Experiment 2

*Start Date: Nov. 2, 2022
*Report Due Date: Nov. 16, 2022

Notes:

- Each experiment must be done individually. You can search material through Internet but remember to mark it.

- Write an experiment report to describe and analyze the experimental observations.

- You should pack your files including **source code, report, readme file, and other related files** in one .zip/.rar/.7z file.

- Please submit on the Web Learning platform. Do NOT Email or Wechat your report to the instructor or TAs.

- No late report is accepted. No exceptions.

- You can write the report using English or Chinese.

Task:

Learn about how to detect the fake news using the BERT model.

Goal:

In this experiment, you are expected to clean the data, analyze the data and design your model.

Data:

Please check the following train and test dataset. You can download it via this link.

Experiment Step:

- Do some EDA (Explore Data Analyze) work to analyze the data. There is an example.

- Build your BERT based classification model using Transformers package from Hugging Face. You can follow the demo here but DO NOT copy it directly.

- Try some data augment tricks such as replacing some words with similar words. You can refer to these paper for more method: AEDA and EDA.

- Try to modify the structure of your model based on the observation of your EDA discussion. Analyze the result and get your conclusion.

- (Optional) You are encouraged to choose a simple model and give a comparative experiment.