Title: Sentiment Analysis of News Headlines Using the VADER Sentiment Analyzer

Abstract: This thesis explores the application of sentiment analysis techniques to analyze the sentiment of news headlines. Sentiment analysis is a valuable tool in understanding the emotional tone and polarity of textual data, and it has numerous applications in various domains, including journalism, finance, and social media. In this study, we focus on sentiment analysis of news headlines and employ the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analyzer, a widely used tool for this purpose.

Introduction: In today's information-driven world, news headlines play a crucial role in shaping public perception and decision-making. Understanding the sentiment conveyed by news headlines can provide valuable insights into the prevailing public opinion and market sentiment. Sentiment analysis, a subfield of natural language processing (NLP), enables us to automate the process of sentiment classification, allowing for efficient analysis of large volumes of text.

Methodology: To perform sentiment analysis on news headlines, we utilized the VADER sentiment analyzer. VADER is a lexicon and rule-based tool designed to gauge sentiment in text by providing four key sentiment scores: Positive, Negative, Neutral, and Compound. The Compound score is a normalized representation of the overall sentiment of the text, ranging from -1 (extremely negative) to 1 (extremely positive),

with 0 indicating neutrality. The VADER sentiment analysis process was implemented in Python using the Natural Language Toolkit (NLTK) library.Th is as follows:

```
In 876
```

```python
def calculate_average_scores(score_df):
    # Select numeric columns for averaging
    numeric_cols = score_df.select_dtypes(include=['float64', 'int64']).columns

    # Calculate the average scores per ticker
    avg_scores = score_df.groupby('Ticker')[numeric_cols].mean()

    # Round the average scores to two decimal places
    avg_scores = avg_scores.round(2)

    # Sort the average scores by 'Compound' in descending order
    avg_scores = avg_scores.sort_values('Compound', ascending=False)

    return avg_scores

# Calculate and display the average scores
average_scores = calculate_average_scores(score_df)
print(average_scores)
```

Executed in 38ms, 17 Oct at 18:31:07

```
        Negative  Neutral  Positive  Compound
Ticker
AMZN        0.04     0.81      0.16      0.19
AAPL        0.05     0.82      0.13      0.13
TSLA        0.06     0.85      0.09      0.05
```

Discussion: The sentiment analysis code provided above calculates sentiment scores for a given set of news headlines. These scores include Positive, Negative, Neutral, and

Compound scores. The Compound score, in particular, provides an overall sentiment score for each headline, allowing for a quick assessment of the sentiment conveyed by the text.

Conclusion: In conclusion, sentiment analysis of news headlines using the VADER sentiment analyzer is a valuable approach for understanding the emotional tone and polarity of news articles. The code presented in this study successfully applies the VADER sentiment analysis technique to news headlines, providing a comprehensive sentiment analysis of each headline. This analysis can aid in various applications, such as news sentiment monitoring, market sentiment analysis, and trend prediction.

Citation: The sentiment analysis code used in this study is based on the VADER sentiment analyzer and the Natural Language Toolkit (NLTK) library. The VADER sentiment analyzer was developed by C.J. Hutto and Eric Gilbert and is widely used in the field of sentiment analysis.

Hutto, C.J. and Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media.

Note: Please ensure that you provide the appropriate citation for the code and the VADER sentiment analyzer as per the academic guidelines of your institution.