

# MSMHNSM - Code Execution Instructions

Rohit Viswakarma, Pidishetti

## Directory setup:

Make sure that your current working directory (CWD) on google collab has the following files before executing the source code as these are very important files in order to execute the source code,

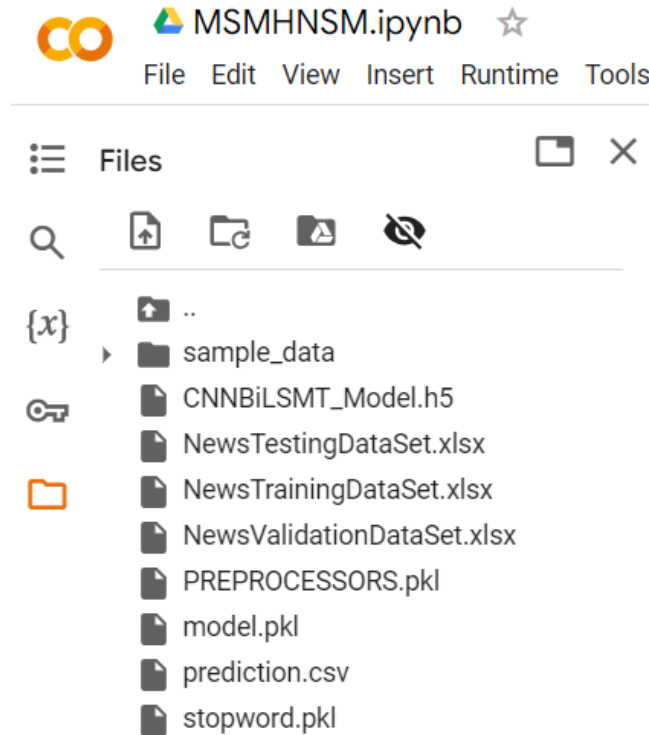


Fig: Illustrating the list of files that are essential for code execution.

- **CNNBiLSMT\_Model.h5:** This file is not mandatory as the code functionality is written in such a way that it first searches for this file when executing the cell related to model training, if not found then it will generate another deep learning model, if found In order to save time and computation, this pre-trained model will be loaded into the code for data classification.
- **\*.xlsx:** These files are essential for CNN+BiLSTM model training, testing and validation.
- **PREPROCESSORS.pkl:** This is a very important binary file that consists of all the class, file loaders that are necessary to carry out text based classification task related to this study.
- **model.pkl:** This file is generated only when the model is trained, and contains all the information related to model's accuracy, precision and other metrics related to its performance as well. Make sure you include this file when using a pre-trained model instead of training the model.
- **prediction.csv:** The file with all the predicted values i.e., labels. (Need not to be included) as this will be automatically generated while testing the performance of the model.
- **stopword.pkl:** This file is a very essential file as it contains all the information regarding the stopwords i.e., words that are used very frequently in the english language, the data is stored in binary context.