

Soumyodeep Dey

+91-9830193919 | soumyodeepdey96@gmail.com | [LinkedIn](#) | [GitHub](#) |

Education

MS(R) | Artificial Intelligence

Indian Institute of Technology, Delhi

07/2023 - Present

CGPA: 7.86

B.Tech | Computer Science and Engineering

National Institute of Technology, Durgapur

07/2018 - 05/2022

CGPA: 7.99

Master's Thesis

Dynamic Spatio-Temporal Graph Neural Network for Cold Wave Prediction

- Created a dual-task ST-GNN for both cold wave classification and temperature regression, implementing a hybrid GRU-GraphSAGE model in PyTorch Geometric.
- Implemented a novel dynamic graph formulation where the graph topology evolves to model spatial dependencies, leading to a 36% F1-score improvement over statically formulated graphs and 25% F1-score improvement over time-series based Transformer models like Informer.

Internships

Research Intern at Samsung Research Institute, Bangalore

Jan 2022 – July 2022

- Conducted research in cross-device user behavior by analyzing large-scale search logs and time-series sensor data. Developed Support Vector Machine and Gradient Boosted Decision Tree models to predict search task continuation.

Research Intern at IIT Patna

Mar 2021 – Aug 2021

- Created the CESAMARD dataset by crawling and annotating multimodal Amazon reviews for complaint, emotion, and sentiment. Developed an attention-based multimodal model using adversarial multi-task learning leveraging SBERT, ResNet, and a dual attention mechanism with BiLSTMs achieving F1-score of 89.07%, beating single task baselines.

Projects

Image Captioning using Encoder-Decoder Frameworks

- Implemented an encoder-decoder architecture for image captioning on the Flickr8k dataset using PyTorch, experimenting with ResNet + LSTM and ViT + Transformer models. Used beam search for inference, achieving BLEU score up to 0.35. Deployed using Streamlit for real-time image uploads and caption generation.

Vehicle Insurance Prediction Pipeline

- Implemented an MLOps pipeline to train and deploy an XGBoost model for predicting vehicle insurance purchases. The system automates the entire workflow, including data ingestion, validation, feature engineering, and model training. Implemented a CI/CD pipeline with GitHub Actions to deploy the final model as a Dockerized FastAPI service on AWS.

GNN-based Movie Recommendation System

- Built LightGCN-based movie recommender using PyTorch Geometric, modeling user-item interactions as a bipartite graph. Trained using Bayesian Personalized Ranking loss function and deployed as a web interface using Streamlit.

RAG Chatbot for research domain question answering

- Built and deployed an end to end Retrieval Augmented Generation system for answering questions on papers. The backend, built with FastAPI, uses LangChain for orchestration, OpenAI for generation, and Pinecone for vector search.

End-to-End Machine Learning pipeline for Text Summarization

- Built an MLOps pipeline to train and deploy a T5-based text summarization model. Utilized AWS S3 for data storage and MLflow for experiment tracking and model registry. Implemented a CI/CD pipeline with GitHub Actions to trigger automated model retraining on new data pushes. Deployed the final model as a REST API using FastAPI and Docker.

Publications

Sentiment and Emotion-Aware Multi-Modal Complaint Identification

A. Singh, **S. Dey**, A. Singha, S. Saha. *AAAI Conference on Artificial Intelligence*, 2022.

Mitigating DoS Attack in MANETs Considering Node Reputation with AI

S. Joardar, N. Sinhababu, **S. Dey**, P. Choudhury. *Journal of Network and Systems Management*, 2023.

Technical Skills

Languages: C++, Python, SQL **Databases:** MongoDB, Pinecone, Neo4j **Libraries:** NumPy, Pandas, Matplotlib, Scikit-learn, NLTK, XGBoost, OpenCV, HuggingFace **Frameworks:** PyTorch, TensorFlow, Flask, LangChain, LangGraph
Tools/Technologies: Jupyter, Git, Docker, Kubernetes, MLflow, DVC, Apache Spark, FastAPI, AWS (SageMaker, S3, EC2)
Relevant Courses: Machine Learning, Deep Learning for NLP, Optimization, Graph Machine Learning, Reinforcement Learning, Biomedical Data Analysis, Special Modules on Visual Computing, Ethics in AI, Theoretical Machine Learning