COURSE CONTENT

LARGE LANGUAGE MODELS								
VI Semester: CSE								
Course Code	**Category**	**Hours / Week**			**Credits**	**Maximum Marks**		
ACSD47	Core	**L**	**T**	**P**	**C**	**CIA**	**SEE**	**Total**
		3	0	0	3	40	60	100
Contact Classes:48	**Total Tutorials: Nil**	**Total Practical Classes: Nil**				**Total Classes: 48**		
Prerequisites: Artificial Intelligence								

I. COURSE OVERVIEW:

This course delves into the principles, architectures, and applications of Large Language Models (LLMs) with a focus on transformer-based models. Students will understand the inner workings of LLMs, such as GPT, BERT, and their applications in natural language processing (NLP) tasks. The course will explore model training, fine-tuning, deployment strategies, and the ethical concerns related to their use. It is designed to equip students with the necessary skills to develop and work with LLMs in real-world settings.

II. COURSE OBJECTIVES:

The students will try to learn:

I. The key principles and architectures of Large Language Models, particularly transformer- based models.
II. How to train, fine-tune, and deploy LLMs for a variety of NLP tasks.
III. Explore applications of LLMs in areas such as text generation, machine translation, and conversational AI.
IV. Address ethical concerns related to bias, fairness, and transparency in LLMs.
V. Evaluation and optimization of LLMs for real-world deployment, considering computational efficiency and scalability.

III. COURSE OUTCOMES:

At the end of the course, students should be able to:

CO1 Understand the fundamental principles of large language models and the transformer architecture.
CO2 Explain the training process of LLMs, including data pre-processing, model training, and optimization.
CO3 Implement basic LLM applications, such as text generation, sentiment analysis, and question answering.
CO4 Critically evaluate the ethical concerns surrounding the use of LLMs, including biases and fairness.
CO5 Apply fine-tuning techniques to adapt pre-trained models to specific tasks.
CO6 Discuss the future of LLMs and the challenges in scaling and deploying them effectively.

IV. COURSE CONTENT:

MODULE 1: INTRODUCTION TO LANGUAGE MODELS & NLP FUNDAMENTALS (9)
Language models and its Evolution: from n-grams to neural networks, Introduction to NLP tasks: tokenization, POS tagging, Named Entity Recognition (NER). Key challenges in NLP and the role of deep learning.

MODULE II: THE TRANSFORMER ARCHITECTURE & ATTENTION MECHANISM (10)
Introduction to neural networks and deep learning models, The Transformer model: Encoder- decoder architecture, The Attention mechanism and its variations (Self-attention, multi-head attention), Position Encoding and how Transforms handle sequences.

MODULE III: TRAINING LARGE LANGUAGE MODELS (10)
Pre-training vs Fine-tuning large models, Transfer learning in NLP, Techniques for training LLMs (Data collection, tokenization, batching), Optimization methods (Adam, Gradient clipping, learning rate scheduling).

Evaluation and Benchmarks, Parameter-efficient fine-tuning, Domain adaptation, Instruction tuning and reinforcement learning with human feedback (RLHF)

MODULE IV: APPLICATIONS OF LARGE LANGUAGE MODE (9)
Text generation (e.g., GPT, Open AI's GPT-3), Text summarization and translation (e.g., T5, BART), Question answering systems and conversational AI (e.g., BERT), Sentiment analysis and classification tasks.

MODULE V: ETHICAL CONSIDERATIONS AND DEPLOYMENT OF LLMS (10)
Ethical issues in LLMs: Bias, fairness, and misinformation, Ensuring transparency and interpretability in AI models, Environmental impact: The cost of training large models, Model deployment and optimization strategies (cloud, edge deployment), Future challenges in LLM research (scalability, safety).

V. TEXT BOOKS:

1. Jacob Eisenstein, Natural Language Processing, MIT Press, 1st Ed., 2019.
2. Delip Rao & Brian McMahan, Natural Language Processing with PyTorch, O'Reilly, 1st Ed., 2019.
3. Sebastian Raschka et al., Machine Learning with PyTorch and Scikit-Learn, Packt, 2nd Ed., 2022.

VI. REFERENCE BOOKS:

1. Ian Goodfellow et al., Deep Learning, MIT Press, 1st Ed., 2016 (for deep learning foundations relevant to LLMs).
2. Yoav Goldberg, Neural Network Methods for Natural Language Processing, Morgan & Claypool, 1st Ed., 2017.

VII. ELECTRONICS RESOURCES:

1. https://www.coursera.org/learn/generative-ai-with-llms
2. https://web.stanford.edu/class/cs25/
3. https://ocw.mit.edu/

VIII. MATERIALS ONLINE

1. Course outline description
2. Tech-talk topics
3. Assignments
4. Definition and terminology
5. Tutorial question bank
6. Model question paper – I
7. Model question paper – II
8. Lecture notes
9. Early lecture readiness videos (ELRV)
10. Power point presentations