import pandas as pd

import nltk

from nltk.sentiment.vader import SentimentIntensityAnalyzer

# Download VADER lexicon

nltk.download('vader\_lexicon')

# Load the dataset

data = pd.read\_csv('imdb\_movie\_reviews.csv') # Replace 'imdb\_movie\_reviews.csv' with your dataset file path

# Initialize the VADER sentiment analyzer

sid = SentimentIntensityAnalyzer()

# Function to get sentiment score for each review

def get\_sentiment\_score(text):

scores = sid.polarity\_scores(text)

return scores['compound']

# Add a new column for sentiment score

data['sentiment\_score'] = data['review'].apply(get\_sentiment\_score)

# Function to categorize sentiment

def categorize\_sentiment(score):

if score >= 0.05:

return 'Positive'

elif score <= -0.05:

return 'Negative'

else:

return 'Neutral'

# Add a new column for sentiment category

data['sentiment\_category'] = data['sentiment\_score'].apply(categorize\_sentiment)

# Print the first few rows of the dataset with sentiment scores and categories

print(data[['review', 'sentiment\_score', 'sentiment\_category']].head())