SEEK + UNSW COMP6713 -Group Project 2025 Term 1

SEEK is very excited to be sponsoring this year's course project! The project theme is **Generative AI for Job Ad Enrichment**.
- **Task:** Your task will be to use large, pre-trained, generative language models to identify and extract key pieces of relevant information from job advertisements.
- **Data**: SEEK will provide a dataset of labelled and unlabelled data from job ads recently posted on SEEK. Participants will be required to sign an NDA prior to receiving the dataset.
- **Prizes:** SEEK will be sponsoring two cash prizes: $500 to the team with the best technical solution, and $500 to the team with the best business solution.

Marking for the group project will be done only by personnel at UNSW. Therefore, ensure that you follow the scope specifications and evaluation rubric as provided by the course team.

If you opt for this industry project,
- All team members will send the signed NDA to aditya.joshi@unsw.edu.au before the data can be sent. The dataset itself will be sent ONLY at the start of the project AFTER SEEK has validated the NDAs. Delays in this case are beyond the control of the course team.
- Project teams that win the prizes agree to send the code to SEEK.

Background

At the heart of SEEK's business is the capability to match job seekers with relevant job openings. Part of this capability is our ability to extract key pieces of relevant information from job advertisements, so that candidates can easily filter and identify jobs that meet their preferences.

Generative pre-trained LLMs are an attractive solution to this problem due to their strengths in understanding natural language text and reasoning. Therefore, SEEK would like to gain a deeper understanding of the pros and cons of different applications of generative AI to this problem of information extraction from job ads.

Research Questions

As SEEK develops Generative AI solutions to extract key pieces of information from job ads, we are interested in answering questions such as:
1. **What is the Cost vs Speed vs Accuracy trade-off?** Different solutions may be better than others on one or two of these, but no solution will be the best on all

three dimensions.
• To answer this question, you should evaluate multiple solutions and frame this trade-off as a decision to be made by business leaders about which solution should go into production. Your evaluation should include concrete metrics for all three dimensions: Cost: Track and report API call costs, token usage, and compute resources required; Speed: Measure average processing time per job ad and maximum throughput; Accuracy: Compare performance against the labelled test set using standard metrics.

2. **Should we use open-weight or proprietary models?** Like most businesses, SEEK would prefer to avoid "vendor lock-in", where we become dependent on a single provider (for example: OpenAI). Open-weight models provide a greater degree of control and flexibility, but they may not perform as well as proprietary models, and self-hosting your own models (as opposed to calling them via an external API) adds complexity to the build.
• To answer this question, you should compare the performance of some open-source and some proprietary models, and discuss the trade-offs that they bring.

3. **Is fine-tuning worth it?** Although pre-trained foundation models are very effectively used as-is, for some use cases there are still significant gains which can come from fine-tuning the LLM on your own data. However, fine-tuning can be costly, and using a fine-tuned model may mean that you miss out on the opportunities to upgrade to the latest new model releases.
• To answer this question, you should consider the uplift that fine-tuning provides, by comparing a fine-tuned LLM against its base (non-fine-tuned) counterpart. You might also consider how fine-tuned models might provide cost savings by comparing a smaller, fine-tuned model against a larger, non-fine-tuned model.

4. **Is our solution unbiased, fair, and transparent?** SEEK has a diverse range of hirers and candidates from many different backgrounds, with different needs, speaking many different languages. We would like to ensure that our AI solutions provide a benefit to all users in a fair and transparent way.
• To answer this question, you should consider multiple solutions and identify their strengths and weaknesses with respect to bias, fairness, and transparency.

The data

You will be provided with a dataset of job ads that have recently been posted on SEEK. This dataset consists of three parts:
1. Development set (unlabelled)
2. Development set (labelled)
3. Test set (labelled)

You may use the two development sets to aid you in developing your solution. The test set must *only* be used to evaluate your final submission.

The labelled data sets illustrate the information extraction tasks. There are three separate tasks:
1. **Salary** – identify the currency, amount, and frequency of salary
2. **Work arrangements** – identify if the position is Remote, On-Site, or Hybrid

3. **Seniority** – identify the seniority level of the position

Your task

Your task is to build an AI solution that accepts job ads as input and outputs one or more of the structured outputs that we have defined. Your solution must follow the following guidelines:

1. You *must* provide at least one solution based on generative (decoder) LLM.
2. You *must* evaluate your solution(s) on the provided test set and report the test metrics for your solution(s).
3. You *must* address at least one of the provided research questions.

Note that the COMP6713 team requires you to submit a simple, pre-neural solution.

How winners will be determined

SEEK will assess the solution for the cash prize, while marking will be done solely by the course team at UNSW. Your solution will be assessed similarly to how data science initiatives in industry settings are assessed, by a combination of technical and business considerations. The goal of this exercise is *not* to simply try to achieve the highest evaluation metrics possible. We are looking for a well-reasoned and well-communicated analyses of the research questions that have been posed, based on a foundation of data collected from your technical implementation.

We will assess on these two dimensions, equally weighted towards the final score:

1. The quality of the technical implementation(s) that you have included in your solution. Have you followed the current best practices to achieve a good result using the data provided and the current technology available?
2. The quality of your analysis and discussion in response to the questions posed above in this document. Is your analysis based on a solid foundation of good data and metrics? Is the discussion clearly communicated? Have you made a good consideration of the business needs and how they connect to the technical implementation?

Submission and Evaluation

Teams must follow the submission and evaluation requirements as communicated by the course team. **Marking will be done completely at UNSW.** Group projects will be uploaded to Moodle.

The prize

SEEK will award two $500 prizes:

- $500 to the **best technical solution**. This prize will be given to the team whose solution we believe exemplifies the best practices in modern NLP, including the solution design, implementation, evaluation and other technical aspects.
- $500 to the **best business solution**. This prize will be given to the team who we believe did the best job of analysing and communicating the trade-offs of different technical solutions from the perspective of a business leader at SEEK.