

VISHAL SHARMA

jamvishalsharma.vs@gmail.com | +91-7890942588 | <https://github.com/vishalssharma> | <https://linkedin.com/in/vishalssharma>

SKILLS

Python | PyTorch | scikit-learn | Machine Learning | Natural Language Processing | Deep Learning | Prompt Engineering | LLMs | PEFT | Transformers | LangChain | LoRA | FastAPI | Spark | Big Data | Docker | SQL | Elasticsearch | Git | AWS

WORK EXPERIENCE

Informatica, Bangalore: CLAIRE AI | Machine Learning Engineer

Sep 2021 - Present

- Responsible for developing, fine-tuning, testing and deploying large language models (LLMs), NLP solutions and analytics.
- Led the end-to-end development and deployment of CLAIREGPT, right from data generation, data curation, fine-tuning and optimizing two open-source LLMs for specialized tasks including text translation, classification and summarization.
- Engineered and refined prompt strategies, conducted in-depth performance benchmarking, developed interactive interfaces facilitating seamless model evaluation and user interaction for dev and qa testing.
- Contributed to four key projects - CLAIREGPT, CDGC, CDI, and MDM - demonstrating a strong track record of delivering impactful machine learning solutions across diverse domains.

Informatica: CLAIRE AI | Software Development Intern

Sep 2020 - Aug 2021

- Built and implemented multiple projects from proof-of-concepts (POCs) to production for AI integration across diverse products.
- Designed an NLP solution specifically for generating mock data for banking clients, ensuring statistical accuracy and pattern fidelity while preserving data confidentiality.
- Led data masking and desensitization initiatives to uphold data privacy, and performed advanced NLP tasks, including Named Entity Recognition (NER) and Part-of-Speech (POS) tagging, to enhance AI capabilities.

Higradius | Software Development Intern

Apr 2020 - Aug 2020

- Developed an autonomous Accounts Receivable (AR) system that significantly boosted productivity and efficiency, enabling AR analysts to handle a higher volume of transactions with greater speed and accuracy.

PROJECTS

CLAIREGPT - Multi Step Reasoning and NL2SQL

CLAIREGPT takes in user queries through chat interface and using LLM best practices, retrieves the required information to show to the user.

- Performed data generation, data curation, prompt engineering, fine-tuning LLMs using SOTA techniques to process user queries in a multi-step pipeline.
- The first model handled tasks such as greeting the user, intent identification, query segmentation and query summarization by referencing previous chat history. This improved the user interaction experience and laid the foundation for downstream tasks.
- With a second specialized LLM for Natural Language to intermediate DSL and SQL (NL2SQL) translation, enrichments etc. Integrated with the Reasoning LLM model, this system converted summarized user queries into SQL and DSL, enabling precise querying and used a RAG based approach for content enrichment.

Data Governance and Catalog - Schema Matching

- Enhanced a rule based system with an ML-powered solution to predict similar, joinable and unionalbe tables across schemas using just the metadata and data profiling results.
- Handled over 1 million plus records using Apache Spark in cluster mode achieving a 70% reduction in processing time using an intelligent blocking layer, saving thousands of dollars in operation costs annually.
- Leveraged AWS cloud services such as S3, EC2 and SageMaker to build, deploy and scale ML models.

EDUCATION

KIIT University, Bhubaneswar, India

Graduated: 2021

Bachelor of Technology (B.Tech) in Computer Science | CGPA: **9.22**

- Relevant Coursework: Object Oriented Programming, Data Structures and Algorithms, Operating Systems, Computer Networks, Machine Learning.