

Indian Institute of Technology Kanpur
Department of Computer Science and Engineering
Proposal for a New Course

1. **Course Title:** Foundations of Modern AI
2. **Course No.:** CS778
3. **Credits:** 3-0-0-0 [9].
4. **Duration of Course:** Full Semester
5. **Prerequisites:** CS203 (or equivalent) and CS771 (or equivalent). Instructor's consent will be needed for registering in this course.
6. **Proposing Department:** CSE
7. **Other Departments/IDPs who may be interested in the proposed course:** EE
8. **Other faculty members interested in teaching the proposed course:**
9. **Proposing Instructor:** Sayak Ray Chowdhury, Department of CSE.

10. Objectives: This course examines the theoretical foundations behind the rapid rise of generative AI, particularly large language models (LLMs) such as ChatGPT, Gemini, Llama, and Claude, which have significantly integrated into daily life and are expected to continue reshaping the digital landscape. The course will focus on the fundamental machine learning principles behind these models, with an emphasis on preference-based learning—the backbone of the training pipeline of most LLMs. Moreover, many advanced models leverage RL and planning techniques for improved reasoning and decision-making. Thus, in addition to understanding the core mechanisms of generative AI, the course will explore the growing role of reinforcement learning (RL) in enhancing AI capabilities, providing students with insights into cutting-edge developments in AI training. This research-oriented course is designed as an exploratory deep dive into the field, primarily facilitated through lectures and reading of research papers. Additionally, students will gain hands-on experience working with LLMs, enabling them to develop a practical understanding of these technologies and their broader societal implications.

11. Contents:

S. No.	Broad Title	Topics to be covered	No. of Lectures
1.	Introduction	Basic Terminology of Bandits and RL: states, actions, policy, value function, Q function, Markov decision processes	2
2.	Learning from preference data	Preference Elicitation, Social choice theory, Preference models, Statistical estimation algorithms	6

3.	Bandit algorithms	Contextual bandits, Experimental design, Dueling choice bandits	4
4.	Reinforcement Learning	Policy gradient, Natural policy gradient, Trust region policy optimization, Proximal policy optimization, online vs offline RL	4
5.	Large Language Models	Autoregressive language models, Transformer architecture, Reward function learning, Supervised fine-tuning, Alignment algorithms: RLHF, DPO, IPO, GRPO, In-context learning	8
6.	Special topics	Inference-time alignment/Diffusion models in AI/AI privacy and safety	4

Reference Books: None. References will be provided.

Textbooks / Monographs / Surveys:

1. Machine Learning from Human Preferences. *Sang T. Truong and Sanmi Koyejo*.
<https://ai.stanford.edu/~sttruong/mlhp/>
2. Preference-based Online Learning with Dueling Bandits: A Survey. *Viktor Bengs, Robert Busa-Fekete, Adil El Mesaoudi-Paul, Eyke Hullermeier*.
<https://arxiv.org/abs/1807.11398>

Sayak Ray Chowdhury

Dated: 04.03.2025

Proposer: Prof. Sayak Ray Chowdhury

Dated: _____

DPGC Convener: Prof. Piyush Rai

The course is approved / not approved

SPGC/SUGC Chairperson:

Dated: _____