Mahir Efe KAYA
2016400195

ASSIGNMENT 3

(i) Describe the model you used to encode the genres of the books:

- All the data retrieval functions and the tasks is stored in the reader.py.

- I found the regarding parts of the context with python's str's find. I found the starting and ending positions of the related parts to get the text from the whole context. Sometimes I had to do more than one find to seperate a particular part. There was a case of having a description too long, it had to store the description text in a seperate block. I have controlled these exceptions with controlling outer block's scope.

- Find genres was quite easy I have to find the outer block's scope and take the all the elements in the list.

- I used tf-idf model to encode the genres, in which the number of the users recommending a certain degree was the term frequency. Similar done for idfs.

(ii) Describe the model parameters (minimum/maximum thresholds, number of terms, weight variants, α, etc.)

- I used the alpha value(α) to find the best value for the combination of the genres' tfidf and description's tfidf.
- To find the best value for alpha, I started from 0 and incremented it by 0.05