ISHITA RAJ

ishitaraj5173@gmail.com ◇ ishita-raj-134348200 ◇ +91 (8969121788)

EDUCATION

Dual Major in Electrical & Electronics Engineering and Mathematics Aug 2021 - July 2026
BITS Pilani *India*

Minor in Data Science Aug 2023 - July 2025
BITS Pilani *India*

RESEARCH EXPERIENCE

Undergraduate Student Researcher Aug 2024 - *Current*
NLP Research Group | *Prof. Prajna Devi Upadhyay*

- **Machine Translation of Low Resource Languages**
  - Developed AI-based translation tools for Mundari, a low-resource tribal language, under the **Ministry of Tribal Affairs** initiative to enhance digital accessibility for underrepresented languages.
  - Fine-tuned Meta's Fairseq m2m model for Mundari, integrating a custom contrastive loss function to enhance cross-lingual representation learning and mitigate overfitting in low-data regimes.
- **Custom-Built Semantic Framework: Developing Open Information Extraction and Word2Vec Models from First Principles**
  - Developed a Supervised Model for Open Information Extraction (Open IE), converting sentences into structured tuples (< subject, relation, object, time, location >) and evaluated its accuracy using the CaRB metric which gave a **precision of 64.6%**.
  - Implemented Word2Vec (Skip-Gram & CBOW) with Softmax and Negative Sampling, analyzing MRR on WikiText-103 to optimize window size, observing a **9% MRR drop** in Skip-Gram (0.1780→0.1628) and a CBOW improvement (0.1255→0.1618) with larger contexts.

Student Research Associate Dec 2024 - *Current*
Neural Systems & Pattern Recognition Lab | *Prof. Aneesh Sreevallabh Chivukula*

- **Computational Sansktrit**
  - Developed a novel binary embedding framework that leverages transaction/word/graph representations using *Dhatu Patha* methodologies to generate verb forms.
  - Engineered a robust encoding scheme using OKM methodologies paired with custom software-based decoding, ensuring efficient and accurate data transformation.
- **FPGA Implementation of Hand Written Text**
  - Implemented an adversarial network for FPGA-based handwritten digit recognition using reconfigurable IP cores and optimized memory techniques for real-time inference.
  - Leveraged adversarial training to enhance model robustness and security while ensuring efficient hardware utilization on resource-constrained platforms.

Principal Component Decomposition of Electrophysiological Signals: Advanced Spike Sorting with MountainSort5 Aug 2024 - Dec 2024
NeuroPharmacology Lab |*Prof. Srinivas Prasad K*

- Developed a data science pipeline to analyze electrophysiological data from 24 brain activity channels, utilizing spike sorting to filter noise and extract accurate spikes with associated timestamps.
- Developed an adaptive clustering approach leveraging PCA, t-SNE, and UMAP, improving spike classification granularity for low-SNR datasets.

STUDY REPORTS

- Analysis of Spotify's Hybrid Server Architecture and Application of Queuing Models in Improving User Satisfaction

WORK EXPERIENCE

Summer Research Intern May 2023 - Jul 2023
Bhaskaracharya National Institute for Space Applications and Geo-informatics *Gandhinagar, Gujarat*

- **3D Visualization of Large-Scale Weather Data for Regional Pattern Analysis**
  - Developed an interactive 3D terrain-based weather visualization that transforms 50,000+ data points into a dynamic geospatial map, enabling real-time analysis of temperature, wind speed, pressure, humidity, and density over varying elevations.
  - Leveraged Plotly's WebGL rendering for smooth interactivity, optimizing large-scale data visualization for climate analysis.
  - **Tools Used:** Python, Plotly, Pandas, NetCDF, Cesium.js

OTHER PROJECTS — *All Projects*

Sentimental Analysis using IMDB Review Oct 2024 - Nov 2024
Personal Project

- Developed a cutting-edge NLP module to deliver granular sentiment analysis, readability scoring, and lexical feature extraction, seamlessly exporting insights to Excel.
- Optimized data ingestion workflows with rigorous input sanitization, comprehensive error handling, and modular design, ensuring scalable and reliable performance.

Automated NLP Pipeline for Web Content Extraction and Analysis Jun 2024 - July 2024
Personal Project

- Engineered robust data handling—from URL input via Excel to comprehensive processing and export of structured results in Excel—enabling efficient, scalable insights from diverse online content for computational linguistics research.
- Developed an end-to-end web scraping and text analysis pipeline in Python, to automate HTML retrieval, text storage and perform sentiment analysis, readability assessments, and feature extraction.

Time Series Analysis for Prediction of Daily Maximum Temperature May 2024 - Jun 2024
CSIS Department — Prof. Akanksha Rathore

- Trained ARIMA models on daily max temperature data, reducing Mean Absolute Error (MAE) by **47%** when increasing training data from 3 to 27 data points.
- Demonstrated superior short-term predictive performance compared to a weather app, with ARIMA achieving lower prediction errors over multiple test cases.

SKILLS

Programming Languages	Python, C/C++, Matlab, HTML/CSS
Software	Vivado, LTSpice, Simulink
Libraries	PyTorch, NLTK, SciPy, GeoPandas, Mountainsort5

CERTIFICATIONS

- **PyTorch for Deep Learning Bootcamp -** Udemy - ZTM
- **Supervised Machine Learning -** Standford - Coursera
- **Introduction to Artificial Intelligence (AI) -** IBM

POSITIONS OF RESPONSIBILITIES

Student Coordinator May 2023 – Aug 2024
Admissions Division, BITS Pilani Hyderabad Campus

- Streamlined the document verification and onboarding process for 1200+ freshmen, ensuring 100% compliance and reducing processing time by 20%.
- Led and coordinated a team of 90+ volunteers, successfully optimizing the admission process for 1200+ students, resulting in a 15% increase in efficiency.