HONG KONG BAPTIST UNIVERSITY

Department of Mathematics

MATH7370 & 4665 Research Methods

Homework 2: AI Cross-disciplinary Terminology Translation

Due: 4 March. 2026

Selected Project Topics

UG1: Recommender System is a data-driven problem where the goal is to guess what each user may like, often by learning patterns from a table of past ratings and completing missing entries to suggest similar movies or products.

UG2: Bayesian Optimization is a sequential optimization problem where each experiment is slow or costly. The method learns a smooth approximation of an unknown function and uses it to decide the next test point that may reduce the function value most effectively.

UG3: Dynamic Pricing is a simple machine-learning problem where the model learns how sales respond to price, often by fitting or updating a small regression model as new data arrive, to decide which price may raise overall revenue.

UG4: Proximal and ADMM Methods address structured optimization problems by splitting them into smaller subproblems that can be solved alternately or in parallel. Each update combines linear-algebra steps with proximity or penalty terms to enforce constraints. These methods illustrate how modern convex optimization ideas connect to large-scale AI training frameworks.

RPg: Text-to-Image Generation is a modern AI framework that unites modeling, language, and optimization ideas. A large neural network first learns to represent the distribution of natural images through an encoder–decoder structure with controlled noise (the diffusion process). Training couples this visual model with a text encoder so that short text prompts, converted into token sequences, can guide the denoising steps toward consistent images. The complete system can be viewed as iterative optimization in a high-dimensional cross-modal space—fitting both spatial (image) and temporal (noise-schedule) dimensions together.

RPg students must first **work as a group** to understand and present the **big picture** of how text-to-image generation operates, clearly enough that undergraduate peers can follow. After a shared overview is achieved, each member may focus on one component. Given sets of images and text $\{I, T\}$, for example,

- *CLIP and Vision Transformer* explore how large neural networks represent visual and textual information in a shared space. A Vision Transformer encodes images as sequences of visual tokens ($I \to E_{\text{img}}(I)$), while a language Transformer encodes text as word tokens ($T \to E_{\text{text}}(T)$).
- *Representation learning:* the encoder–decoder and latent-space structure $(E_{\text{vae}}, D_{\text{vae}})$ that model the distribution of image data.

- *Diffusion or generative dynamics:* forward and reverse processes ($x_t \leftrightarrow x_{t-1}$) formulated by stochastic differential equations (inverse Brownian motion), or their ODE/PDE analogues.
- *Training formulation:* given the encoders ($E_{\text{text}}, E_{\text{img}}$) and the latent representation ($E_{\text{vae}}, D_{\text{vae}}$), one can define an optimization problem that couples noisy image recovery with text-conditioned guidance in the shared latent space.
- If you want a brainteaser, consider how to extend this model to allow *image-guided text-to-image generation*.

General Expectations: Each student will select **one topic only** from the list provided and complete the report using LaTeX. The task is to connect real understanding with its mathematical formulation and computation. The goal is clarity of thought, not use of terminology (e.g., narrow set of buzzwords $\rightarrow$ genuine understanding).

1. **Real-world Motivation:** Explain in plain (layman) words what situation the topic represents and what question it answers. Describe briefly what the data represent and what outcome is desired (e.g., optimizer, surrogate, or decision).
2. **Mathematical and Statistical View:** Express the idea in analytical form. Identify the main variables and model the relation among them using equations or expectations. Explain key concepts using multiple terms where suitable (e.g., loss = objective = cost).
3. **From Model to Computation:** Translate the formulation into a linear algebra and/or algorithmic viewpoint suitable for coding. Comment on realistic computation for large data (e.g., iterative, *stochastic* approximate, or online updates).
4. **Interpretation and Reflection:** Summarize what the model or calculation reveals about the original problem. Any concise and honest insight is acceptable (e.g., Ridge regression = Tikhonov regularization = least-squares fit with a penalty).

Each submission should show how an applied situation becomes a model, how the model becomes a computation, and include *Mickey mouse examples* for proof of concept whenever possible.