Automatic Text Summarization

Why?

As a person that spends a lot of time reading various things on the internet, when I started searching for ideas about Information Retrieval and Search Engines this project got my attention immediately. If this project works the way I expect it to be, I am sure I will use it for my own benefit too.

What?

Text Summarization is a pretty known thing under Information Retrieval concepts. Google for example uses a kind of Document Summarization to write the text that we see under the links when we search something. The objective is summarizing a text, which shortens the length but the information stays the same.

How?

The text summarization method I will use will be "Extraction", which basically picks the key points from the text, which is picking the related sentence without modifying it, thus forming a shorter paragraph. Various algorithms will be used for this and comparisons will be made.

When?

Here is the timeline I propose:
(Note that this week is Week 7)
- Week 8, 9: (10 April – 21 April)
Extracting the dataset I want to use which is several Q&A's from
https://www.reddit.com/r/askscience/
- Week 10, 11, 12: (24 April – 12 May)
Learning and implementing various (at least 3) algorithms for text summarization.
- Week 13, 14: (15 May – 26 May)
Create the UI prior to presentation.
- Week 15: (29 May – 2 June)
This is the presentation week

Erhan Tezcan
14011062