into impactful applications contributing to thought leadership.

### •Drug–Drug Interaction Prediction using Graph Neural Networks | RDKit, PyTorch Geometric

2025

- Parsed DrugBank SMILES/XML, converted molecules to graphs via **RDKit**; engineered GNN-ready dataset demonstrating data engineering skills and hands-on execution.
- Trained **Graph Autoencoder** applying cutting-edge machine learning to learn molecular embeddings and infer adverse drug interactions; developed algorithmic solution bridging graph theory and deep learning.
- Evaluated model with standard link-prediction metrics (ROC-AUC / AP); demonstrated sharp analytical skills and ability to implement solutions creating real business impact in healthcare domain.
- Deployed interactive **Streamlit** app with real-time inference on **Hugging Face Spaces**; built scalable data & AI solution showing technical expertise and deployment capabilities.

### •Clinical Trial Data Quality Tool – Novo Nordisk GBS Hackathon | Python,ML

Sep–Oct 2025

- Led cross-functional team developing PRAGNA, automated eCRF comparison tool; demonstrated collaborative spirit and strong communication skills managing stakeholders under tight deadlines.
- Built algorithmic solution implementing automated ETL pipeline and data quality validation framework; applied cutting-edge data processing techniques reducing manual effort.
- Took ownership of project delivery driving real business impact for clinical trial submissions; demonstrated hands-on mentality and ability to contribute from day one while learning from experienced domain experts.

## ACADEMIC AND PROFESSIONAL CONTRIBUTIONS

---

### • Novo Nordisk GBS Hackathon 2025 – Team Lead

Sep-Oct 2025

*Led development and delivery of working prototype under tight timelines; demonstrated leadership, collaborative mindset, and ability to drive data & AI transformation projects creating tangible outcomes.*

### • IEEE CIES 2025 – Published Research

2025

*Published peer-reviewed paper on large-scale NLP pipeline; contributed to thought leadership demonstrating technical communication skills and ability to translate complex analytical work into impactful applications.*