

FakeNewsNetPreprocessing

July 14, 2024

1 FakeNewsNetPreprocessing

This notebook preprocesses the FakeNewsNet dataset created by Shu et al. [7]. The dataset is hosted in a GitHub repository[8] and contains political and entertainment news along with social context and spatiotemporal data. The dataset includes the GossipCop and PolitiFact subsets, which will be explored, cleaned, processed, and saved to Google Drive for use in subsequent machine learning tasks in the Classifier notebook.

1.1 Import necessary libraries

```
[1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import nltk
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
from collections import Counter
import re
from sklearn.utils import shuffle

nltk.download('stopwords')
```

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.
```

```
[1]: True
```

1.2 Download the datasets

```
[2]: # Downloading datasets from the GitHub repository
!wget "https://raw.githubusercontent.com/KaiDMML/FakeNewsNet/master/dataset/
↳gossipcop_fake.csv" -O gossipcop_fake.csv
!wget "https://raw.githubusercontent.com/KaiDMML/FakeNewsNet/master/dataset/
↳gossipcop_real.csv" -O gossipcop_real.csv
!wget "https://raw.githubusercontent.com/KaiDMML/FakeNewsNet/master/dataset/
↳politifact_fake.csv" -O politifact_fake.csv
!wget "https://raw.githubusercontent.com/KaiDMML/FakeNewsNet/master/dataset/
↳politifact_real.csv" -O politifact_real.csv

--2024-07-14 07:18:28-- https://raw.githubusercontent.com/KaiDMML/FakeNewsNet/m
aster/dataset/gossipcop_fake.csv
Resolving raw.githubusercontent.com (raw.githubusercontent.com)...
185.199.108.133, 185.199.109.133, 185.199.110.133, ...
Connecting to raw.githubusercontent.com
(raw.githubusercontent.com)|185.199.108.133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 12538704 (12M) [text/plain]
Saving to: 'gossipcop_fake.csv'

gossipcop_fake.csv 100%[=====>] 11.96M 79.1MB/s in 0.2s

2024-07-14 07:18:28 (79.1 MB/s) - 'gossipcop_fake.csv' saved [12538704/12538704]

--2024-07-14 07:18:28-- https://raw.githubusercontent.com/KaiDMML/FakeNewsNet/m
aster/dataset/gossipcop_real.csv
Resolving raw.githubusercontent.com (raw.githubusercontent.com)...
185.199.111.133, 185.199.109.133, 185.199.110.133, ...
Connecting to raw.githubusercontent.com
(raw.githubusercontent.com)|185.199.111.133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 19978126 (19M) [text/plain]
Saving to: 'gossipcop_real.csv'

gossipcop_real.csv 100%[=====>] 19.05M 96.8MB/s in 0.2s

2024-07-14 07:18:28 (96.8 MB/s) - 'gossipcop_real.csv' saved [19978126/19978126]

--2024-07-14 07:18:28-- https://raw.githubusercontent.com/KaiDMML/FakeNewsNet/m
aster/dataset/politifact_fake.csv
Resolving raw.githubusercontent.com (raw.githubusercontent.com)...
185.199.109.133, 185.199.110.133, 185.199.108.133, ...
Connecting to raw.githubusercontent.com
(raw.githubusercontent.com)|185.199.109.133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 3286418 (3.1M) [text/plain]
Saving to: 'politifact_fake.csv'
```

```
politifact_fake.csv 100%[=====>] 3.13M --.-KB/s in 0.08s
```

```
2024-07-14 07:18:29 (38.7 MB/s) - 'politifact_fake.csv' saved [3286418/3286418]
```

```
--2024-07-14 07:18:29-- https://raw.githubusercontent.com/KaiDMML/FakeNewsNet/master/dataset/politifact_real.csv
```

```
Resolving raw.githubusercontent.com (raw.githubusercontent.com)...
```

```
185.199.109.133, 185.199.111.133, 185.199.108.133, ...
```

```
Connecting to raw.githubusercontent.com
```

```
(raw.githubusercontent.com)|185.199.109.133|:443... connected.
```

```
HTTP request sent, awaiting response... 200 OK
```

```
Length: 8278658 (7.9M) [text/plain]
```

```
Saving to: 'politifact_real.csv'
```

```
politifact_real.csv 100%[=====>] 7.89M 52.3MB/s in 0.2s
```

```
2024-07-14 07:18:29 (52.3 MB/s) - 'politifact_real.csv' saved [8278658/8278658]
```

1.3 Load the Data

```
[3]: # Load the datasets
gossipcop_fake = pd.read_csv('gossipcop_fake.csv')
gossipcop_real = pd.read_csv('gossipcop_real.csv')
politifact_fake = pd.read_csv('politifact_fake.csv')
politifact_real = pd.read_csv('politifact_real.csv')

# Quick check of the data loaded
print(gossipcop_fake.head())
print(gossipcop_real.head())
print(politifact_fake.head())
print(politifact_real.head())
```

```
          id                               news_url \
0  gossipcop-2493749932  www.dailymail.co.uk/tvshowbiz/article-5874213/...
1  gossipcop-4580247171  hollywoodlife.com/2018/05/05/paris-jackson-car...
2  gossipcop-941805037  variety.com/2017/biz/news/tax-march-donald-tru...
3  gossipcop-2547891536  www.dailymail.co.uk/femail/article-3499192/Do-...
4  gossipcop-5476631226  variety.com/2018/film/news/list-2018-oscar-nom...
```

```
          title \
0  Did Miley Cyrus and Liam Hemsworth secretly ge...
1  Paris Jackson & Cara Delevingne Enjoy Night Ou...
2  Celebrities Join Tax March in Protest of Donal...
3  Cindy Crawford's daughter Kaia Gerber wears a ...
4  Full List of 2018 Oscar Nominations - Variety
```

	tweet_ids		
0	284329075902926848\t284332744559968256\t284335...		
1	992895508267130880\t992897935418503169\t992899...		
2	853359353532829696\t853359576543920128\t853359...		
3	988821905196158981\t988824206556172288\t988825...		
4	955792793632432131\t955795063925301249\t955798...		
	id	news_url	\
0	gossipcop-882573	https://www.brides.com/story/teen-mom-jenelle-...	
1	gossipcop-875924	https://www.dailymail.co.uk/tvshowbiz/article-...	
2	gossipcop-894416	https://en.wikipedia.org/wiki/Quinn_Perkins	
3	gossipcop-857248	https://www.refinery29.com/en-us/2018/03/19192...	
4	gossipcop-884684	https://www.cnn.com/2017/10/04/entertainment/c...	
	title	\	
0	Teen Mom Star Jenelle Evans' Wedding Dress Is ...		
1	Kylie Jenner refusing to discuss Tyga on Life ...		
2	Quinn Perkins		
3	I Tried Kim Kardashian's Butt Workout & Am For...		
4	Celine Dion donates concert proceeds to Vegas ...		
	tweet_ids		
0	912371411146149888\t912371528343408641\t912372...		
1	901989917546426369\t901989992074969089\t901990...		
2	931263637246881792\t931265332022579201\t931265...		
3	868114761723936769\t868122567910936576\t868128...		
4	915528047004209152\t915529285171122176\t915530...		
	id	news_url	\
0	politifact15014	speedtalk.com/forum/viewtopic.php?t=51650	
1	politifact15156	politics2020.info/index.php/2018/03/13/court-o...	
2	politifact14745	www.nscdscamps.org/blog/category/parenting/467...	
3	politifact14355	https://howafrica.com/oscar-pistorius-attempts...	
4	politifact15371	http://washingtonsources.org/trump-votes-for-d...	
	title	\	
0	BREAKING: First NFL Team Declares Bankruptcy O...		
1	Court Orders Obama To Pay \$400 Million In Rest...		
2	UPDATE: Second Roy Moore Accuser Works For Mic...		
3	Oscar Pistorius Attempts To Commit Suicide		
4	Trump Votes For Death Penalty For Being Gay		
	tweet_ids		
0	937349434668498944\t937379378006282240\t937380...		
1	972666281441878016\t972678396575559680\t972827...		
2	929405740732870656\t929439450400264192\t929439...		
3	886941526458347521\t887011300278194176\t887023...		
4	915205698212040704\t915242076681506816\t915249...		
	id	news_url	\
0	politifact14984	http://www.nfib-sbet.org/	

```

1 politifact12944 http://www.cq.com/doc/newsmakertranscripts-494...
2 politifact333 https://web.archive.org/web/20080204072132/htt...
3 politifact4358 https://web.archive.org/web/20110811143753/htt...
4 politifact779 https://web.archive.org/web/20070820164107/htt...

```

```

                                title \
0      National Federation of Independent Business
1                                comments in Fayetteville NC
2 Romney makes pitch, hoping to close deal : Ele...
3 Democratic Leaders Say House Democrats Are Uni...
4      Budget of the United States Government, FY 2008

```

```

                                tweet_ids
0  967132259869487105\t967164368768196609\t967215...
1  942953459\t8980098198\t16253717352\t1668513250...
2
3
4  89804710374154240\t91270460595109888\t96039619...

```

1.4 Data Inspection

```

[4]: # Display basic information about each dataset
gossipcop_fake.info()
gossipcop_real.info()
politifact_fake.info()
politifact_real.info()

```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5323 entries, 0 to 5322
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   id           5323 non-null   object
1   news_url     5067 non-null   object
2   title        5323 non-null   object
3   tweet_ids    5135 non-null   object
dtypes: object(4)
memory usage: 166.5+ KB
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 16817 entries, 0 to 16816
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   id           16817 non-null   object
1   news_url     16804 non-null   object
2   title        16817 non-null   object
3   tweet_ids    15759 non-null   object
dtypes: object(4)

```

```

memory usage: 525.7+ KB
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 432 entries, 0 to 431
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0   id           432 non-null    object
1   news_url     428 non-null    object
2   title        432 non-null    object
3   tweet_ids    392 non-null    object
dtypes: object(4)
memory usage: 13.6+ KB
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 624 entries, 0 to 623
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0   id           624 non-null    object
1   news_url     567 non-null    object
2   title        624 non-null    object
3   tweet_ids    409 non-null    object
dtypes: object(4)
memory usage: 19.6+ KB

```

1.5 Initial Exploratory Data Analysis (EDA)

1.5.1 Visualization of Class Distribution

```

[5]: # Label the data
gossipcop_fake['label'] = 'fake'
gossipcop_real['label'] = 'real'
politifact_fake['label'] = 'fake'
politifact_real['label'] = 'real'

# Add dataset identifier
gossipcop_fake['dataset'] = 'GossipCop'
gossipcop_real['dataset'] = 'GossipCop'
politifact_fake['dataset'] = 'PolitiFact'
politifact_real['dataset'] = 'PolitiFact'

# Combine the datasets
combined = pd.concat([
    gossipcop_fake[['label', 'dataset']],
    gossipcop_real[['label', 'dataset']],
    politifact_fake[['label', 'dataset']],
    politifact_real[['label', 'dataset']]
])

```

```

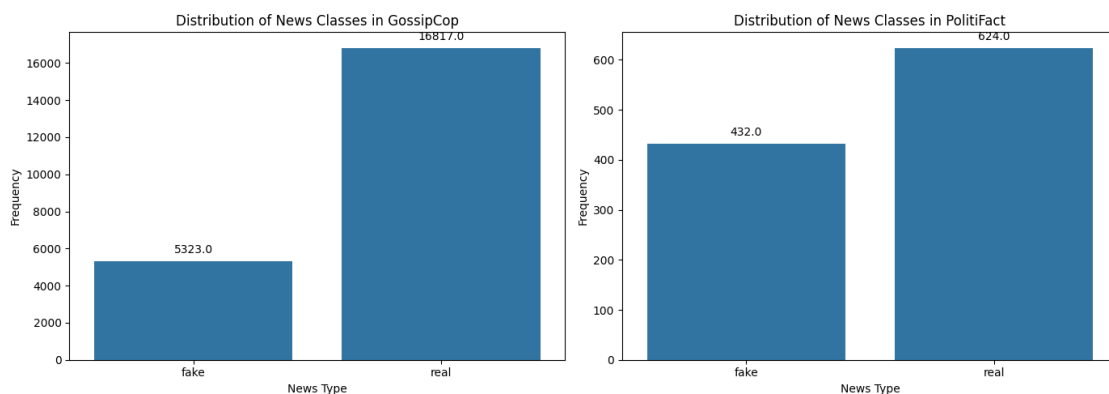
# Plotting
fig, axes = plt.subplots(1, 2, figsize=(14, 5))

# Plot for GossipCop
gossipcop_plot = sns.countplot(x='label', data=combined[combined['dataset'] == 'GossipCop'], ax=axes[0])
axes[0].set_title('Distribution of News Classes in GossipCop')
axes[0].set_xlabel('News Type')
axes[0].set_ylabel('Frequency')
# Annotate bars with count
for p in gossipcop_plot.patches:
    gossipcop_plot.annotate(f'{p.get_height()}', (p.get_x() + p.get_width() / 2., p.get_height()), ha = 'center', va = 'center', xytext = (0, 10), textcoords = 'offset points')

# Plot for PolitiFact
politifact_plot = sns.countplot(x='label', data=combined[combined['dataset'] == 'PolitiFact'], ax=axes[1])
axes[1].set_title('Distribution of News Classes in PolitiFact')
axes[1].set_xlabel('News Type')
axes[1].set_ylabel('Frequency')
# Annotate bars with count
for p in politifact_plot.patches:
    politifact_plot.annotate(f'{p.get_height()}', (p.get_x() + p.get_width() / 2., p.get_height()), ha = 'center', va = 'center', xytext = (0, 10), textcoords = 'offset points')

plt.tight_layout()
plt.show()

```



The datasets have class imbalance. There are approximately 3 times more records of real news than fake news in the GossipCop dataset. In the PolitiFact dataset, the number of fake news articles is about 30% less than that of real news articles. This will require a strategy to balance the dataset.

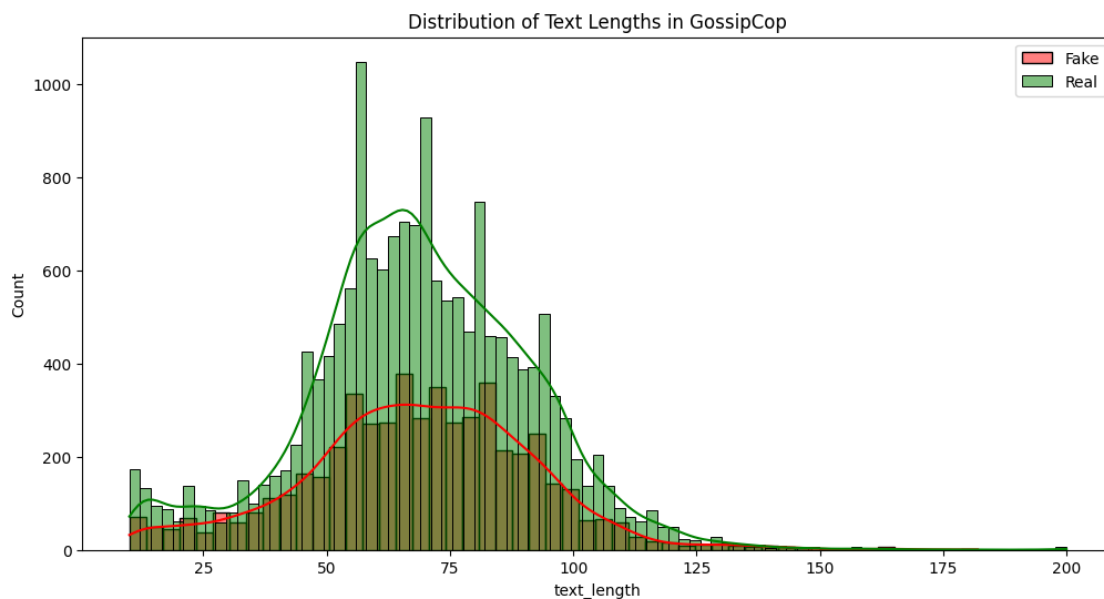
1.5.2 Text Length Distribution

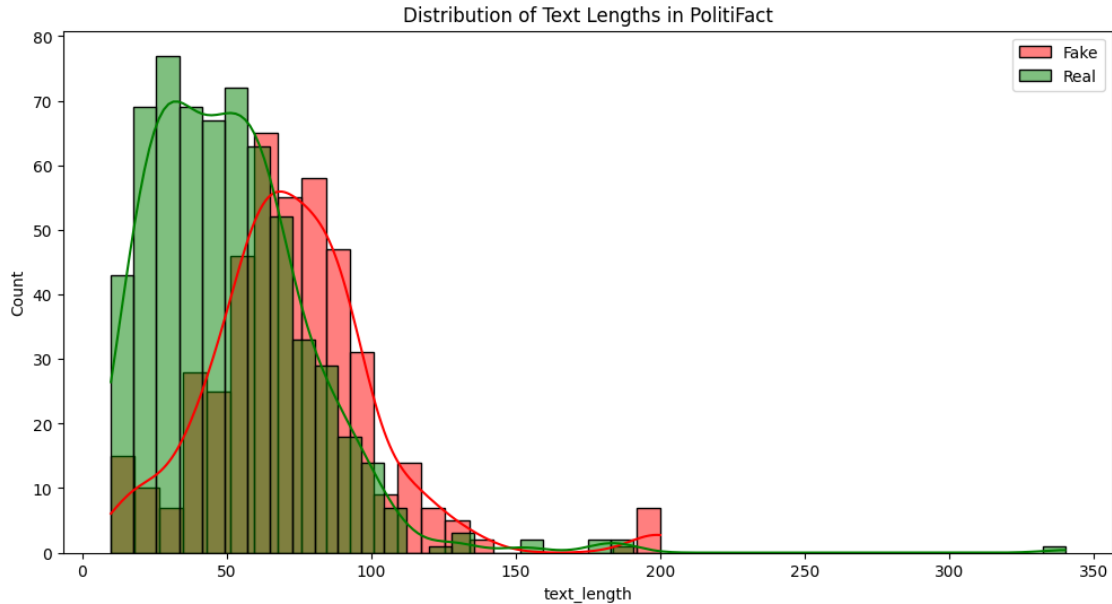
I will check text length distribution in title column of both datasets and compare the real and fake news.

```
[6]: # Add text length as a new feature
gossipcop_fake['text_length'] = gossipcop_fake['title'].apply(len)
gossipcop_real['text_length'] = gossipcop_real['title'].apply(len)
politifact_fake['text_length'] = politifact_fake['title'].apply(len)
politifact_real['text_length'] = politifact_real['title'].apply(len)

# Visualize the distribution of text length
plt.figure(figsize=(12, 6))
sns.histplot(gossipcop_fake['text_length'], color='red', label='Fake', kde=True)
sns.histplot(gossipcop_real['text_length'], color='green', label='Real',
             ↪kde=True)
plt.title('Distribution of Text Lengths in GossipCop')
plt.legend()
plt.show()

plt.figure(figsize=(12, 6))
sns.histplot(politifact_fake['text_length'], color='red', label='Fake',
             ↪kde=True)
sns.histplot(politifact_real['text_length'], color='green', label='Real',
             ↪kde=True)
plt.title('Distribution of Text Lengths in PolitiFact')
plt.legend()
plt.show()
```





The text length distributions indicate that real news generally has shorter titles compared to fake news in both GossipCop and PolitiFact datasets.

1.5.3 Word Frequency Analysis

```
[7]: # Function to get the most common words
def get_most_common_words(text_list, num=20):
    all_words = ' '.join(text_list).split()
    common_words = Counter(all_words).most_common(num)
    return pd.DataFrame(common_words, columns=['Word', 'Frequency'])

# Apply the function to each dataset
fake_words_gossipcop = get_most_common_words(gossipcop_fake['title'])
real_words_gossipcop = get_most_common_words(gossipcop_real['title'])
fake_words_politifact = get_most_common_words(politifact_fake['title'])
real_words_politifact = get_most_common_words(politifact_real['title'])

# Plotting the results
fig, axes = plt.subplots(2, 2, figsize=(15, 10))

sns.barplot(x='Frequency', y='Word', data=fake_words_gossipcop, ax=axes[0, 0],
            color='red')
axes[0, 0].set_title('Most Common Words in Fake News - GossipCop')

sns.barplot(x='Frequency', y='Word', data=real_words_gossipcop, ax=axes[0, 1],
            color='green')
axes[0, 1].set_title('Most Common Words in Real News - GossipCop')
```

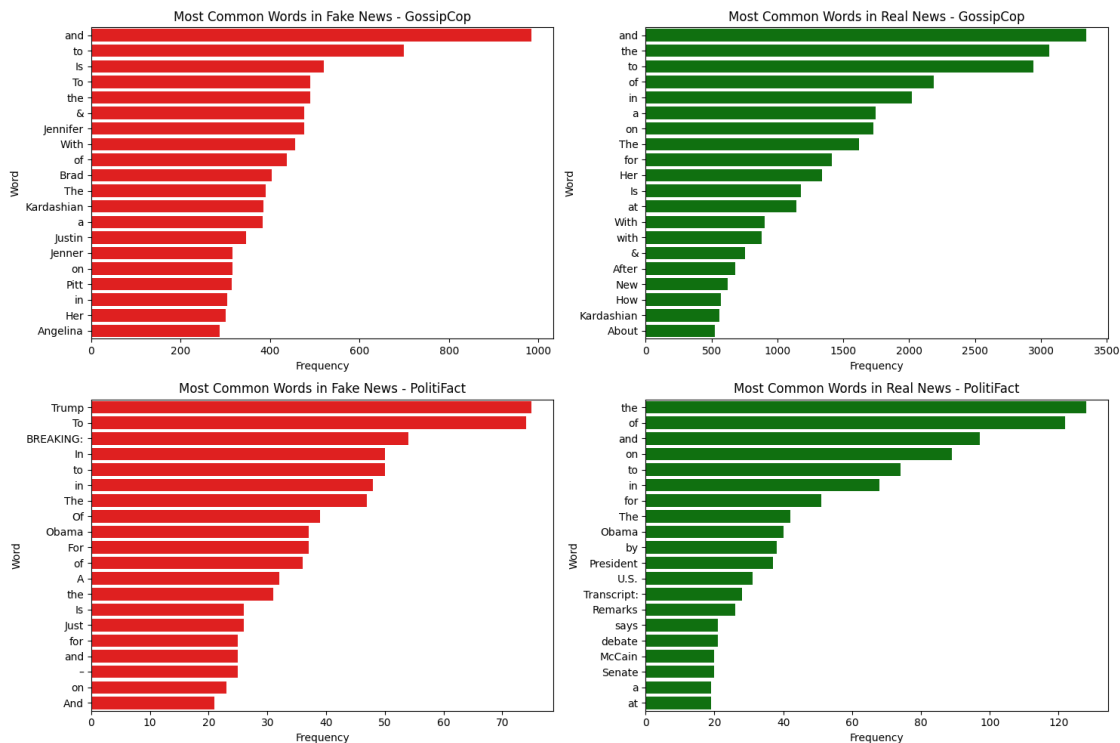
```

sns.barplot(x='Frequency', y='Word', data=fake_words_politifact, ax=axes[1, 0],
            color='red')
axes[1, 0].set_title('Most Common Words in Fake News - PolitiFact')

sns.barplot(x='Frequency', y='Word', data=real_words_politifact, ax=axes[1, 1],
            color='green')
axes[1, 1].set_title('Most Common Words in Real News - PolitiFact')

plt.tight_layout()
plt.show()

```



The most common words in fake news GossipCop are: and, to, Is. The most common words in real news GossipCop are: and, the, to. The most common words in fake news PolitiFact are: Trump, To, BREAKING:. The most common words in real news PolitiFact are: the, of, and. The results contain stopwords and punctuation and text in title column will need preprocessing.

1.6 Data Cleaning

I will clean the data by removing unnecessary columns, which will prepare the datasets to include only relevant features (text column) for model training. Most machine learning algorithms require input data, including labels, to be numeric, so I will convert labels (label column) to numeric values. I will check for missing values in all columns of each dataset which will ensure the data's integrity. Cleaning step will ensure the compatibility with the machine learning algorithms that will be used

in classifier notebook.

```
[8]: # Drop unneeded columns
columns_to_drop = ['id', 'news_url', 'tweet_ids', 'text_length', 'dataset']
gossipcop_fake.drop(columns=columns_to_drop, inplace=True)
gossipcop_real.drop(columns=columns_to_drop, inplace=True)
politifact_fake.drop(columns=columns_to_drop, inplace=True)
politifact_real.drop(columns=columns_to_drop, inplace=True)

# Convert labels to numeric values
label_mapping = {'fake': 0, 'real': 1}
gossipcop_fake['label'] = gossipcop_fake['label'].map(label_mapping)
gossipcop_real['label'] = gossipcop_real['label'].map(label_mapping)
politifact_fake['label'] = politifact_fake['label'].map(label_mapping)
politifact_real['label'] = politifact_real['label'].map(label_mapping)

# Check for missing values in all columns of each dataset
missing_values_gossipcop_fake = gossipcop_fake.isnull().sum()
missing_values_gossipcop_real = gossipcop_real.isnull().sum()
missing_values_politifact_fake = politifact_fake.isnull().sum()
missing_values_politifact_real = politifact_real.isnull().sum()

print("Missing values in GossipCop Fake:\n", missing_values_gossipcop_fake)
print("\nMissing values in GossipCop Real:\n", missing_values_gossipcop_real)
print("\nMissing values in PolitiFact Fake:\n", missing_values_politifact_fake)
print("\nMissing values in PolitiFact Real:\n", missing_values_politifact_real)

# check the datasets
print(gossipcop_fake.head())
print(gossipcop_real.head())
print(politifact_fake.head())
print(politifact_real.head())
```

Missing values in GossipCop Fake:

```
title    0
label    0
dtype: int64
```

Missing values in GossipCop Real:

```
title    0
label    0
dtype: int64
```

Missing values in PolitiFact Fake:

```
title    0
label    0
dtype: int64
```

Missing values in PolitiFact Real:

```
title    0
label    0
dtype: int64
```

	title	label
0	Did Miley Cyrus and Liam Hemsworth secretly ge...	0
1	Paris Jackson & Cara Delevingne Enjoy Night Ou...	0
2	Celebrities Join Tax March in Protest of Donal...	0
3	Cindy Crawford's daughter Kaia Gerber wears a ...	0
4	Full List of 2018 Oscar Nominations - Variety	0

	title	label
0	Teen Mom Star Jenelle Evans' Wedding Dress Is ...	1
1	Kylie Jenner refusing to discuss Tyga on Life ...	1
2	Quinn Perkins	1
3	I Tried Kim Kardashian's Butt Workout & Am For...	1
4	Celine Dion donates concert proceeds to Vegas ...	1

	title	label
0	BREAKING: First NFL Team Declares Bankruptcy O...	0
1	Court Orders Obama To Pay \$400 Million In Rest...	0
2	UPDATE: Second Roy Moore Accuser Works For Mic...	0
3	Oscar Pistorius Attempts To Commit Suicide	0
4	Trump Votes For Death Penalty For Being Gay	0

	title	label
0	National Federation of Independent Business	1
1	comments in Fayetteville NC	1
2	Romney makes pitch, hoping to close deal : Ele...	1
3	Democratic Leaders Say House Democrats Are Uni...	1
4	Budget of the United States Government, FY 2008	1

1.7 Feature Engineering

1.7.1 Text Preprocessing

```
[9]: # Initialize stopwords and stemmer
stop_words = set(stopwords.words('english'))
stemmer = PorterStemmer()

def preprocess_text(text):
    # Convert text to lowercase
    text = text.lower()
    # Remove punctuation and non-word characters
    text = re.sub(r'\W+', ' ', text)
    # Stemming and stopword removal
    words = text.split()
    filtered_words = [stemmer.stem(word) for word in words if word not in
    ↪stop_words]
    return ' '.join(filtered_words)
```

```

# Apply preprocessing to each title
gossipcop_fake['title'] = gossipcop_fake['title'].apply(preprocess_text)
gossipcop_real['title'] = gossipcop_real['title'].apply(preprocess_text)
politifact_fake['title'] = politifact_fake['title'].apply(preprocess_text)
politifact_real['title'] = politifact_real['title'].apply(preprocess_text)

```

1.7.2 Visualize the Most Common Words After Preprocessing

```

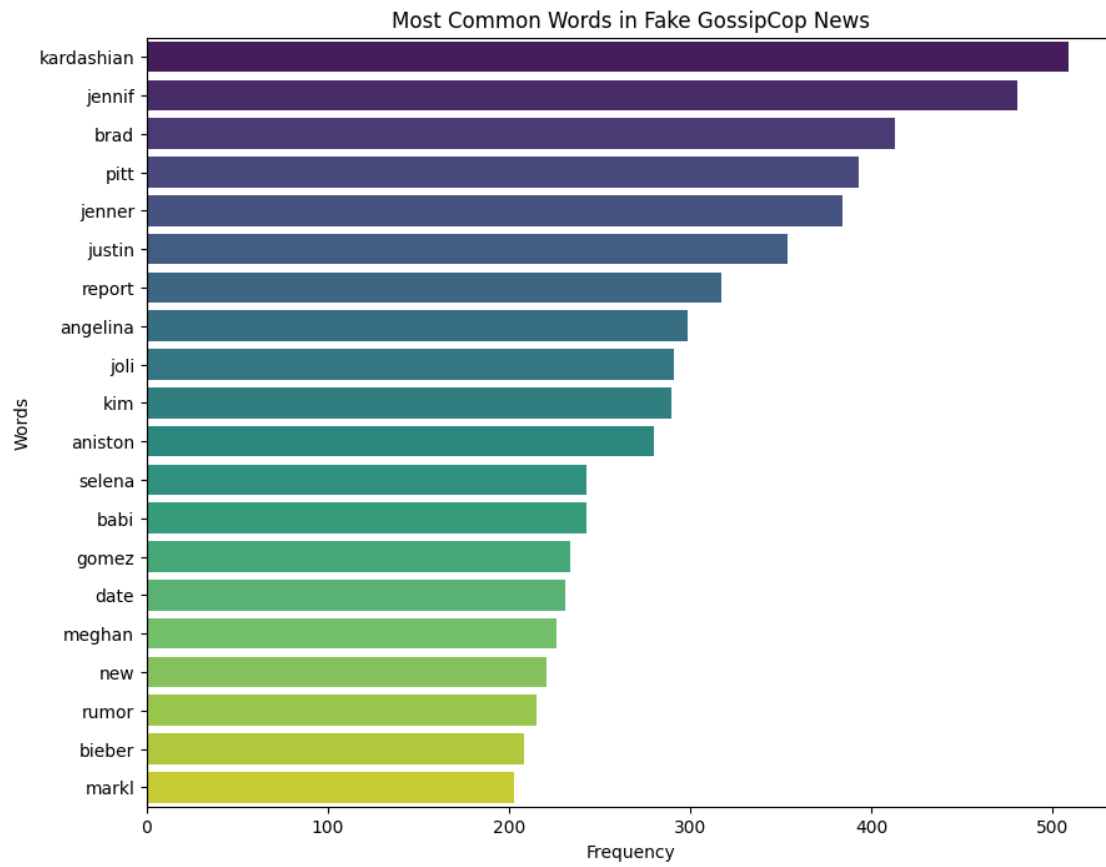
[10]: # Function to plot most common words
def plot_most_common_words(preprocessed_texts, title, num=20):
    all_words = ' '.join(preprocessed_texts).split()
    freq_dist = Counter(all_words)
    common_words = freq_dist.most_common(num)

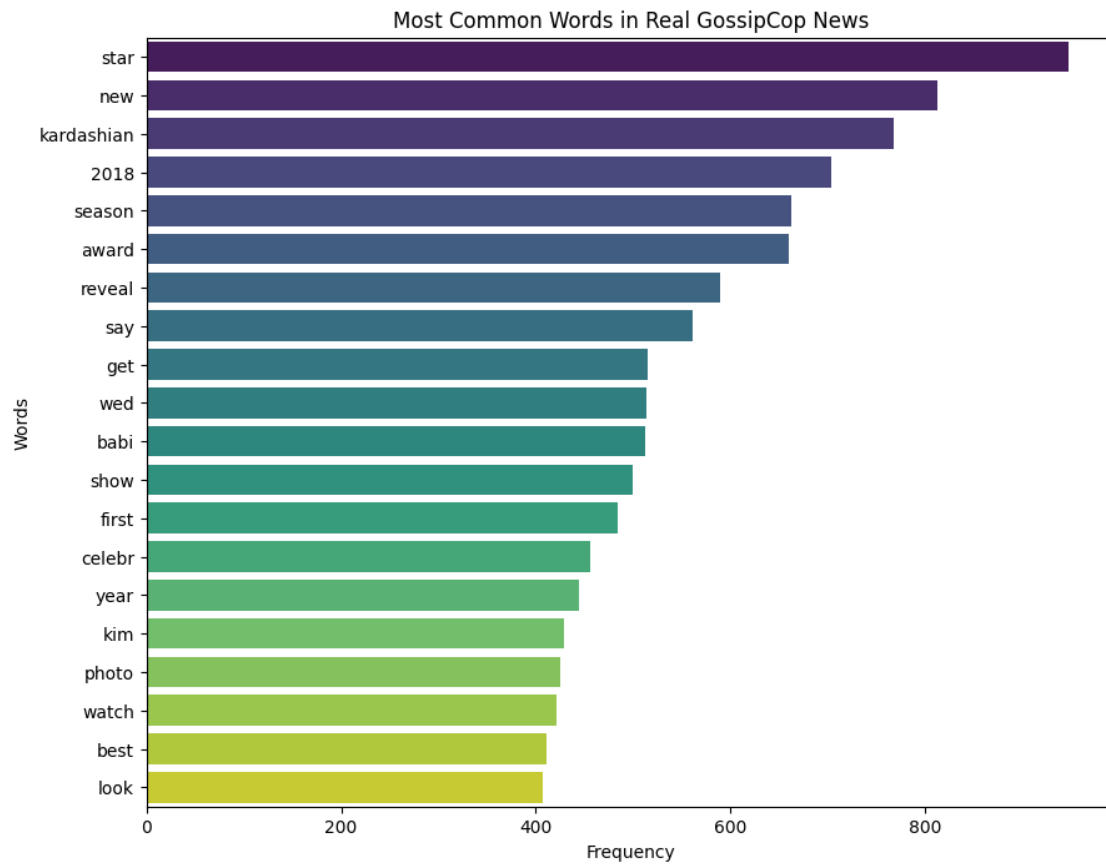
    words = [word[0] for word in common_words]
    counts = [word[1] for word in common_words]
    data = pd.DataFrame({'Word': words, 'Frequency': counts})

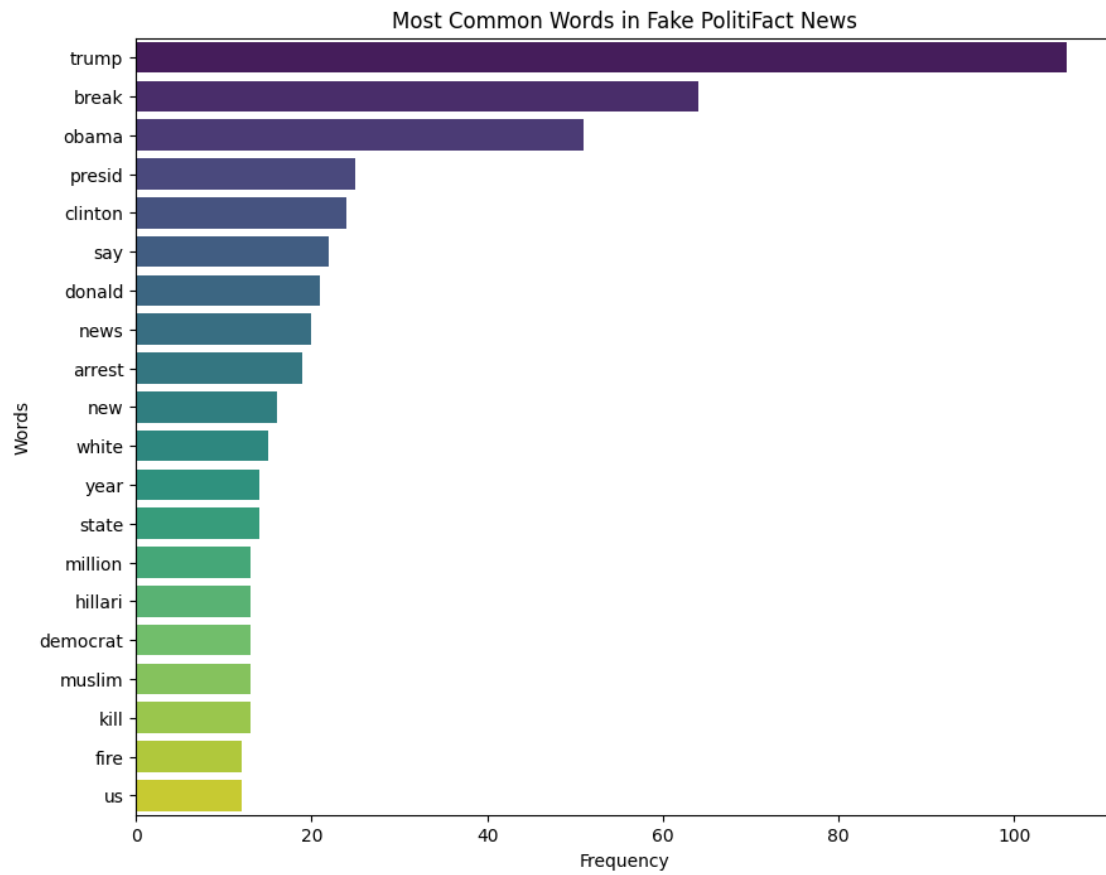
    plt.figure(figsize=(10, 8))
    sns.barplot(data=data, y='Word', x='Frequency', hue='Word', dodge=False,
        palette="viridis")
    plt.legend([], [], frameon=False) # Hides the legend as it's redundant in
    this context
    plt.title(f'Most Common Words in {title}')
    plt.xlabel('Frequency')
    plt.ylabel('Words')
    plt.show()

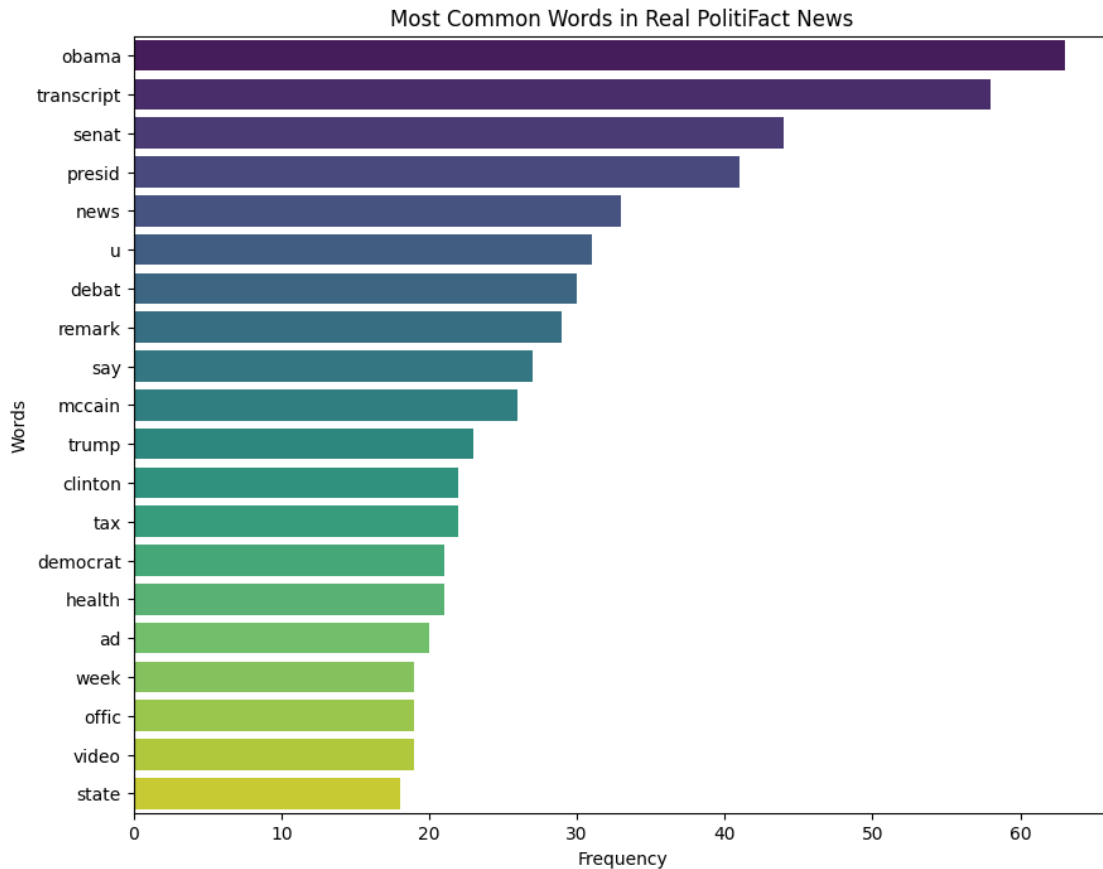
# Apply the function to visualize data
plot_most_common_words(gossipcop_fake['title'], 'Fake GossipCop News')
plot_most_common_words(gossipcop_real['title'], 'Real GossipCop News')
plot_most_common_words(politifact_fake['title'], 'Fake PolitiFact News')
plot_most_common_words(politifact_real['title'], 'Real PolitiFact News')

```









The most common words in fake news GossipCop are: kardashian, jennif, brad. The most common words in real news GossipCop are: star, new, kardashian. The most common words in fake news PolitiFact are: trump, break, obama. The most common words in real news PolitiFact are: obama, transcript, senat.

1.8 Combine and Shuffle the Datasets

```
[11]: # Combine fake and real datasets for GossipCop and shuffle
gossipcop_combined = pd.concat([gossipcop_fake, gossipcop_real]).
    ↪reset_index(drop=True)
gossipcop_combined = shuffle(gossipcop_combined, random_state=42)

# Combine fake and real datasets for PolitiFact and shuffle
politifact_combined = pd.concat([politifact_fake, politifact_real]).
    ↪reset_index(drop=True)
politifact_combined = shuffle(politifact_combined, random_state=42)

# Check the combined datasets
print(gossipcop_combined.head())
print(politifact_combined.head())
```

		title	label
11080	lea michel hairstylist mix textur spray coconu...		1
291	thoma markl princ harri polit miss daughter me...		0
17231	2019 sag award nomin see full list nomine varieti		1
16382	see megghan markl royal coat arm symbol hide wi...		1
9364	kyli jenner visit shaman life kyli season final		1

		title	label
260	world popular candi remov shelv octob 2017		0
832	brows congression bill		1
846	suprem court vacanc video		1
1007	u import export		1
88	die 78 year old cia agent admit kill marilyn m...		0

1.9 Save Preprocessed datasets to Google drive

```
[12]: # Mount Google Drive
from google.colab import drive
drive.mount('/content/drive')

# Save preprocessed data to Google Drive
gossipcop_combined.to_csv('/content/drive/My Drive/gossipcop_preprocessed.csv',
    ↪index=False)
politifact_combined.to_csv('/content/drive/My Drive/politifact_preprocessed.
    ↪csv', index=False)
```

Mounted at /content/drive

1.10 References

[7] SHU, K., MAHUDESWARAN, D., WANG, S., LEE, D., and LIU, H. 2018. FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. <https://doi.org/10.48550/arXiv.1809.01286>

[8] SHU, K., MAHUDESWARAN, D., WANG, S., LEE, D., and LIU, H. 2018. FakeNewsNet: A Data Repository with News Content, Social Context, and Dynamic Information for Studying Fake News on Social Media. <https://github.com/KaiDMML/FakeNewsNet>

```
[ ]: # Install LaTeX packages necessary for converting notebooks to PDF
!apt-get update
!apt-get install -y texlive-xetex texlive-fonts-recommended
    ↪texlive-plain-generic texlive-latex-extra pandoc

# Convert the notebook to PDF
!jupyter nbconvert --to pdf "/content/drive/My Drive/Colab Notebooks/
    ↪FakeNewsNetPreprocessing.ipynb"
```

Get:1 <https://cloud.r-project.org/bin/linux/ubuntu/jammy-cran40/> InRelease
[3,626 B]

Hit:2 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2204/x86_64

```

InRelease
Get:3 http://security.ubuntu.com/ubuntu jammy-security InRelease [129 kB]
Hit:4 http://archive.ubuntu.com/ubuntu jammy InRelease
Ign:5 https://r2u.stat.illinois.edu/ubuntu jammy InRelease
Get:6 https://r2u.stat.illinois.edu/ubuntu jammy Release [5,713 B]
Get:7 https://r2u.stat.illinois.edu/ubuntu jammy Release.gpg [793 B]
Get:8 http://archive.ubuntu.com/ubuntu jammy-updates InRelease [128 kB]
Hit:9 https://ppa.launchpadcontent.net/deadsnakes/ppa/ubuntu jammy InRelease
Get:10 https://ppa.launchpadcontent.net/graphics-drivers/ppa/ubuntu jammy
InRelease [24.3 kB]
Hit:11 http://archive.ubuntu.com/ubuntu jammy-backports InRelease
Hit:12 https://ppa.launchpadcontent.net/ubuntugis/ppa/ubuntu jammy InRelease
Get:13 https://r2u.stat.illinois.edu/ubuntu jammy/main amd64 Packages [2,544 kB]
Get:14 http://security.ubuntu.com/ubuntu jammy-security/universe amd64 Packages
[1,127 kB]
Get:15 https://r2u.stat.illinois.edu/ubuntu jammy/main all Packages [8,181 kB]
Get:16 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 Packages [2,263
kB]
Get:17 http://security.ubuntu.com/ubuntu jammy-security/main amd64 Packages
[1,998 kB]
Get:18 https://ppa.launchpadcontent.net/graphics-drivers/ppa/ubuntu jammy/main
amd64 Packages [48.1 kB]
Get:19 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 Packages
[1,410 kB]
Fetched 17.9 MB in 2s (7,257 kB/s)
Reading package lists... Done
W: Skipping acquire of configured file 'main/source/Sources' as repository
'https://r2u.stat.illinois.edu/ubuntu jammy InRelease' does not seem to provide
it (sources.list entry misspelt?)
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono fonts-
texgyre
  fonts-urw-base35 libapache-pom-java libcmark-gfm-extensions0.29.0.gfm.3
libcmark-gfm0.29.0.gfm.3
  libcommons-logging-java libcommons-parent-java libfontbox-java libfontenc1
libgs9 libgs9-common
  libidn12 libijs-0.35 libjbig2dec0 libkpathsea6 libpdfbox-java libptexenc1
libruby3.0 libsynchronet2
  libteckit0 libtexlua53 libtexluajit2 libwoff1 libzip-0-13 lmodern pandoc-data
poppler-data
  preview-latex-style rake ruby ruby-net-telnet ruby-rubygems ruby-webrick ruby-
xmlrpc ruby3.0
  rubygems-integration tlutils teckit tex-common tex-gyre texlive-base texlive-
binaries
  texlive-latex-base texlive-latex-recommended texlive-pictures tipa xfonts-

```

encodings xfonts-utils

Suggested packages:

fonts-noto fonts-freefont-otf | fonts-freefont-ttf libavalon-framework-java
libcommons-logging-java-doc libexcalibur-logkit-java liblog4j1.2-java texlive-
luatex
pandoc-citeproc context wkhtmltopdf librsvg2-bin groff ghc nodejs php python
libjs-mathjax
libjs-katex citation-style-language-styles poppler-utils ghostscript fonts-
japanese-mincho
| fonts-ipafont-mincho fonts-japanese-gothic | fonts-ipafont-gothic fonts-
arphic-ukai
fonts-arphic-uming fonts-nanum ri ruby-dev bundler debhelper gv | postscript-
viewer perl-tk xpdf
| pdf-viewer xzdec texlive-fonts-recommended-doc texlive-latex-base-doc
python3-pygments
icc-profiles libfile-which-perl libspreadsheet-parseexcel-perl texlive-latex-
extra-doc
texlive-latex-recommended-doc texlive-pstricks dot2tex prerex texlive-
pictures-doc vprerex
default-jre-headless tipa-doc

The following NEW packages will be installed:

dvisvgm fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono fonts-
texgyre
fonts-urw-base35 libapache-pom-java libcmark-gfm-extensions0.29.0.gfm.3
libcmark-gfm0.29.0.gfm.3
libcommons-logging-java libcommons-parent-java libfontbox-java libfontenc1
libgs9 libgs9-common
libidn12 libijs-0.35 libjbig2dec0 libkpathsea6 libpdfbox-java libptexenc1
libruby3.0 libsyntax2
libteckit0 libtexlua53 libtexluajit2 libwoff1 libzip-0-13 lmodern pandoc
pandoc-data
poppler-data preview-latex-style rake ruby ruby-net-telnet ruby-rubygems ruby-
webrick ruby-xmlrpc
ruby3.0 rubygems-integration tlutils teckit tex-common tex-gyre texlive-base
texlive-binaries
texlive-fonts-recommended texlive-latex-base texlive-latex-extra texlive-
latex-recommended
texlive-pictures texlive-plain-generic texlive-xetex tipa xfonts-encodings
xfonts-utils

0 upgraded, 58 newly installed, 0 to remove and 45 not upgraded.

Need to get 202 MB of archives.

After this operation, 728 MB of additional disk space will be used.

Get:1 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-droid-fallback all
1:6.0.1r16-1.1build1 [1,805 kB]

Get:2 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-lato all 2.0-2.1
[2,696 kB]

Get:3 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 poppler-data all
0.4.11-1 [2,171 kB]

Get:4 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 tex-common all 6.17 [33.7 kB]
Get:5 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-urw-base35 all 20200910-1 [6,367 kB]
Get:6 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libgs9-common all 9.55.0~dfsg1-0ubuntu5.7 [752 kB]
Get:7 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libidn12 amd64 1.38-4ubuntu1 [60.0 kB]
Get:8 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libijs-0.35 amd64 0.35-15build2 [16.5 kB]
Get:9 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libjbig2dec0 amd64 0.19-3build2 [64.7 kB]
Get:10 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libgs9 amd64 9.55.0~dfsg1-0ubuntu5.7 [5,028 kB]
Get:11 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libkpathsea6 amd64 2021.20210626.59705-1ubuntu0.2 [60.4 kB]
Get:12 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libwoff1 amd64 1.0.2-1build4 [45.2 kB]
Get:13 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 dvisvgm amd64 2.13.1-1 [1,221 kB]
Get:14 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 fonts-lmodern all 2.004.5-6.1 [4,532 kB]
Get:15 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 fonts-noto-mono all 20201225-1build1 [397 kB]
Get:16 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 fonts-texgyre all 20180621-3.1 [10.2 MB]
Get:17 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libapache-pom-java all 18-1 [4,720 B]
Get:18 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libcmark-gfm0.29.0.gfm.3 amd64 0.29.0.gfm.3-3 [115 kB]
Get:19 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libcmark-gfm-extensions0.29.0.gfm.3 amd64 0.29.0.gfm.3-3 [25.1 kB]
Get:20 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libcommons-parent-java all 43-1 [10.8 kB]
Get:21 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libcommons-logging-java all 1.2-2 [60.3 kB]
Get:22 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 libfontenc1 amd64 1:1.1.4-1build3 [14.7 kB]
Get:23 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libptexenc1 amd64 2021.20210626.59705-1ubuntu0.2 [39.1 kB]
Get:24 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 rubygems-integration all 1.18 [5,336 B]
Get:25 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 ruby3.0 amd64 3.0.2-7ubuntu2.7 [50.1 kB]
Get:26 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 ruby-rubygems all 3.3.5-2 [228 kB]
Get:27 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 ruby amd64 1:3.0~exp1 [5,100 B]

Get:28 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 rake all 13.0.6-2 [61.7 kB]
Get:29 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 ruby-net-telnet all 0.1.1-2 [12.6 kB]
Get:30 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 ruby-webrick all 1.7.0-3 [51.8 kB]
Get:31 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 ruby-xmlrpc all 0.3.2-1ubuntu0.1 [24.9 kB]
Get:32 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libruby3.0 amd64 3.0.2-7ubuntu2.7 [5,113 kB]
Get:33 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libsyntax2 amd64 2021.20210626.59705-1ubuntu0.2 [55.6 kB]
Get:34 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libteckit0 amd64 2.5.11+ds1-1 [421 kB]
Get:35 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libtexlua53 amd64 2021.20210626.59705-1ubuntu0.2 [120 kB]
Get:36 <http://archive.ubuntu.com/ubuntu> jammy-updates/main amd64 libtexluajit2 amd64 2021.20210626.59705-1ubuntu0.2 [267 kB]
Get:37 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libzip-0-13 amd64 0.13.72+dfsg.1-1.1 [27.0 kB]
Get:38 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 xfonts-encodings all 1:1.0.5-0ubuntu2 [578 kB]
Get:39 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 xfonts-utils amd64 1:7.7+6build2 [94.6 kB]
Get:40 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 lmodern all 2.004.5-6.1 [9,471 kB]
Get:41 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 pandoc-data all 2.9.2.1-3ubuntu2 [81.8 kB]
Get:42 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 pandoc amd64 2.9.2.1-3ubuntu2 [20.3 MB]
Get:43 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 preview-latex-style all 12.2-1ubuntu1 [185 kB]
Get:44 <http://archive.ubuntu.com/ubuntu> jammy/main amd64 t1utils amd64 1.41-4build2 [61.3 kB]
Get:45 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 teckit amd64 2.5.11+ds1-1 [699 kB]
Get:46 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 tex-gyre all 20180621-3.1 [6,209 kB]
Get:47 <http://archive.ubuntu.com/ubuntu> jammy-updates/universe amd64 texlive-binaries amd64 2021.20210626.59705-1ubuntu0.2 [9,860 kB]
Get:48 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 texlive-base all 2021.20220204-1 [21.0 MB]
Get:49 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 texlive-fonts-recommended all 2021.20220204-1 [4,972 kB]
Get:50 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 texlive-latex-base all 2021.20220204-1 [1,128 kB]
Get:51 <http://archive.ubuntu.com/ubuntu> jammy/universe amd64 libfontbox-java all 1:1.8.16-2 [207 kB]

```

Get:52 http://archive.ubuntu.com/ubuntu jammy/universe amd64 libpdfbox-java all
1:1.8.16-2 [5,199 kB]
Get:53 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-
recommended all 2021.20220204-1 [14.4 MB]
Get:54 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-pictures
all 2021.20220204-1 [8,720 kB]
Get:55 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-latex-extra
all 2021.20220204-1 [13.9 MB]
Get:56 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-plain-
generic all 2021.20220204-1 [27.5 MB]
Get:57 http://archive.ubuntu.com/ubuntu jammy/universe amd64 tipa all 2:1.3-21
[2,967 kB]
Get:58 http://archive.ubuntu.com/ubuntu jammy/universe amd64 texlive-xetex all
2021.20220204-1 [12.4 MB]
Fetched 202 MB in 7s (27.1 MB/s)
Extracting templates from packages: 100%
Preconfiguring packages ...
Selecting previously unselected package fonts-droid-fallback.
(Reading database ... 123576 files and directories currently installed.)
Preparing to unpack .../00-fonts-droid-fallback_1%3a6.0.1r16-1.1build1_all.deb
...
Unpacking fonts-droid-fallback (1:6.0.1r16-1.1build1) ...
Selecting previously unselected package fonts-lato.
Preparing to unpack .../01-fonts-lato_2.0-2.1_all.deb ...
Unpacking fonts-lato (2.0-2.1) ...
Selecting previously unselected package poppler-data.
Preparing to unpack .../02-poppler-data_0.4.11-1_all.deb ...
Unpacking poppler-data (0.4.11-1) ...
Selecting previously unselected package tex-common.
Preparing to unpack .../03-tex-common_6.17_all.deb ...
Unpacking tex-common (6.17) ...
Selecting previously unselected package fonts-urw-base35.
Preparing to unpack .../04-fonts-urw-base35_20200910-1_all.deb ...
Unpacking fonts-urw-base35 (20200910-1) ...
Selecting previously unselected package libgs9-common.
Preparing to unpack .../05-libgs9-common_9.55.0~dfsg1-0ubuntu5.7_all.deb ...
Unpacking libgs9-common (9.55.0~dfsg1-0ubuntu5.7) ...
Selecting previously unselected package libidn12:amd64.
Preparing to unpack .../06-libidn12_1.38-4ubuntu1_amd64.deb ...
Unpacking libidn12:amd64 (1.38-4ubuntu1) ...
Selecting previously unselected package libijs-0.35:amd64.
Preparing to unpack .../07-libijs-0.35_0.35-15build2_amd64.deb ...
Unpacking libijs-0.35:amd64 (0.35-15build2) ...
Selecting previously unselected package libjbig2dec0:amd64.
Preparing to unpack .../08-libjbig2dec0_0.19-3build2_amd64.deb ...
Unpacking libjbig2dec0:amd64 (0.19-3build2) ...
Selecting previously unselected package libgs9:amd64.
Preparing to unpack .../09-libgs9_9.55.0~dfsg1-0ubuntu5.7_amd64.deb ...

```

```

Unpacking libgs9:amd64 (9.55.0~dfsg1-0ubuntu5.7) ...
Selecting previously unselected package libkpathsea6:amd64.
Preparing to unpack .../10-libkpathsea6_2021.20210626.59705-1ubuntu0.2_amd64.deb
...
Unpacking libkpathsea6:amd64 (2021.20210626.59705-1ubuntu0.2) ...
Selecting previously unselected package libwoff1:amd64.
Preparing to unpack .../11-libwoff1_1.0.2-1build4_amd64.deb ...
Unpacking libwoff1:amd64 (1.0.2-1build4) ...
Selecting previously unselected package dvisvgm.
Preparing to unpack .../12-dvisvgm_2.13.1-1_amd64.deb ...
Unpacking dvisvgm (2.13.1-1) ...
Selecting previously unselected package fonts-lmodern.
Preparing to unpack .../13-fonts-lmodern_2.004.5-6.1_all.deb ...
Unpacking fonts-lmodern (2.004.5-6.1) ...
Selecting previously unselected package fonts-noto-mono.
Preparing to unpack .../14-fonts-noto-mono_20201225-1build1_all.deb ...
Unpacking fonts-noto-mono (20201225-1build1) ...
Selecting previously unselected package fonts-texgyre.
Preparing to unpack .../15-fonts-texgyre_20180621-3.1_all.deb ...
Unpacking fonts-texgyre (20180621-3.1) ...
Selecting previously unselected package libapache-pom-java.
Preparing to unpack .../16-libapache-pom-java_18-1_all.deb ...
Unpacking libapache-pom-java (18-1) ...
Selecting previously unselected package libcmark-gfm0.29.0.gfm.3:amd64.
Preparing to unpack .../17-libcmark-gfm0.29.0.gfm.3_0.29.0.gfm.3-3_amd64.deb ...
Unpacking libcmark-gfm0.29.0.gfm.3:amd64 (0.29.0.gfm.3-3) ...
Selecting previously unselected package libcmark-gfm-
extensions0.29.0.gfm.3:amd64.
Preparing to unpack .../18-libcmark-gfm-
extensions0.29.0.gfm.3_0.29.0.gfm.3-3_amd64.deb ...
Unpacking libcmark-gfm-extensions0.29.0.gfm.3:amd64 (0.29.0.gfm.3-3) ...
Selecting previously unselected package libcommons-parent-java.
Preparing to unpack .../19-libcommons-parent-java_43-1_all.deb ...
Unpacking libcommons-parent-java (43-1) ...
Selecting previously unselected package libcommons-logging-java.
Preparing to unpack .../20-libcommons-logging-java_1.2-2_all.deb ...
Unpacking libcommons-logging-java (1.2-2) ...
Selecting previously unselected package libfontenc1:amd64.
Preparing to unpack .../21-libfontenc1_1%3a1.1.4-1build3_amd64.deb ...
Unpacking libfontenc1:amd64 (1:1.1.4-1build3) ...
Selecting previously unselected package libptexenc1:amd64.
Preparing to unpack .../22-libptexenc1_2021.20210626.59705-1ubuntu0.2_amd64.deb
...
Unpacking libptexenc1:amd64 (2021.20210626.59705-1ubuntu0.2) ...
Selecting previously unselected package rubygems-integration.
Preparing to unpack .../23-rubygems-integration_1.18_all.deb ...
Unpacking rubygems-integration (1.18) ...
Selecting previously unselected package ruby3.0.

```



```

Preparing to unpack .../24-ruby3.0_3.0.2-7ubuntu2.7_amd64.deb ...
Unpacking ruby3.0 (3.0.2-7ubuntu2.7) ...
Selecting previously unselected package ruby-rubygems.
Preparing to unpack .../25-ruby-rubygems_3.3.5-2_all.deb ...
Unpacking ruby-rubygems (3.3.5-2) ...
Selecting previously unselected package ruby.
Preparing to unpack .../26-ruby_1%3a3.0~exp1_amd64.deb ...
Unpacking ruby (1:3.0~exp1) ...
Selecting previously unselected package rake.
Preparing to unpack .../27-rake_13.0.6-2_all.deb ...
Unpacking rake (13.0.6-2) ...
Selecting previously unselected package ruby-net-telnet.
Preparing to unpack .../28-ruby-net-telnet_0.1.1-2_all.deb ...
Unpacking ruby-net-telnet (0.1.1-2) ...
Selecting previously unselected package ruby-webrick.
Preparing to unpack .../29-ruby-webrick_1.7.0-3_all.deb ...
Unpacking ruby-webrick (1.7.0-3) ...
Selecting previously unselected package ruby-xmlrpc.
Preparing to unpack .../30-ruby-xmlrpc_0.3.2-1ubuntu0.1_all.deb ...
Unpacking ruby-xmlrpc (0.3.2-1ubuntu0.1) ...
Selecting previously unselected package libruby3.0:amd64.
Preparing to unpack .../31-libruby3.0_3.0.2-7ubuntu2.7_amd64.deb ...
Unpacking libruby3.0:amd64 (3.0.2-7ubuntu2.7) ...
Selecting previously unselected package libsyntax2:amd64.
Preparing to unpack .../32-libsyntax2_2021.20210626.59705-1ubuntu0.2_amd64.deb
...
Unpacking libsyntax2:amd64 (2021.20210626.59705-1ubuntu0.2) ...
Selecting previously unselected package libteckit0:amd64.
Preparing to unpack .../33-libteckit0_2.5.11+ds1-1_amd64.deb ...
Unpacking libteckit0:amd64 (2.5.11+ds1-1) ...
Selecting previously unselected package libtexlua53:amd64.
Preparing to unpack .../34-libtexlua53_2021.20210626.59705-1ubuntu0.2_amd64.deb
...
Unpacking libtexlua53:amd64 (2021.20210626.59705-1ubuntu0.2) ...
Selecting previously unselected package libtexluajit2:amd64.
Preparing to unpack
.../35-libtexluajit2_2021.20210626.59705-1ubuntu0.2_amd64.deb ...
Unpacking libtexluajit2:amd64 (2021.20210626.59705-1ubuntu0.2) ...
Selecting previously unselected package libzip-0-13:amd64.
Preparing to unpack .../36-libzip-0-13_0.13.72+dfsg.1-1.1_amd64.deb ...
Unpacking libzip-0-13:amd64 (0.13.72+dfsg.1-1.1) ...
Selecting previously unselected package xfonts-encodings.
Preparing to unpack .../37-xfonts-encodings_1%3a1.0.5-0ubuntu2_all.deb ...
Unpacking xfonts-encodings (1:1.0.5-0ubuntu2) ...
Selecting previously unselected package xfonts-utils.
Preparing to unpack .../38-xfonts-utils_1%3a7.7+6build2_amd64.deb ...
Unpacking xfonts-utils (1:7.7+6build2) ...
Selecting previously unselected package lmodern.

```

```

Preparing to unpack .../39-lmodern_2.004.5-6.1_all.deb ...
Unpacking lmodern (2.004.5-6.1) ...
Selecting previously unselected package pandoc-data.
Preparing to unpack .../40-pandoc-data_2.9.2.1-3ubuntu2_all.deb ...
Unpacking pandoc-data (2.9.2.1-3ubuntu2) ...
Selecting previously unselected package pandoc.
Preparing to unpack .../41-pandoc_2.9.2.1-3ubuntu2_amd64.deb ...
Unpacking pandoc (2.9.2.1-3ubuntu2) ...
Selecting previously unselected package preview-latex-style.
Preparing to unpack .../42-preview-latex-style_12.2-1ubuntu1_all.deb ...
Unpacking preview-latex-style (12.2-1ubuntu1) ...
Selecting previously unselected package t1utils.
Preparing to unpack .../43-t1utils_1.41-4build2_amd64.deb ...
Unpacking t1utils (1.41-4build2) ...
Selecting previously unselected package teckit.
Preparing to unpack .../44-teckit_2.5.11+ds1-1_amd64.deb ...
Unpacking teckit (2.5.11+ds1-1) ...
Selecting previously unselected package tex-gyre.
Preparing to unpack .../45-tex-gyre_20180621-3.1_all.deb ...
Unpacking tex-gyre (20180621-3.1) ...
Selecting previously unselected package texlive-binaries.
Preparing to unpack .../46-texlive-
binaries_2021.20210626.59705-1ubuntu0.2_amd64.deb ...
Unpacking texlive-binaries (2021.20210626.59705-1ubuntu0.2) ...
Selecting previously unselected package texlive-base.
Preparing to unpack .../47-texlive-base_2021.20220204-1_all.deb ...
Unpacking texlive-base (2021.20220204-1) ...
Selecting previously unselected package texlive-fonts-recommended.
Preparing to unpack .../48-texlive-fonts-recommended_2021.20220204-1_all.deb ...
Unpacking texlive-fonts-recommended (2021.20220204-1) ...
Selecting previously unselected package texlive-latex-base.
Preparing to unpack .../49-texlive-latex-base_2021.20220204-1_all.deb ...
Unpacking texlive-latex-base (2021.20220204-1) ...
Selecting previously unselected package libfontbox-java.
Preparing to unpack .../50-libfontbox-java_1%3a1.8.16-2_all.deb ...
Unpacking libfontbox-java (1:1.8.16-2) ...
Selecting previously unselected package libpdfbox-java.
Preparing to unpack .../51-libpdfbox-java_1%3a1.8.16-2_all.deb ...
Unpacking libpdfbox-java (1:1.8.16-2) ...
Selecting previously unselected package texlive-latex-recommended.
Preparing to unpack .../52-texlive-latex-recommended_2021.20220204-1_all.deb ...
Unpacking texlive-latex-recommended (2021.20220204-1) ...
Selecting previously unselected package texlive-pictures.
Preparing to unpack .../53-texlive-pictures_2021.20220204-1_all.deb ...
Unpacking texlive-pictures (2021.20220204-1) ...

```