National College of Ireland

Configuration Manual

MSc Research Project
MSc of Artificial Intelligence

Muhammad Hassan Shakeel
Student ID: 23215992

School of Computing
National College of Ireland

Supervisor: Dr. Anh Duong Trinh

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Muhammad Hassan Shakeel
Student ID:	23215992
Programme:	MSc of Artificial Intelligence
Year:	2024
Module:	MSc Research Project
Supervisor:	Dr. Anh Duong Trinh
Submission Due Date:	18/12/2024
Project Title:	Configuration Manual
Word Count:	484
Page Count:	4

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Muhammad Hassan Shakeel
Date:	24th January 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☑
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☑
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☑

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Muhammad Hassan Shakeel
23215992

1 System Requirements

- Use Google Colab pro for fine-tuning and evaluation of LLM models
- Select L4 GPU with "High RAM" toggle button enabled
- The selected environment should have 53 GB of RAM, 22.5 GB of GPU RAM and 235.7 GB of total disk space.

2 Data Acquisition

Although data is available in "MedQuAD-master" directory inside the zipped directory. But dataset is publicly available to download as well at https://github.com/abachaa/MedQuAD (Accessed at: 17/12/2024) (Abacha; 2024).

3 Project Structure

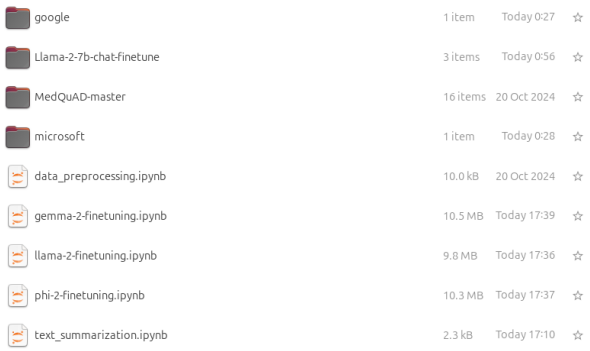

Figure 1: Project Structure

The project structure shown in Fig. 1 is explained below:

- Directories "google", "Llama-2-7b-chat-finetune" and "microsoft" contains fine-tuned model parameter configurations that we can load in our code files and use later.
- "MedQuAD-master" directory contains the raw dataset we downloaded from GitHub link shared earlier.

- The dataset is in xml format, therefore the file data-preprocessing.ipynb converts data into CSV format and it also filters out data rows that contain empty string in the answer column.
- Afterwards, text_summarization.ipynb file is executed to generate extractive summary of the answers provided in the dataset.
- By this point data preprocessing is done, so gemma-2-finetuning.ipynb, llama-2-finetuning.ipynb and phi-2-finetuning.ipynb files can be run in any order as well as it can be run in parallel as well.

4 Code Execution Guidelines

4.1 Preprocessing guidelines

- Make sure to set current directory as root directory of the project and avoid changing any directory/file names.
- Execute data_preprocessing.ipynb file cell-by-cell. When all the cells are executed in sequential order, a new file MedQuAD.csv will be saved in root directory of the project.
- Execute text_summarization.ipynb file in sequential order to generate summary of the answers. Summaries will be appended in the CSV file that was generated by data_preprocessing.ipynb in the last step.

4.2 Fine-tuning guidelines

- For fine-tuning of the LLM model, execute each of the three LLM model ipynb files until the cell shown in fig. 2. Execute the cell shown in the fig. 2 as well if you want to save the fine-tuned model configurations.
- For evaluation of LLM models, continue execution until the cell shown in fig. 3. Cell shown in the figure calculates BLEU score.

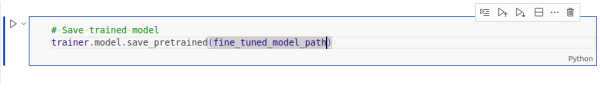

Figure 2: Cell - To save the fine-tuned model

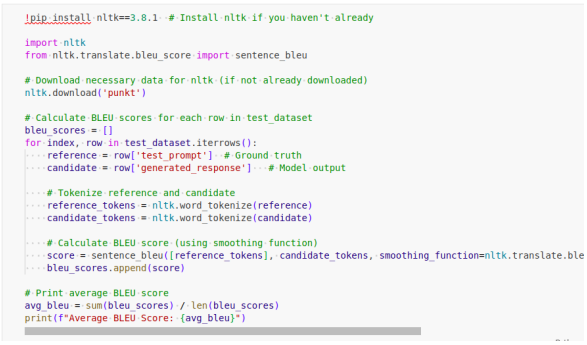

Figure 3: Cell - For evaluation of the LLM

4.3 Load Fine-tuned LLM

- If you have you not executed any cell of the LLM fine-tuning code files. Make sure to execute until and including the cell shown in fig. 4.
- Then execute the cell shown in the fig. 5 to load fine-tuned LLM.
- To evaluate the loaded model, execute three cells shown in the fig 6. Also, execute the cell shown in the fig. 3

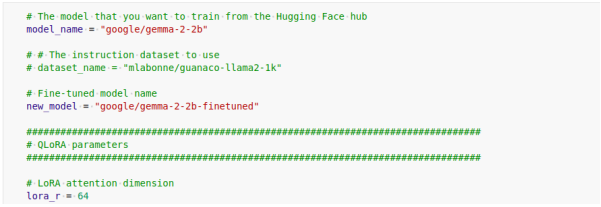

Figure 4: Prerequisite of loading saved model

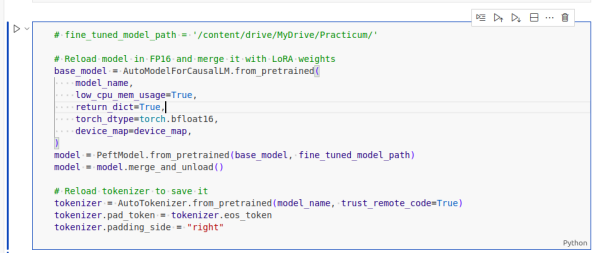

Figure 5: Load Fine-tuned LLM

Figure 6: Evaluation steps of Fine-tuned model

For fine-tuning and loading fine-tuned LLM, same steps will be applied for all three LLM files.

References

Abacha, A. B. (2024). Medquad: A medical question answering dataset, `https://github.com/abachaa/MedQuAD`. A comprehensive dataset for medical question answering.