ICT302 Generative AI: Theory and Practice

Level: 3

Credit Units: 5 Credit Units

Language: ENGLISH

Presentation Pattern: EVERY JAN

Synopsis:

This course explores Large Language Models (LLMs) from their foundational concepts to practical applications, while reinforcing ethical considerations and safeguarding measures, to equip students to be socially responsible users and advocates. We begin with the historical evolution of LLMs and introduce the fundamental building blocks of these models, including Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and Transformers. We then explore the practical applications of LLMs, focusing on dialogue-based and vision-based models, while addressing the crucial topics of misuse, adversarial attacks and hallucinations. With an emphasis on data privacy and safety, we provide actionable guidelines for workplace deployment. Lastly, we engage in discussions on policies on generative AI, preparing students to navigate the dynamic landscape of AI ethics. By the end of the course, students will understand the process of integrating LLMs into workplaces and evaluate the safety and ethical considerations of AI systems, ensuring responsible and informed AI use in the modern world.

Topics:

- Introduction to large language models (LLMs) and prompting
- Language modelling and embeddings
- The building blocks of LLMs 1, Recurrent Neural Networks (RNNs).
- The building blocks of LLMs 2, Convolutional Neural Networks (CNNs) and Transformers.
- Dialogue-based LLM applications.
- Vision-based LLM applications.
- Misuse and adversarial attacks.
- Cases of hallucinations, misinformation, and disinformation.
- Data privacy and safety
- Vulnerabilities of neural-based models
- Challenges in deploying safe AI systems.
- Ethical implications of generative AI

Learning Outcome:

- Demonstrate the underlying theory behind language modelling and embeddings, and highlight their role in LLMs.
- Differentiate between RNNs, CNNs, and Transformers and their respective applications.
- Discuss the risks associated with adversarial attacks, misinformation, disinformation, and other forms of misuse in LLMs.
- Examine ethics and policies on generative AI.
- Evaluate the potential of large language models (LLMs) in solving specific problems.
- Analyse guidelines for ensuring data privacy, safety, and ethical considerations in AI system deployment.
- Evaluate and mitigate potential risks to ensure the safety of AI systems in various workplace scenarios.

Assessment Strategies - Regular Semester (Daytime Class):

Components	**Description**	**Weightage Allocation (%)**
Overall Continuous Assessment	QUIZ 1	6
	QUIZ 2	6
	QUIZ 3	6
	TUTOR-MARKED ASSIGNMENT 1	12
Overall Examinable Components	Written Exam	70
Total		**100**

*The information listed is subject to review and change.