Book Recommender System

DOCUMENTATION

Dataset : https://www.kaggle.com/zygmunt/goodbooks-10k

AIM: BOOK RECOMMENDATION ENGINE
Step1 : Import the dataset named books.csv to jupyter notebook for analysing and making recommendations
Step2 : data cleansing and preprocessing process are done(removed missing values from the dataset)
Step3 : visulaised the behaviour of data and info of data regrading the data structure type of columns and all.
Step 4: visulaised most popular books for the ease of recommending.
Step 5 : using NLP properties clean the dataset for cleaning the text

NLP is a popular library which is used for the Natural language processing,NLP consists of so many other varieties of libraries,packages for the sake of TEXT processing.
Eg: NLTK,Stopwords removing.

Step 6 : As we know that if convert the text to numeric format then the machine will easily categorizes the titles of books.So that we need to perform TFIDF(term frequency-inverse document frequency) vectorizer.it's widely used for converting words to vector and we can easily make array for each words.Now each word will represent as vectors(eg: good=0.324,bad=0.55,book=0.87)

Equations TF=Number of times term t appears in a document) / (Total number of terms in the document
IDF(t) = log_e(Total number of documents / Number of documents with term t in it).

Step 7:Cosine similarity
Cosine similarity is used to find the most similar number (words which are converted to numericals)corresponding to each movie name.

OUTPUT:

```
books7 = get_recommendations_books('harry potter and the goblet of fire', cosine_sim_author)
author_book_shows(books7)
```

```
harry potter and the order of the phoenix
harry potter and the chamber of secrets
harry potter and the goblet of fire
harry potter and the deathly hallows
harry potter and the halfblood prince
harry potter boxed set books
harry potter and the prisoner of azkaban
the casual vacancy
the tales of beedle the bard
complete harry potter boxed set
```