Question
1. Do we need to use single-spaced or double-spaced?
2. Do we need to include the graph we made in this pdf?

<u>Introduction:</u>

       With the development of technology, there are many sources of information on the Internet. People can easily obtain information from their smartphones and personal computers, which leads us to a convenient society. However, not all information is correct. Fake news is widespread in the news industry and has become a global issue. A recent example is the news of the pandemic caused by COVID-19. According to Amy Watson, who is a research expert covering the media, 80 percent of consumers in the United States have seen fake news about the pandemic and have been confused. Fake news can harm you in terms of three perspectives. First, fake news is harmful to students. The professor requires you to write your report using reliable resources, but if you use fake news as a reference, you may get a lower grade. Next, there are many fake news stories related to medical care and illness. This incorrect information will directly cause serious and life-threatening damage to you. Finally, fake news makes it difficult for people to know the truth. It is not easy to distinguish between fake news and real news. Fake news helps people have biases and biases make people vulnerable to misinformation. Fake news has a major impact on their relationships and their daily lives. To solve these problems, we analyze the data that summarizes real and fake news with titles and content, and finally, we create a fake news detection algorism to help fight the spread of misinformation before the 2020 election.

Files:

Script.ipynb

This file downloads the dataset programmatically and loads the dataset into a data frame through pandas.

Data_Analysis.ipynbk

This file makes some graphs and generates analysis based on the graphs.

Machine Learning Classifier and prediction model. ipynb, machine learning 2. ipynb

This file uses a classifier model and includes two machine learning.

README.md

This file includes an explanation of how to reproduce the result of analysis and machine learning.

Analysis:

Using a groupby operation, we counted the number of fake and real news and made a bar graph which shows that the total number is relatively the same.
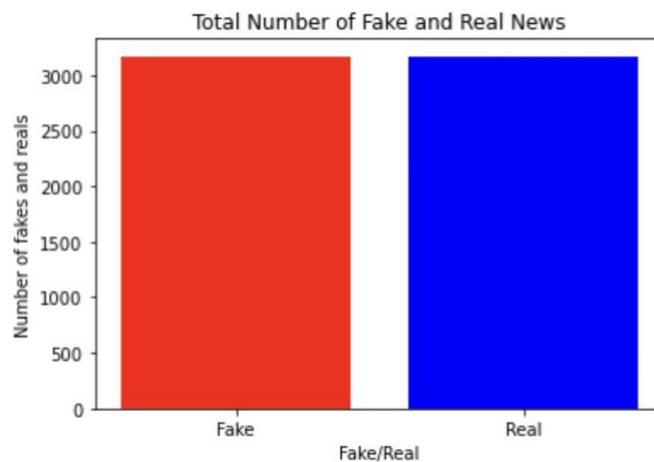

Figure 1: *Total Number of Fake and Real News*

Machine Learning:

Using TfidfVectorizer, we divided the data into a word and find meaningful words. Then, using Passive Aggressive Classifier, we classified the data and calculate the accuracy of the predicted value.