数据整理

图示

描述已自动生成

1. \*\*Training the Naive Bayes Classifier\*\*:

- The script reads sentences from `pos.txt` and `neg.txt` and processes them.

- An IDF-like preprocessing is applied to the words.

- The sentences are converted into feature sets and used to train a Naive Bayes Classifier.

2. \*\*Sentiment Analysis of Text Files\*\*:

- The script reads text files from the specified directory `data/txt`.

- Each line in these files is processed and its sentiment is classified as either positive (`1`) or negative (`-1`).

- A cumulative sentiment score is calculated for each file.

**Training the naïve bayes classifier的词频云图**



* On the left, the word cloud shows the most frequent positive words found in the positive reviews, based on the predefined list of positive words.
* On the right, the word cloud for the negative reviews displays the most frequent negative words, again filtered by a predefined list of negative words.

**Assignment4-男性-云图**

图片包含 文本

描述已自动生成

**Assignment4-女性-云图**

