SHREYA SHARMA

ss.shreyasharma1609@gmail.com | +1 (640) 204-9203 | linkedin.com/in/shreya916 | github.com/lshreyasharmal | https://lshreyasharmal.github.io

EDUCATION

Masters in Computer Science, Johns Hopkins University, USA **Aug 2022 – May 2024 (Expected)**

Courses – Machine and Deep Learning, Natural Language Processing (Transformers, Machine Translation), Cloud Computing, Software System Design, Parallel Computing for Data Science, Artificial Intelligence System Design

Bachelors in Computer Science and Engineering, IIIT Delhi, India **Aug 2015 – Aug 2019**

Courses - Data Structures, Advanced Algorithms, Operating Systems, Machine Learning, Natural Language Processing, Databases, Big Data Analytics

TECHNICAL SKILLS

- **Programming Languages** – Python (Proficient), Java (Proficient), C/C ++(Intermediate), R, JavaScript, HTML/CSS, SQL, MATLAB
- **Data Science Tools** – PyTorch, TensorFlow, scikit-learn, scipy, matplotlib, pandas, HuggingFace, LangChain, Whisper API, OpenAI, OpenCV, Llama-2, Apache Spark, Apache Hadoop, Apache Kafka, Dask, LaTeX, TensorBoard, Jupyter Notebooks, Google Colab
- **Software Engineering Tools** – ReactJS, Django, Flask, Bootstrap, Spring-Boot, Reactive Spring, REST services, Celery, AWS, boto3, Grafana, Git, Docker

PROFESSIONAL EXPERIENCE

AI/ML Research Engineer **CLSP, Johns Hopkins University** **Sept 2022 – Present**

Entailment Verification

• Designed and Developed a ReactJS application for an Amazon Mechanical Turk Task to collect human annotations for a new benchmark for common sense reasoning – to identify flawed scientific proofs produced from generative AI systems to answer MCQs based on reasoning.
• Employed OpenAI API (GPT-3.5) using LangChain framework to score truthfulness and validation of scientific statements.
• Tuned BERT-based and prompt-tuned T5 models for the classification of premise statements as entailments to a hypothesis or not, achieving a 0.73 f1-score and 0.90 ROC_AUC with Roberta-base model.

Open-Domain Long-Form Question Answering

• Fine-tuned T5-base model on ELI5 data for answering questions that require information from multiple documents to get a conclusive answer.
• Implemented Sparse-Retrieval Index (BM25) and Dense-Retrieval Index (FAISS) to fetch relevant Wikipedia articles.
• Evaluated models using ROUGE, BLEU, and BERT Score, ensuring robust performance in open-domain question answering.

Senior Software Developer **MakeMyTrip.com** **June 2019 – July 2022**

Booking Reports Management System

• Developed travel booking and reporting back-end micro-services on Spring Boot and Reactive Spring.
• Implemented advanced features using JPA Criteria queries for search, filter, count, and download data from multiple databases, reducing response time by **0.6 seconds and boosting platform traffic by 25%.**

Multi-Level Approval Workflow System

• Developed an end-to-end multi-level approval workflow with Activiti and Spring Boot for approving bookings made on the platform.
• Created a dynamic 'Policy Management System' for corporate travel policies and implemented efficient business logic to automate approval flow for **30,000 daily booking requests**.

Expense Management System

• Developed Reactive Spring micro-service and MySQL database for managing user reimbursements in personal and business travel bookings. **Received GoTripper of the Month award for project success.**
• Developed wrapper APIs for in-house lines of businesses (flights, hotels, buses, cabs) integrated into MakeMyTrip's corporate platform, resulting in a **10% boost in client acquisition.**

Software Integrations

• Built an automated pipeline to manage **1M daily** wallet transactions with a fallback Kafka system achieving **35% reduction in issues.**

PUBLICATIONS

- **S. Sharma** and M. Mohania, Comparative Analysis of Entity Identification and Classification of Indian Epics, ICMI'22, ACM
- A. Iyer, K. Gupta, **S. Sharma** et al., Integrative Analysis and Machine Learning Characterization of Single Circulating Tumor Cells, J Clin Med, Apr'20
- S. Trivedi, S. Bhola, …, **S. Sharma**, Predictive Maintenance of Air Conditioning Systems Using Supervised Machine Learning, ISAP'19

PROJECTS

Comprehend.it [Report]

• An education platform build on Django to transcribe videos and generate quizzes/notes/summaries from videos using Whisper API, Llama-2 and GPT-3.5 integrated using LangChain for prompt-engineering.

LangPrep – Language Preparation using AWS [Link]

• Django application designed for learning word pronunciation in diverse languages. Hosted on AWS, it utilizes Amazon Textract for Optical Character Recognition (OCR) to extract text from documents, while Amazon Polly converts the text to speech. All documents are stored in an S3 bucket.

Language-Agnostic Sentence Embedding Generator [Report]

• Trained a GAN network to create a language agnostic embedding generator. Trained multi-lingual BERT based models as embedding generator and dense classifiers as the discriminator to identify source language of a text.

Unsupervised Monocular Depth Estimation from Video [Poster]

• Enhanced depth estimation for monocular videos using self-attention and vision transformer, outperforming the baseline on KITT dataset. Implemented self-attention layers and ViT as image encoder with adversarial loss.

LEADERSHIP

- Treasurer at Graduate Association of Women in Computer Science & Electrical and Computer Engineering, JHU, 2023-2024
- Course Assistant for Introduction to Human Computer Interaction taught at Johns Hopkins University.