



# YUSUF BAIG

+1 (929) 701-6567 • yb2510@nyu.edu •  YusufBaig7 •  mirzayusuf • <https://yusufbaig7.github.io/>

## EDUCATION

**New York University (Courant Institute of Mathematical Sciences)**, New York City, NY May 2026  
Master of Science in Computer Science  
*Coursework:* Operating system, Programming Languages, Natural Language Processing

**Manipal Institute of Technology**, Udupi, India May 2022  
Bachelor of Technology in Information and Technology, Minor in Computational Intelligence  
*Coursework:* Linear Algebra, Distributed Systems, Statistics, Machine Learning, Computer Vision

## EXPERIENCE

**Multimodal Digital Media Analysis (MIDAS) Lab** **Delhi, India | February 2022 – August 2024**  
Research Assistant – Legal AI Advisor: Dr. Jessy Li, Dr. Rajiv Ratn Shah

- Implemented GNNs for Legal Judgment Prediction, leveraging node classification techniques with advanced XLNet-based embeddings to model judicial case structures, achieving a macro F1 score of 75%.
- Developed and evaluated zero-shot Legal Judgment Prediction methods using multiple LLMs, leveraging techniques such as Rhetorical Role structuring, Contextualized Case-aware Templates (CCaT), and metadata integration to simulate judicial reasoning processes, achieving a 10% improvement in F1 score compared to traditional approaches.

**Carl Zeiss R&D** **Bangalore, India | January 2022 – June 2024**  
Software R&D Engineer – Image Processing and Backend

- Studied and implemented advanced image processing techniques, mainly calculating a 3D volume's Contrast Discrimination Function and Modulation Transfer Function for an Aerospace client.
- Formulated and executed a complex set of algorithms targeting pixel intensity adjustments, resulting in a streamlined workflow that saved an average of 10 hours in a 30-hour project timeframe during the image processing phase.

**Samsung Prism Project** **Udupi, India | August 2021 – December 2021**  
Undergraduate Research Assistant - Natural Language Processing

- Developed a Multilingual Word Sense Disambiguation system by enhancing the Lesk algorithm, improving accuracy across three languages, and created 'hindiwsd,' a Python-based framework for transliteration, spell correction, POS tagging, and word sense disambiguation of Hindi text, enabling advanced natural language processing for Hinglish code-mixed text.

## PUBLICATIONS AND RESEARCH PROJECTS

**Zero-shot Legal Judgment Prediction: How ready are LLMs on legalese?**  
*[UNDER REVIEW]*

**Exploring Graph Neural Networks for Indian Legal Judgment Prediction**  
*PrePrint: arXiv 2023*

**HindiWSD: A package for word sense disambiguation in Hinglish & Hindi**  
*LREC 2022*

**Severity Classification of Mental Health-Related Tweets**  
*IEEE 2021, ICML CA2MH 2021*

**Curb Your Carbon Emissions: Benchmarking Carbon Emissions in Machine Translation**  
*arXiv 2021*

**Assessing the Carbon Intensity of Models Across Different Languages**  
*ICML WiML 2021*

## PROJECTS

**Development of the MAPLE Score for Evaluating LLMs in Mathematical Reasoning**

- Designed the MAPLE score, a holistic metric to assess reasoning misalignment in LLMs by integrating error rates, redundancy, and validity.
- Developed a multi-stage evaluation framework to identify reasoning pitfalls and improve LLM performance on complex mathematical tasks.

## TEACHING EXPERIENCE

CSCI-UA 381 Programming Tools for the Data Scientist January 2025 - Present

## SKILLS

<b>Languages</b>	Python, C++/C, C#, MATLAB, SQL, Java, HTML, CSS, $\LaTeX$ , Racket
<b>ML Frameworks</b>	PyTorch, HuggingFace, pandas, LangChain, NumPy, scikit-learn, seaborn, matplotlib
<b>Frameworks</b>	HPC, Git, GitHub, Docker, Flask, Jira, BitBucket, .NET, MongoDB, AWS, Linux