# Question Generation for Discussion Facilitation

**Divyanshu Bhartiya**
IIT Kanpur
divbhar@iitk.ac.in

**Radhika Kedia**
DA-IICT
radhika_kedia@daiict.ac.in

**Ashudeep Singh**
IIT Kanpur
ashudeep@iitk.ac.in

**Biman Gujral**
DA-IICT
biman_gujral@daiict.ac.in

## 1 Motivation

In contrast to IRE approach (Initiate Response Evaluate) , where teacher asks a question, the student responds to it and the teacher judges it on the spot. This method of education doesn't allow long retention of knowledge and the student is not able to understand the concept behind the facts. However , discussion on the other end facilitates long retention and sharing of ideas. Also students learn to communicate and construct explanations and learn the ability to reason. Hence we aim to encourage discussion through question generation . Hence , we aim at building questions that can help us achieve brain storming and healthy discussion among students. It can also be used as a self-tutoring system. Also the tool can be used by the teachers to automatically generate questions for evaluation purposes.

## 2 Technical Details

The tool will take a chapter and a summary of the respective chapter as input and will output a list of subjective questions. We divided this task into two phases : Sentence extraction and Question Generation.

### 2.1 Sentence Extraction

This module extracts "important" and "abstract" sentences from the summary with respect to the chapter. We used summary for extraction purposes because chapter itself contains a lot of data and one sentence may not represent a good abstraction of the chapter. The individual sentences of the chapter are more factual rather than conceptual. To verify our reasoning we used Jaccard Coefficient to find the similarity between each sentence of the chapter with the whole document. Though there were a number of spikes , the selected sentences weren't subjective and couldn't essentially trigger discussion.
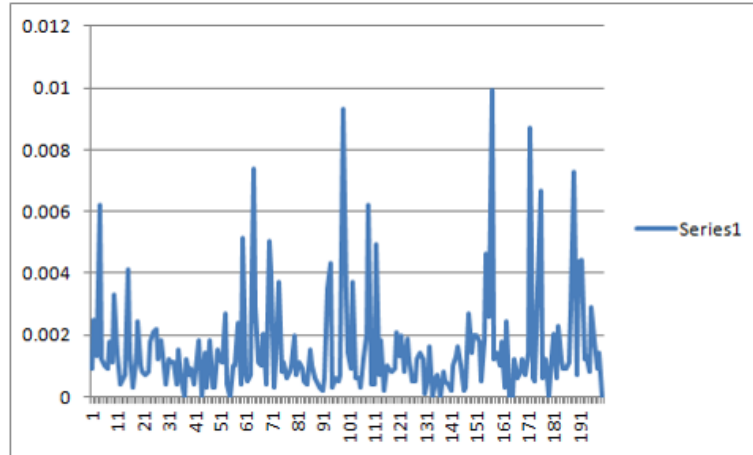
Figure 1: Jaccard Scores applied on Chapter 1 from The Tipping Point

The summary being a concise and abstract representation of the chapter can provide sentences which are possible candidates for generating questions that encourage good discussion. Hence , we aim at selecting sentences from summary which have the highest measure of similarity from the chapter. For this we used several textual features:

## 2.2 LSA

Latent Semantic Analysis analyzes relationships between terms and documents. Each sentence is considered a separate document. The LSA algorithm assumes the documents to be divided into 'k' topics. It analyzes the frequency of terms in the documents and returns a vector for every term and every document , the component of vector telling the similarity of the document with that topic. The motivation behind using LSA has been the thought that a chapter may contain several well defined topics in its contents. For example a scientific literature contains explicit topics like technologies, experiments, results etc. Keeping in mind that a chapter may have such clear defined topics, we thought of using LSA since it may be able to return a vector for each sentence of the summary with its weight in each topic. Selecting few highest weight sentence from each category can give an overall good amount of information.

## 2.3 Norm

It may also be the case that instead of sentence representing a chapter , a sentence may contain a lot of information . In other words , a sentence may reflect all the categories and may not lie specifically in a particular category. The norm calculated over the vectors obtained from applying LSA may represent the amount of total information. The information in a sentence is represented by the length of vector . This may be the case in a fictional literature where there are no well defined categories but there may be sentences which are worthy of question formation.

2

## 2.4 Tf-Idf

Term Frequency-Inverse Document Frequency is used to measure the importance of a word . This feature is used in case if the chapter contains some words which are important and should be known by students. The weight of term should be a good measure in biological texts where biological terms are given importance and should be grasped by students. Tf-Idf finds the weight of words from their occurrences in the summary multiplied by a factor that discounts its appearances in the whole chapter. This feature is used keeping in mind the basic structure of certain type of texts where some words and terms are deemed important.

## 2.5 Jaccard Coefficient

This feature measures similarity between documents as intersection divided by the union. It considers the term frequencies of each document and uses the tf vectors as sets .It computes the sum of weights of shared terms divided by the weight of terms in either document. This feature is used to select sentences which have a high similarity/resemblance to the chapter. Since the selected sentences give a broader picture of the chapter due to their high Jaccard score , these sentences are apt for question generation. In short jaccard score tries to find sentences that are best abstraction of chapter by treating them as sets.

## 2.6 Cosine Similarity

This feature is used within summary itself. The cosine distance finds the angle between two vectors , the more the score, the more the similarity. In other words, the cosine distance is calculated as a dot product of vectors divided by the product of magnitudes. This feature is used in a summary thinking of the fact that a summary itself contain some flow of information which is carried on through the course of summary. The cosine distance is calculated between a sentence of the summary and the sentence above it , over the course of summary. The sentences with higher cosine distance may be said to contain the information carried from before the summary and hence , the selected sentence is more informative and hence should be selected.

# 3 Question Generation

Once the sentences are selected we try to generate subjective questions from the sentences. We use the Heilman's tool to generate question and then make our modifications to make the questions subjective.

## 3.1 Heilman's Tool

- Generates direct and objective questions, factual and reading purpose questions

- Uses overgenerating transformations , that is a sentence is repeatedly tried to generate redundant questions

## 3.2 Modifications to the tool

- We introduced "why" and opinion based questions to be constructed from the sentences extracted. "why" questions were framed from recognizing the term "because" and "due to" and doing necessary transformations to make it a question. Opinion based questions were framed from sentences where the author conveys some information. This was implemented by figuring out author's name and suitable verbs with it like "states", "argues", etc.

- We changed the questions which have yes/no answers to reasoning questions, hence promoting discussion by giving them an incentive to discuss about it.

- Addition of clauses like "explain" or "discuss in detail" was done by analyzing the requisite type of questions and adding them at end of questions. Adding them may lead the students to go deep into the concepts and ideas of chapter.

- Adding the why clause to questions which contain person's name as answer leads to know more about the author or the characters.

## 3.3 Subjectivity Analysis

To rank the questions on the basis of abstraction and ability to trigger discussion, we introduced the concept of subjectivity. Subjectivity is calculated by measuring the subjectivity score of each word and normalizing it over the sentence. To measure the subjectivity score of each word, we used the Senti Word-Net database. Senti Word-Net is a corpus of words which have a positivity and negativity score assigned to each of them grouped by parts of speech. The objectivity score as mentioned with it was $1 - (positivity + negativity)$. Hence the subjectivity is $(positivity + negativity)$. The subjectivity score has been then used to rank the questions generated.