Natural Language Processing - Theory and Applications [CS-2385] [CS-2385-1]

Faculty Name

Lipika Dey [lipika.dey@ashoka.edu.in], Diya Roongta [diya.roongta_ug24@ashoka.edu.in], Niranjan Rajesh [niranjan.rajesh_asp24@ashoka.edu.in];

Overview

Natural Language Processing - Theory and Applications

1. Basics of NLP : Syntax and Semantics - Parsing, Dependency, Named Entity Recognition

2. Deep Neural Models :

2.1 Representing Word meaning with word vectors

2.2 From Neural Networks for word prediction to building large language models - LSTMs, Transformers, GPTs, Large Language Models (LLM)

3. NLP applications - machine translation, question answering, summarization

4. Responsible Language Models - dealing with bias and hallucination in LLMs

5. Designing LLM-based applications - Best practices

5.1 fine-tuning, Retrieval Augmented Generation, Chain of reasoning

Learning Outcomes

Students will learn the fundamentals of NLP, difference between analytical and generative tasks. They will also learn how language models are trained, resource requirements

Understanding of how language models can be used to obtain insights from large volumes of text for decision making ; a

The shortcomings of current models will also be discussed along with ethical implications

By end of the course students should be able to design their own applications using standard libraries

Requirements (Reading List and other materials)

Machine Learning, Coding in Python / R

Grading Rubric

Grading will be based on Continuous evaluation over multiple components

1. Coding assignments - coding with NLP libraries in Python / R - grading based on submission of code, results and report along with insights obtained from task - 2 small coding assignments, 1 project of choice - in groups of 2

2. Paper presentation - Application based papers will be shared - students will have to read and make presentations to entire class - in groups of 2

3. Written exams - individual - 1 mid-semester, 1 end-semester

Approximate division of marks -

Coding and result presentation - 40

Reading and Paper presentation - 20

Mid semester exam - 15

End semester examination - 2

Attendance Policy

Expected - 70% attendance