AUTOMATIC MEETING SUMMARIZATION

SYNOPSIS

GROUP MEMBERS:

Dhruv Pai (ENG17CS0069)
Meghana V (ENG17CS0125)
Muppalla Bhargavi (ENG17CS0134)
Muskan Asmath (ENG17CS0135)

INTRODUCTION :

Summarization is the task of condensing a piece of text to a shorter version, reducing the size of the initial text while at the same time preserving key informational elements and the meaning of content. Since manual text summarization is a time expensive and generally laborious task, the automatization of the task is gaining increasing popularity and therefore constitutes a strong motivation for academic research.

PROBLEM STATEMENT :

A web application that summaries a meeting using a NLP algorithm and emails the meeting summary.

OPERATIONAL FUNCTIONS:

The operational functions of automatic text summarization are:

- Taking a record of a meeting and converting the speech into text.
- Summarizing the text through the model.
- Summarized text should be mailed to the participants in the form of PDF.

ABSTRACT :

A Machine learning-based project that uses NLP algorithm. The interactive Web application takes a recording of a meeting and converts it to text by using the speech recognition api which will be temporarily stored in a MongoDB database . The text summarization model then summarizes the text fed to it by using a NLP algorithm using Python as backend and serves using the Flask micro framework. Using a web driver called selenium the summary will be mailed using a bot gmail account thereby automating the procedure, the summary will be mailed in form of a PDF using a specific format to streamline and automate the process.

SCOPE OF THIS PROJECT:

The project intends to eliminate the hectic job of making notes of a meeting for future reference and automate the process thereby reducing human error.

EXISTING SYSTEMS:

The current systems that exist in the market are mostly text summarizers and do not include features such as voice to text or even the automated email facility. Also the current system only works for simple web pages i.e. the amount of time it takes to summaries increases exponentially with the length of the document, our model will attempt to solve the problem.

CURRENT ISSUES /PROBLEM IN EXISTING SYSTEM:

The current issues can be jotted down as there is no way of recording meetings when the meeting is actually going on.There are no automatic summarization applications which makes man's work easier. There are no easy interfaces to do the above job.

PROPOSED SOLUTION:

- We can customize the API by adding more options to manipulate the output. ie, summary length, ignoring list text, etc
- Display list of sentences instead of paragraph.
- Create a chrome plugin and highlight the sentences.

HARDWARE AND SOFTWARE REQUIREMENTS:

SOFTWARE:

- FRONT END : Javascript, HTML,CSS
- DATABASE : MongoDB
- BACKEND : Flask,Python 3.7,++
- WEB DRIVER: Selenium using Python 3.7
- LIBRARIES : Tensorflow

HARDWARE:

- Gpu- nvidia gtx 980 or above
- Ram - 4 gb or above
- System - linux/windows(>32)

POSSIBLE OUTCOMES:

Example shown below:

Original Text/URL

"Peter and Elizabeth took a taxi to attend the night party in the city. While in the party, Elizabeth collapsed and was rushed to the hospital. Since she was diagnosed with a brain injury, the doctor told Peter to stay besides her until she gets well. Therefore, Peter stayed with her at the hospital for 3 days without leaving."

Summary

"Peter and Elizabeth attended the night party in the city. While in the party, Elizabeth collapsed and was rushed to the hospital. Therefore, Peter stayed with her at the hospital for 3 days without leaving."