Cheng Chengliang

SENTIMENT EXTRACTION AND RECOGNITION ALGORITHM

This article is dedicated to analyse techniques used to determine the emotional tone or attitude of a piece of text. It's a way to understand whether a particular text expresses a positive, negative, or neutral sentiment

INTRODUCTION

Sentiment analysis is typically performed using natural language processing techniques and machine learning algorithms. The process involves analyzing the words, phrases, and context of a piece of text to identify the sentiment conveyed by it.

There is one common method involves using a pre-built sentiment lexicon, which is a collection of words and phrases that are associated with particular sentiments. These lexicons are usually curated by human experts and assigned a positive or negative score, which reflects their sentiment.

When performing sentiment analysis based on text, the algorithm will use the lexicon to assign a sentiment score to each word in the text. The scores are then aggregated to calculate an overall sentiment score for the text. This score can be used to determine whether the text is positive, negative, or neutral.

I. AFINN LEXICON

AFINN is a lexicon that assigns a sentiment score to each word in a text based on its association with positive or negative sentiment. The scores range from -5 to +5, with negative scores indicating negative sentiment and positive scores indicating positive sentiment.

Example:

```
from afinn import Afinn
afinn1 = Afinn()
text = "I love this product! It's the
best thing I've ever bought."
score = afinn1.score(text)
print(score)
```

In this example, we're using the Afinn library in Python to perform sentiment analysis on a piece of text. We create an instance of the Afinn class and then pass our text to the score() method to get a sentiment score. The resulting score is 5, which indicates a very positive sentiment.

II. TextBlob LEXICON

TextBlob is a Python library that includes a pre-built sentiment lexicon for sentiment analysis based on text. The lexicon assigns a sentiment score to each word in a text based on its association with positive or negative sentiment. The scores range from -1 to +1, with negative scores indicating negative sentiment and positive scores indicating positive sentiment.

TextBlob uses a similar approach to AFINN, but it also includes additional natural language processing features such as part-of-speech tagging and noun phrase extraction, which can help to improve the accuracy of sentiment analysis.

Here's an example of how TextBlob can be used for sentiment analysis based on text:

```
from textblob import TextBlob
text = "I hate this product. It was a
complete bad thing."
blob2 = TextBlob(text)
score = blob2.sentiment.polarity
print(score)
```

In this example, the sentiment score is -0.8, which indicates a very negative sentiment.

TextBlob is a powerful tool for natural language processing and sentiment analysis, and it can be customized to suit specific applications. It also includes features such as language detection and translation, which can be useful for analyzing text in multiple languages.

III. CONCLUSION

Sentiment analysis can be applied to a wide range of text sources, including social media posts, customer reviews, news articles, and more.

Pre-trained sentiment lexicons are readily available and can be quickly integrated into a sentiment analysis pipeline without the need for manual curation or annotation. Pre-trained lexicons are usually based on large datasets and have a wide coverage of words and phrases, making them useful for analyzing large volumes of text. Pre-trained lexicons have been developed and tested on large datasets, so their accuracy and reliability are usually well-established.

1. Pre-trained vs. custom sentiment lexicons? [Electronic resource]. - Access mode: https://www.linkedin.com/advice/0/what-advantages-disadvantages-using-pre-trained. - Access data: 06.04.2023.

Cheng Chengliang, master student of department of information technologies in automated systems, BSUIR, cclstrawberry@gmail.com

Scientific supervisor: German Yuliya Olegovna, PhD of technical sciences, associate professor, BSUIR, jgerman@bsuir.by