1. FAKE AND REAL NEWS
all_TI-CNN_data.csv: https://arxiv.org/abs/1806.00749 find in the footnote in this research paper

The dataset in this paper contains 20,015 news, i.e. **11,941 fake news** and **8,074 real news**. It is available on google drive1. If it is not available, you can also get the data set from one drive 2. For fake news, it contains text and metadata scraped from more than 240 websites by the Megan Risdal on Kaggle3. The real news is crawled from the well known authoritative news websites, i.e., the New York Times, Washington Post, etc. The dataset contains multiple information, such as the title, text, image, author and website. To reveal the intrinsic differences between real and fake news, we solely use the title, text and image information.

2. FAKE NEWS

Getting_Real_about_Fake_News(big data)： https://www.kaggle.com/datasets/mrisdal/fake-news

The dataset contains text and metadata from 244 websites and represents 12,999 posts in total from the past 30 days. The data was pulled using the webhose.io API; because it's coming from their crawler, not all websites identified by the BS Detector are present in this dataset. Each website was labeled according to the BS Detector as documented here. Data sources that were missing a label were simply assigned a label of "bs". There are (ostensibly) no genuine, reliable, or trustworthy news sources represented in this dataset (so far), so don't trust anything you read.

3. REAL NEWS
True.csv Fake.csv:
https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset?select=True.csv

20826 rows of true news
17903 rows of fake news