Atharva Dashora

dashoraatharva1@gmail.com, f20210127@hyderabad.bits-pilani.ac.in, 8790401979

EDUCATION

BITS PILANI
Hyderabad, India
B.E in Computer Science with a minor in Data Science
Expected July 2025
CGPA : 8.61/10

SKILLS

LANGUAGES
• C • C++ • Python • Java • C#

FAMILIAR WITH
• Spacy • Numpy • Pandas • TensorFlow • Lucene • HTML • CSS • JavaScript • .NET

LINKS

Github: **Atharva-Dashora**
LinkedIn: **Atharva-Dashora**
Leetcode: **Leetcode**

COURSEWORK

UNDERGRADUATE

Data Structures and Algorithms
Object Oriented Programming
Operating Systems
Database Systems
Machine Learning
Information Retrieval
Computer Architecture
Microprocessors and Interfaces
Computer Networks
Compiler Construction
Design and Analysis of Algorithms
Theory of Computation
Foundations of Data Science
Applied Statistical Methods

INTERESTS

Machine Learning
Data Engineering
Information Retrieval
Natural Language Processing

EXPERIENCE

WALMART GLOBAL TECH | Developer intern
Jan 2025 - Present

- Intern at the Software Engineering Team.

BARCLAYS | Explorer Intern (Developer)
May 2024 - July 2024

- Developed an autosys job to determine why trades go missing post-execution.
- Developed multiple batch ETL pipelines to streamline the analytics for users in post-trade functions which led to a 95% decrease in time.
- Engineered a data display interface that processed and presented outputs from ETL jobs, allowing for real-time analysis of 500K+ records; enhanced data accessibility for 15+ team members in decision-making processes.

ONFINANCE | ML Intern
May 2023 - July 2023

- Fine-tuned an LLM(Falcon 7b) using the knowledge on transfer learning with financial data on stocks for customer support.
- Optimization of training efficiency on lower-end GPUs by using PEFT and QLoRA which led to a 90% decrease in training time.

PROJECTS

RANKING ALGORITHMS | Information Retrieval

- Created a comprehensive suite of models for medical information retrieval, resulting in a streamlined process that facilitated efficient data handling for over 10,000 medical documents.
- Used Knowledge graphs and Entity Based Retrieval Models to expand existing queries and improve accuracy and precision.

UNSUPERVISED MACHINE TRANSLATION | Natural Language Processing

- Developed an unsupervised translation system using denoising autoencoders, adversarial training, and word alignment with Procrustes refinement.
- Designed shared encoder-decoder architectures and improved robustness with noise injection techniques.
- Built a CLI for training, translation, and evaluation across English, French, and German.

SUPERVISED OPENIE | Natural Language Processing

- Built Word2Vec models (Skip-Gram and CBOW) with Softmax and Negative Sampling, achieving an MRR of 0.49 and reducing training time by 60% using optimized batching and preprocessing.
- Designed a supervised OpenIE system using BERT embeddings and a BiLSTM-CRF model, achieving 61% accuracy on validation data with error analysis to refine extractions.
- Conducted extensive data preprocessing, model evaluation, and performance comparison across metrics like MRR and F1-score.

DISTRIBUTED GRAPHDB WITH LOAD BALANCER | Operating System

- Simulated an application for a distributed graph database system involving a load balancer process, a primary server process, two secondary server processes.