**Brief:**

In this assignment, you will get a hands on experience working with PrimeQA toolkit/models to solve an open domain question answering task over a real world dataset: covid-qa. The target domain of this dataset is Covid19 related documents/ publications, over which PrimeQA/models can answer natural language questions. We have designed experiments to focus on domain adaptation aspect of question answering.

**Pre-requisite**: To complete the assignment, we assume you have the following things ready and available with you. Please contact the course TA, if you need help with any of them.

1. Google account and access to Google Collab (available for free).
2. A git account which can be used to access PrimeQA git repository here: https://github.com/primeqa/primeqa (it is a public repository, so an active git account is enough for access)
3. A document index created on covid qa dataset available at a known location < ask TA>
4. (a) Train and (b) test split of covid-qa dataset available at a known location: <ask TA>

5. covid-qa test queries in specific .tsv format as needed by run\_ir –queries flag at a known location <ask TA>

6. covid-qa document data stored in specific .tsv format as needed by run\_ir –collections flag.

**Useful PrimeQA references:**

[R1] How to run a Colbert retriever to search documents given a list of queries and a document index : <https://github.com/primeqa/primeqa/blob/main/primeqa/ir/README.md#dense-index-retrieval-with-colbert-engine>

[R2] How to use a PrimeQA/Reader for training/evaluation on a dataset: <https://github.com/primeqa/primeqa/blob/main/primeqa/mrc/README.md>

[R3] How to train the PrimeQA/Reader on a custom training data:

https://github.com/primeqa/primeqa/blob/main/examples/custom\_mrc/README.md#fine-tuning-model

Assignment 1 with PrimeQA

We’ve supplied you with all the necessary tools and procedures to get started. Here’re the steps to complete the experiments (and assignment !)

* Step1: Get PrimeQA setup on a Google colab.
	+ - Create a new notebook in google collab
		- Mount your google drive to be used for load/store from your google collab notebook.
		- Git clone the PrimeQA repo

 (More help? Check this: [https://medium.com/@ashwindesilva/how-to-use-google-colaboratory-to-clone-a-github-repository-e07cf8d3d22b](https://medium.com/%40ashwindesilva/how-to-use-google-colaboratory-to-clone-a-github-repository-e07cf8d3d22b))

* Step 2: Start your own new worksheet.
* Step3: place the index from the location as mentioned in Pre-req (3) and place it in your desired location which will be used in Step6.

 [Hint: run\_ir script uses a collection of flag values mentioned via –root,

 --experiment and –index to locate the folder for loading/storing the index.

 More precisely, the index is loaded/stored from/in the exact location=

 $root+$experiment+indexes+$index. For example, with the following flag values

 --root A –experiment B –index C , run\_ir script will look to load/store the index in

 /A/B/indexex/C. Place the shared index files accordingly]

* Step4: Check the PrimeQA’s Squad reader model available in HF model hub <https://huggingface.co/PrimeQA/squad-v1-roberta-large> [This step is a no-op. The model can be used simply by referring to its name in Step6]
* Step5: Download the covid QA dataset with our train and test split from the location as mentioned in Pre-req(4) and place it in a desired location which will be used I Step6.
* Step6: Run the end2end pipeline without fine-tuning (yet!):

 for every question in the test set i.e. Pre-req 4(b), do:

* 1. Run a ColBERT retriever with the provided index (pre-req 3) and search for relevant docs for the queries in Pre-req 4(b).

[Hint: Take a look at [R1] to see how to search for documents given

a list of queries and an index.]

* 1. Run reader mentioned in step 4 on the top document returned in step 6a. [Hint : Take a look at [R2], specifically to check how to run a reader model only for evaluation. ]
	2. Report F1 scores
* Step7: Now, you’ll fine-tune the QA system on the training data we’ve provided in Pre-req4(a). Note, you’ll fine-tune ONLY the reader for this assignment.

 [ Hint: Take a look at [R3] to check how you can fine-tune a reader with custom

 Data. ]

* Step8: Repeat step 6 to measure the performance of finetuned reader model.
* Submit the notebook to the class TA.
* You’re done! You’ve successfully created a QA system on a new target domain with PrimeQA. For fun, feel free to try out the notebooks, with some questions that come to your mind.