Dharmendra Mishra

Email: dharm161024@gmail.com / Ph.no (+91) 8770508992 Location: New Delhi
Linkedin: https://www.linkedin.com/in/dharm1287/

PROFILE

Experienced data scientist with a robust background in natural language processing and generative AI. Seeking a dynamic and stimulating role that fosters continuous learning and allows me to apply my knowledge and experience.

TECHNICAL SKILLS

Programming Languages: - Python and SQL.
Machine Learning Frameworks: - sklearn, pytorch, tensorflow
NLP and GenAI Skills: Semantic search, Text Mining, RAG & Agentic Systems, Information Retrieval
Classical ML: Regression (Linear, Logistic), Clustering, Modelling, Optimization
Other Tech: - Git, Basic linux, Azure, Databricks, langchain, Django, Flask, FastAPI

EXPERIENCE

Data Scientist (Aug 2024-Present) **Hexaware, Pune**

- Developing generative AI content creation solutions for product and marketing teams.
- Building an agentic tool that can analyze large unstructured documents, extract all key information, their implications & organize the details helping businesses review and make decisions faster.
- Developed RAG models for intelligent querying of financial reports. Used following metrics: Mean Reciprocal Rank, ROUGE Score, LLM-as-a-judge and Query response time for evaluation.
- Developed NLP applications with LangChain and HuggingFace models for retrieval-augmented Q&A and agentic workflows.
- Experimented with state-of-the-art LLMs (GPT-4, Mistral, Falcon), embedding models, prompting (Zero shot, CoT, ToT) & reranking techniques for improving retrieval and generation for domain-specific tasks.
- Partnered with business stakeholders to understand requirements and develop aligned AI solutions.

Project Engineer – Data Scientist (May 2022 - Aug 2024) **CDAC, Delhi**

- Developed machine learning pipeline for classification of focussed stimuli in BCI.
- Developed the frontend of speller application in Psychopy.
- Utilized Python's Scikit-learn library to implement various machine learning algorithms.
- Incorporated features extracted from preprocessed EEG data into the pipeline.
- Optimized hyperparameters and model performance through cross-validation.
- Evaluated the trained model's performance on unseen data to assess its classification accuracy.
- Implemented image analysis, tissue segmentation, and feature extraction to enhance the interpretability of fMRI and digital neuropathology data.
- Developed and integrated generative models for synthetic data generation and augmentation to enhance training datasets.
- Implemented RAG using LLMware along with ChromaDB to construct the RAG system to serve tasks of retrieving information about financial reports.
  Tools: Python, Scipy, Nipype, Pydicom, MLFlow, MNE & Psychopy, RAG, LLMware.

Python and AI Developer (February 2021 - May 2022) **Slickbooks, Bangalore**

- Performed financial data analysis and predictive modeling.
- Developed semantic models and Named Entity Recognition (NER) pipelines for extracting structured insights from textual data: financial documents, invoices, and reports.
- Designed algorithms for invoice processing and expense categorization.
- Optimized data pipelines to enhance NER accuracy and streamline model training processes.
- Integrated Azure services for OCR (Optical Character Recognition).
- Integrated Azure Cognitive Services for OCR and NLP workflows, aligning with enterprise AI adoption strategies.
- Developed and integrated AI models to automate accounting processes and enhance financial reporting efficiency.
  Tools: Python, Spacy, Azure,

EDUCATION

2021-2022 - Post Graduate Diploma in Artificial Intelligence – CDAC, Pune, India
2016-2020 - Bachelor's of Engineering(Mining) – University Institute of Technology Shahdol, MP, India

CERTIFICATIONS:

- Google Professional Data Engineer, GCP
- Azure AI Engineer Associate, Microsoft
- Data Scientist Associate, Datacamp

TEACHING EXPERIENCE: Course Instructor – Course on Responsible AI under ITEC scheme of MEA

- Designed and delivered comprehensive course content on Responsible AI including bias, fairness, explainability techniques, privacy enhancing strategies, and machine unlearning.
- Developed practical projects and case studies to enhance senior government officials' understanding.
- Led lectures and facilitated discussions on ethical implications, laws and best practices concerning AI usage.

PUBLICATIONS & WORKSHOPS:

- 'Neuroscientific Study with Hybrid models Validation for Deception Detection' (4th International Conference on Emerging Trends and Technologies on Intelligent Systems (ETTIS 2024))
- EEG Signal Analysis Using Quantum Fourier Transform and Quantum Machine Learning (IQCRJ-International QC Research Journal)
- Presented our work: Multimodal learning for Cognitive State Estimation at The 9th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics(NCVPRIPG) held at IIST Thiruvananthapuram