


NLP on Financial Statements

REVIEW

CODE REVIEW

HISTORY


Meets Specifications

Your hard work shows in the project, test cases pass for all methods and project meets all specifications. Happy learning and stay udacious 

10-Ks

✓

The function `get_documents` extracts the documents from the text.

The correct regex pattern has been used to extract the text from the documents 

✓


The function `get_document_type` returns the document type lowercased.

The correct regex type pattern has been used and only the 10-k section text is lowercased ✓

Preprocess the Data

✓

The function `lemmatize_words` lemmatizes verbs.

Good job using WordNetLemmatizer to lemmatize verbs in the list of words provided 

Analysis on 10ks


✓

The function `get_bag_of_words` generates a bag of words from documents.

`get_bag_of_words` has been implemented to generate a bag of words that counts the number of sentiment words in each doc

✓

The function `get_jaccard_similarity` calculates the jaccard similarities for neighboring documents.

jaccard similarities for neighboring documents has been correctly calculated 


✓

The function `tfidf` generate TFIDF vectors for each document.

`get_tfidf` has been implemented correctly to generate TFIDF values from documents for a certain sentiment

✓

The function `get_cosine_similarity` calculates the cosine similarities for each neighboring TFIDF vector/document.

cosine similarities for each neighboring TFIDF vector/documents has been correctly calculated 

[Download Project](#)

Rate this review

