Research Interests and Student Opportunities

Hrishikesh Terdalkar
Assistant Professor
Department of Computer Science and Information Systems
BITS Pilani, Hyderabad Campus, India

https://hrishikeshrt.github.io

I am broadly interested in collaborating on and/or supervising projects that align with one or more of the following areas:

1. **Large Language Models**
   - Evaluation methodologies
   - Fine-tuning strategies
   - Benchmarking
   - Explainability
   - Hallucinations mitigation
2. **NLP-Oriented Tool Building**
   - Development of web-based systems/Python libraries to support NLP workflows
   - Annotation tools
   - Agentic Systems
3. **Knowledge Graphs**
   - Construction (rule-based, automated, or hybrid)
   - Graph repair (error detection and correction)
   - Knowledge-graph-based question answering
4. **Computational Linguistics**
   - Empirical and theoretical exploration of linguistic phenomena in Indian languages, including but not limited to tasks such as
     - Dependency parsing
     - Named entity recognition
     - Grammatical error correction
     - Lexical or semantic grouping
     - Discourse analysis
   - Evaluation of existing, and development of new metrics more suited for various linguistic tasks in Indian languages
5. **Information Retrieval**
   - Traditional retrieval methods
   - Retrieval augmented generation (RAG), including text-based and graph-based variants
   - Generation-supported retrieval pipelines
6. **Creative Text Generation**
   - Generation of poetry, genre-specific prose, and other forms of creative writing
   - Generating novel content, devising metrics for evaluation of the same

Indian Language NLP

- Any task centered around one or more Indian languages

Note

1. These areas often intersect, and interdisciplinary proposals are welcome.
2. My primary focus is on Sanskrit, followed by other Indian languages (in addition to Hindi), and then English.
3. I am generally not interested in projects involving only English.

Useful Prerequisites

If you are a student,

Having prior exposure to some or many of the following concepts and skills will be useful for getting started faster.

1. Foundations of Deep Learning: An understanding of neural networks, the intuition behind backpropagation, and common evaluation practices
2. Transformers and Core NLP Models: Familiarity with transformer architectures, language modeling objectives, and sequence-to-sequence task formulations
3. Practical Experience with LLMs: Hands-on experience running language models locally (e.g., programmatic inference, loading models, working with tokenizers) using frameworks such as PyTorch, Hugging Face, or similar toolkits (ollama, lang-chain)
4. Familiarity with usage of tools such as LaTeX, Git, formats such as JSON, YAML, experience with Python, Jupyter notebooks