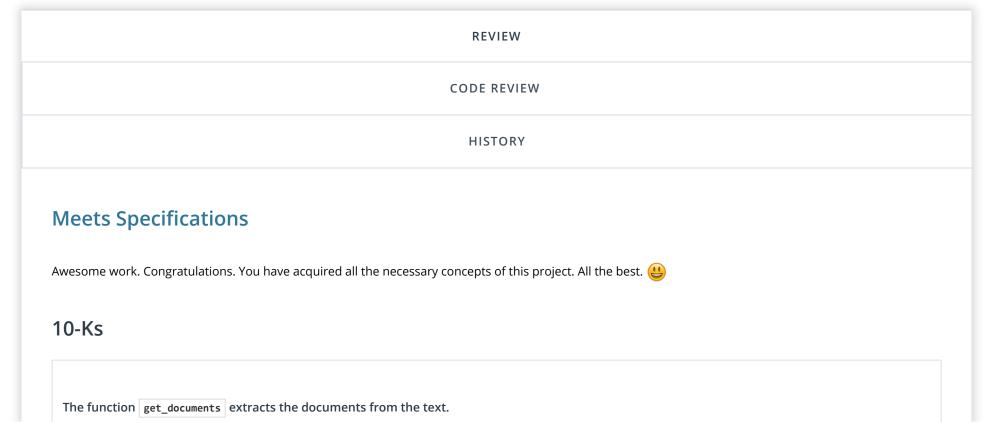


< Return to Classroom

DISCUSS ON STUDENT HUB

NLP on Financial Statements



Right implementation.
The function get_document_type returns the document type lowercased.
Nice and succinct code. 👍

Preprocess the Data

The function lemmatize_words lemmatizes verbs. Correct use of WordNetLemmatizer.

Analysis on 10ks

The function get_bag_of_words generates a bag of words from documents. Correct usage of CountVectorizer.

The function <code>get_jaccard_similarity</code> calculates the jaccard similarities for neighboring documents. Good implementation!

The function tfidf generate TFIDF vectors for each document. Awesome 4The function get_cosine_similarity calculates the cosine similarities for each neighboring TFIDF vector/document. Perfectly implemented 👋 **▶** DOWNLOAD PROJECT

RETURN TO PATH

Rate this review

START