The structure of the zip file:

      Bert
      checkpoint
      CovidSciBert
      CovidTwitterBert
      Data
      LSTM
      Roberta
      cs4248-ensemble.ipynb
      dataset_analysis.ipynb
      instructions.pdf
      result_all_submissions.xlsx

- The checkpoint folder restores model weights.
- The data folder restores the dataset
- In the folder LSTM, running cs4248_LSTM.ipynb will train and evaluate the LSTM model. In addition, the model weights and the metrics will be saved in the current folder.
- In each fold named as a pretrained model, there is one .ipynb notebook. Running the notebook will train and evaluate the model. In addition, the model weights will be saved in **"../checkpoint"**. The training log (training loss and accuracy on validation dataset) will be saved in the current folder.
- The cs4248-ensemble.ipynb defines the ensemble model, running it will give the final F1-score of 0.9860. But make sure the model weights are stored in checkpoint folder.
- The dataset_analysis.ipynb makes analyses of the dataset.
- result_all_submissions.xlsx is the leaderboard of the competition.
- Link for the ensemble model weights: