COMP7015 Artificial Intelligence – Group Project Instructions

1. Overall Requirements

1. **Groups:** Form a group of **1 to 5** students. Forming groups with students from other sections is allowed. (At least 3 members in a team are recommended)
2. **Milestones:**
   - **Group Registration & Topic Selection:** Due by 11:59 pm, 24th October 2025. Please register your group members and chosen topic via the following link: https://hkbuchtl.qualtrics.com/jfe/form/SV_eS9K0CSvz9A9JvU.
   - **Final Submission:** Due by 11:59 pm, 21st November 2025. This includes your source code and project report. One submission for each team will be enough.
   - **In-person Presentation:** Scheduled for 22nd and 23rd November 2025. A detailed schedule will be announced after the group registration deadline.
3. **Final Submission Package:**
   - **Project Report:** A PDF document of at most five A4 pages (single column). It should describe your project's motivation, methods, results, and a discussion. A mandatory section must detail the contribution of each group member.
   - **Source Codes:** A single .zip file. Your code will be evaluated in the FSC 8/F lab environment, so ensure it runs smoothly there. Acknowledge all major third-party libraries (e.g., PyTorch, TensorFlow, Hugging Face Transformers) in your report.
4. **Presentation:** Each group will have approximately 8 minutes for their presentation, followed by a Q&A session. Every member must present their part of the work. The use of visualizations (figures, graphs, demos) is highly encouraged to clearly convey your project's story.
5. **Academic Integrity:**
   - We have a zero-tolerance policy for plagiarism. All submissions will be checked by anti-plagiarism software. Copying code from online sources or generative AI tools without proper citation and understanding is strictly prohibited.
   - Submission of work from other courses or previous projects is considered self-plagiarism and is not allowed.

2. Topics

You may choose one of the following three topics. For Topics 1 and 2, you are expected to build on the suggested tasks. For Topic 3, you have the freedom to define your own project, keeping in mind that the project's scope and difficulty will be evaluated.

Topic 1: Human Action Recognition (Computer Vision)

This project focuses on building and evaluating deep learning models for human action recognition using the **HMDB51 dataset**, a standard benchmark in video understanding.

HMDB51 contains short video clips spanning 51 categories of everyday actions (e.g., running, walking, clapping).

Dataset:
1. This topic is based on the HMDB51 Dataset, available for public download at http://serre-lab.clps.brown.edu/resource/hmdb-a-large-human-motion-database/
2. You may find the HMDB51 dataset API from torchvision helpful. For more details, see https://docs.pytorch.org/vision/master/generated/torchvision.datasets.HMDB51.html.
3. To keep the project tractable, you may select at least three categories of your choice and frame the task as a multi-class classification problem. You are encouraged to explore using more classes, but we understand the limit on computing resources, so you will not be penalized for using fewer classes than other groups as long as you use at least three.

Minimum Requirements (you secure 50% scores if the following are done correctly):
1. Data split: Split the data into training, validation, and test sets.
2. Frame extraction & preprocessing: From each video, extract 3–4 frames, combine them into one image, and treat the problem as a static image classification task.
3. Static 2D CNN models: Train at least one CNN from scratch (e.g., a small custom CNN) and one model leveraging transfer learning (e.g., ResNet18/34) for multi-class classification.
4. Basic experimentation: Explore hyperparameters such as learning rate, batch size, and regularization (dropout, data augmentation).
5. Evaluation: Select proper evaluation metrics and analyze the results you obtained.

If you aim to score higher than 50%, pick one direction and explore in-depth:
1. Explore advanced modeling techniques for video data, for example, 3D CNNs or combining other temporal models with CNNs. Explore methods (e.g., regularization, normalization, etc.) that could optimize your model performance.
2. Focus on the same human action recognition task, but use a larger dataset (e.g., UCF101), find a suitable pre-trained model, and fine-tune it. This option requires access to more powerful GPU cards.
3. Any other interesting and creative ideas you might have for human action recognition (you are free to explore other datasets or models). You should explain how it connects to the course content.

Topic 2: Sentiment Analysis (Natural Language Processing)

This project focuses on building and evaluating models for sentiment analysis, a core task in Natural Language Processing. The goal is to classify movie reviews as either positive or negative using the **Large Movie Review Dataset (IMDb)**, a standard benchmark for binary sentiment classification.

Dataset:
1. This topic is based on the **Large Movie Review Dataset (IMDb)**. It contains 25,000 movie reviews for training and 25,000 for testing, which are labeled as either positive or negative. The dataset is available for public download at https://ai.stanford.edu/~amaas/data/sentiment/.

2. The primary task is a binary classification problem (positive/negative). You are encouraged to explore more granular classifications (e.g., predicting 1-10 star ratings) for a more challenging project.

Minimum Requirements (you secure 50% scores if the following are done correctly):

1. Data Split: Split the provided training data into your own training, validation, and test sets to properly evaluate your models.
2. Text Preprocessing & Vectorization: Clean the raw text data (e.g., remove HTML tags, convert to lowercase). Then, perform tokenization, build a vocabulary from your training data, and convert your text sequences into integer sequences. Ensure you handle sequences of varying lengths by implementing padding.
3. Deep Learning Model: Implement and train one recurrent neural network (e.g., an LSTM or GRU) for binary classification. Your implementation must include and compare the following two approaches for the embedding layer: A randomly initialized embedding layer that is trained from scratch, along with the rest of your model, and an embedding layer initialized with pre-trained word embeddings (e.g., GloVe or Word2Vec), which can be either frozen or fine-tuned during training.
4. Basic Experimentation: Explore and report on the effect of key hyperparameters for your deep learning model, such as learning rate, batch size, dropout rate, and the number of recurrent units.
5. Evaluation: Select proper evaluation metrics for classification (e.g., accuracy, precision, recall, F1-score). Analyze and compare the results you obtained from using the trainable embedding versus the pre-trained embedding.

If you aim to score higher than 50%, pick one direction and explore in-depth:

1. Explore more advanced modeling techniques by fine-tuning a large, pre-trained language model (e.g., a variant of BERT or another suitable model). You should thoroughly analyze the trade-offs in terms of performance improvement, computational cost, and implementation complexity when compared to your recurrent model.
2. Focus on a more complex or nuanced NLP task using the same or a different dataset. This could involve performing fine-grained sentiment analysis by predicting a star rating, tackling aspect-based sentiment analysis, or attempting to detect more subtle linguistic features like sarcasm or irony.
3. Propose and implement any other interesting and creative ideas relevant to sentiment analysis or text classification. You could explore different model architectures, investigate methods for model interpretability to understand *why* your model makes certain predictions, or apply your models to a completely different domain, such as social media or product reviews. You must clearly explain how your chosen idea connects to the course content.

Topic 3: Open Topic

If you are interested in another AI problem that aligns with the course content, you may propose your own project. This is an opportunity to explore areas like generative AI, multimodal learning, foundation models, or other advanced deep learning applications.

- You are free to choose any relevant dataset and AI/deep learning methods.
- Your project should demonstrate a clear understanding of the principles and techniques taught in this course.
- Remember that the chosen topic's difficulty, scope, and creativity will be key factors in your evaluation.

3. Evaluation Criteria

Your project will be evaluated based on a holistic assessment of your work across several dimensions:

- **Project Completeness & Model Performance:** How thoroughly the project tasks were completed and the effectiveness of your final model(s) on the chosen task.
- **Creativity & Difficulty:** The novelty of your approach, the technical challenge of the problem, and the sophistication of the methods used.
- **Code Quality:** The readability, organization, and documentation of your source code. Clean and well-structured code is expected.
- **Storytelling (Report & Presentation):** The clarity and depth of your project report and presentation. This includes how well you explain your motivation, describe your methodology, analyze your results, and present your conclusions.