data

June 7, 2024

```
[7]: #
     import pandas as pd

     #
     file_path = './data/pn.tsv'

     data = pd.read_csv(file_path, sep='\t', header=None)

     #
     data.columns = ['ID', 'Label', 'Text', 'Judges', 'Usage']

     #
     data.head()
```

```
[7]:           ID  Label              Text               Judges Usage
     0  pn17q00001      0                      {"0": 3}  test
     1  pn17q00002      0                      {"0": 3}  test
     2  pn17q00003      1                      {"1": 3}  test
     3  pn17q00004      1                      {"1": 3}  test
     4  pn17q00005      1              {"0": 1, "1": 2}  test
```

```
[9]: data = pd.read_csv('./data/pn.tsv', sep='\t', header=None)
     file_path = './data/pn.tsv'

     # Assigning column names to match the provided format
     data.columns = ['ID', 'Label', 'Text', 'Judges', 'Usage']

     # Create a new DataFrame with only the required columns and rename them␣
      ↪accordingly
     converted_data = data[['Text', 'Label']].copy()
     converted_data.columns = ['review', 'sentiment']

     # Map the sentiment labels to text format
     label_mapping = {1: 'positive', 0: 'neutral', -1: 'negative'}
     converted_data['sentiment'] = converted_data['sentiment'].map(label_mapping)

     # Save the DataFrame to a CSV file
```

```
output_path = './data/reviews_with_sentiment.csv'
converted_data.to_csv(output_path, index=False)

# Display the first few rows of the converted data to confirm
converted_data.head()
```

[9]:
```
   review sentiment
0         neutral
1          neutral
2         positive
3      positive
4         positive
```

[8]: