Schedule for CS6240: Semester 2, 2023/24
Multimedia Analysis

Lecturer: CHUA, Tat-Seng   **Lecture:** Tue, 1000-1200 (COM1-VCRM)

Paper/Idea Presentation & Questioning: 22%

2 Short idea/opinion articles: (2x15) 30%;   **Brave New Idea (BNI) paper:** 48%

Wk	Date	Lecture/Tutorial Topics	Remarks
1.	16 Jan	L1: Overview and Intro (Background, history, featured-based vs. DL-based vs. pre-trained models, LMFM, safety & privacy, etc.)	Outline paper reading, presentation, article and BNI paper requirements
2.	23 Jan	L2: Visual Object Recognition, Detection and Vision-Language Models	Guest Lecturer: Ji Wei
3.	30 Jan	L3: Semantic & Temporal Segmentation and Relation Grounding	Release assignment of paper presentation list (306 Jan)
4.	6 Feb	L4: Cross-modal Alignments and Multimodal Scene Graphs	Discuss details of Article 1 Discuss BNI Papers
5.	13 Feb	L5: Few Shot, Meta and Causal Learning	Submission of Article 1 (16 Feb @1700)
6.	20 Feb	L6: Image/ Video QA, and Reasoning	Discuss details of Article 2 Present Article 1
	27 Feb	**Recess Week**	
7.	5 Mar	L7: Condition-based Diffusion Models for MM Generation	Submission of Article 2 (8 Mar @1700)
8.	12 Mar	L8: Large Multimodal Foundation Model (LMFM)	Meeting on BNI papers Present article 2
9.	19 Mar	L9: NExT-GPT & Multimodal Dialogues	Guest Lecturer: Wu Shengqiong
10.	26 Mar	L10: Responsible AI: Trust, Safety, Privacy & Biased in MM	
11.	2 Apr	L11: Multimodal Event (& Fashion) Detection & Forecasting	Guest Lecturer: Ma Yunshan Submission of BNI Papers (5 Apr @1700)
12.	9 Apr	L12: MM Recommendation Presentation of BNI Papers -1	Guest Lecturer: Wang WJ
13.	16 Apr	L13: Presentation of BNI Papers -2	
14.	23 Apr	**End of Course**	* To return all graded course works & assignments.

A) Presentation & Critiques of Papers and Short Idea Articles
Group or individual **(22%)**

1) a) Each student will be assigned to prepare and **present papers on a sub-topic**. The presentation should be in 25 mins. The presentation should cover: objectives, clear literature review, limitations, design/ implementation and results. Most importantly, it should cover key innovations of the paper, answer many of the "why" questions, and future work.

b) The student will submit a report (<=2 pages), within 2 weeks of the presentation, summarizing the **innovative aspects** of the papers with explanation of "**why it works**", and discussions of **future works**.

2) a) Each student will be assigned as asker for the presentation of a sub-topic. The student will need to pose 2-3 questions for each sub-topic. The questions should have **good depth** and help to **uncover insights of paper**.

B) Short Idea/ Opinion Articles (individual): **(2 x 15) = 30%**

a) Each student is required to submit a short idea/opinion article on an assigned topic. The list of sample topics is listed in Part (b). The writeup should cover the background, issues, analysis and insights. The article should be within 4 pages.

b) Some sample topics could be: (i) Can LFM (Large Foundation Model) use public data for training and content generation: what are the issues and guidelines? (ii) Visual content has always been a supplementary feature to text in semantic analysis. Is this just a problem of maturity of visual analysis tools, or a fundamental problem in multimodal semantic understanding? How LFM should be designed to address this problem? (iii) Robust and trust are the key problems to LFM. Is this a fundamental unsolved problem? What can be done to address this problem? (iv) …

c) The deadlines for the submission of 2 articles are:
  - Article 1: by 16 Feb @1700
  - Article 2: by 8 Mar @1700

C) Brave New Idea (BNI) Papers (individual): **48%**

TOPIC: A BNI on a Multimedia Topic with sufficient Validation

AIM:

To propose works that contain original ideas and research vision that draw attention to novel directions in multimedia research. I am particularly looking for papers offering: (i) novel, exploratory solutions with sufficient evidence of proof-of-concept; (ii) visions describing a new or open problem in multimedia research; and (iii) a novel perspective on existing multimedia research.

This is similar to the new BNI track in ACM Multimedia 2024:
  (https://2024.acmmm.org/brave-new-ideas-proposals)

GUIDELINES:
  - The BNI papers are expected to have a high component of novelty. They can also address an understudied, open problem in multimedia, which may receive less attention in the full paper track. However, BNI papers should still support their ideas with sufficient scientific argumentation, and where

appropriate experimentation or proof. The papers also require high clarity in presentation.

- The paper should contain ideas not previously submitted nor published.
- The paper should be within 5 pages, excluding references, in ACM 2-column format.

MARKING CRITERIA:

The papers will be evaluated using the following criteria:

- Novelty: Are people already studying this problem in the multimedia community? In other communities?
- Conceptual leap: How does it change our way of thinking about a problem, a method, or its execution? How much will research in this area advance our state of knowledge?
- Depth of impact: How many supporting sub-problems requiring foundational research might arise from the observations in the submission?
- Breadth of impact: What stakeholders inside or outside of the multimedia community are impacted? What new applications might this idea trigger?

SUBMISSION DEADLINES:

The key deadlines are:

- Meeting to clarify topic: by 20 Feb (Tue) during or after the lecture.
- Final paper is due on: 5 Apr (Fri) @1700)
- Presentation to class (5 mins each): 9 Apr (1100-1200) & 16 Apr (1000-1200)