Course Proposal: Advance NLP: Large Language Models
Indian Institute of Technology, Kanpur
Department of Computer Science and Engineering

Course Number: CS781

Course Title: Advance NLP: Large Language Models (LLMs)

Units: 3-0-0-0-9 (L-T-P-C)

Proposed by: Ashutosh Modi

Pre-requisites : Instructor's consent and
Must: Statistical Natural Language Processing (CS779), Proficiency in Linear Algebra, Probability and Statistics, Proficiency in Python Programming
Desirable: Introduction to Machine Learning (CS771) or equivalent course, Deep Reinforcement Learning (CS780), Probabilistic Machine Learning (CS772)

Level of the course: Ph.D., PG, and 3rd, 4th year UG Students (7xx level)

Course Objectives:
In recent times, Large Language Models (LLMs) have revolutionized the field of Natural Language Processing (NLP). However, the application of LLMs has not just remain limited to NLP but has also advanced other areas like Biology, Chemistry, Economics, etc. This calls for in-depth understanding of LLMs. This course will introduce the fundamentals of LLMs and go in-depth into various techniques to develop LLMs, scaling laws. It will cover various LLM architectures. It will teach how to fine-tune LLMs using parameter efficient techniques, how LLMs could be used in conjunction with external knowledge sources such as vector databases. We will have a more mathematical and rigorous approach towards understanding LLMs.

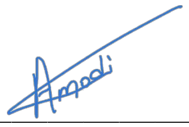
Course Contents (total 40 lectures):

1. Classical Language Modeling (CLM) [3 lectures]: n-grams, smoothing, class-based, brown clustering, etc.
2. Neural Language Modeling (NLM) [4 lectures]: Word Embeddings, Word2Vec, FeedForward Neural LM, Contextualization, Sub-tokenization and Subword information, etc.
3. Transformers for Language Modeling [3 lectures]: Encoder Models, Encoder-Decoder Model, Decoder Models, Pre-trained LMs (PLMs), objective functions for training, etc.
4. Introduction to Large Language Models (LLMs) [3 lectures]: PLMs vs LLMs, LLM families
5. Scaling Laws [3 lectures]: Kaplan's law, Chinchilla Law
6. Training LLMs from Scratch [3 lectures]: Selecting the corpus, cleaning and pre-processing, deciding hyper-parameters using scaling laws, training, etc.

7. Providing Human Feedback [3 lectures]: RLHF, DPO, etc.
8. Emergent Properties in LLMs [4 lectures]: Prompting techniques (zero shot, few shot, etc.), Chain of Thought, Tree of Thought, X of Thought, etc.
9. Parameter Efficient Fine-Tuning (PEFT) [5 lectures]: Transfer Learning, Soft-Prompting, Adaptors, LoRA (and variants).
10. Using LLMs with Vector Databases [4 lectures]: Retireval Augmented Generation and related techniques.
11. Understanding LLM inner workings [5 lectures]: Mechanistic Interpretability

References: Since this is new and emerging area, there are no specific references, this course gleans information from a variety of sources like research papers, tutorials, blogs, etc. Relevant references would be suggested in the lectures.

Date: October 1, 2024



Course Proposer (Ashutosh Modi)

DPGC Convenor

Chairman SPGC

DOAA