

REVIEW ARTICLE



ISSN: 2321-7758

A COLLECTIVE SURVEY ON AUTOMATIC PPT GENERATION SYSTEM

SOMNATH MALGE¹, ROHAN CHAURASIA¹, KETAN SAWALKAR¹, SUMIT RAUT¹, S. R. PATIL²

¹UG Student, Dept. of Computer Engineering, Savitribai Phule Pune University

²Assistant Professor, Dept. of Computer Engineering, Savitribai Phule Pune University

somnathmalge1001@gmail.com; rehu93@gmail.com; ketansawalkar4761@gmail.com; sumit.sraut1@gmail.com; srp.sit@sinhgad.edu

ABSTRACT

Presentation plays an effective role for sharing some information. Presenting PPT is easiest and successful way to communicate with people for sharing information. But generating such PPT is very time consuming process. In this paper we emphasize on generating PPT slides automatically. This paper presents new ways to generate automatic PPT by using some methodologies. The different techniques available for generating automatic PPT are NLP, text mining, feature extraction and Fuzzy based classification techniques. The most of the above mentioned methodologies have performance issues according space and time. This paper concentrating on analyzing techniques and to find a proper way to improve to generate PPT automatically.

Keywords— Natural Language Processing, Text mining, Fuzzy Logic.

©KY PUBLICATIONS



SOMNATH MALGE



ROHAN CHAURASIA



KETAN SAWALKAR



SUMIT RAUT

I. INTRODUCTION

The most effective way of conveying information to audience is power point presentation. Several times they are referred for academic purpose, marketing analysis and resource analysis. Traditional software like MS PowerPoint are being used for formatting of the slides, but not in the data. This generating such PPT is the time consuming process. To avoid this time consumption active research has been done.

This paper presents a framework to generate a PPT not only with the formatting of slides but also in the content. Every individual presenter has unique way of presenting slides

which contain sections like abstract, introduction, related work, experiment and conclusion. Our aim is to generate draft slides which helps people to generate PPT automatically. This leads in less time complexity and human efforts.

Techniques like NLP, text mining, feature extraction and fuzzy logic helps to make the slides more structured ordering sequence of part where each part addresses a specific topic relevant to each other. NLP is a way of programming in terms of natural language. NLP document is a structured document which contents sections and subsections for proper explanation of sentences. It is for making interactive fiction. In statement in NLP the smallest

unit is sentence. In the underlying ontology every sentence is stated in terms of concept. NLP is a top down method of writing software which is based on ontology. An ontology is a concept which is needed to describe task in the topic addressed. This ontology defines the data structures that NLP can use in sentences.

Fuzzy Logic is a method which is based on degree of truth instead of true or false. It is based on Boolean logic on which the modern computer is based. It includes 0 and 1 as extreme cases truth. It also includes the different state of truth between 0 and 1. The working of Fuzzy Logic is similar to our brain work where we aggregate data and form a number of partial truth. A similar kind of process used in artificial intelligence.

II. LITERATURE SURVEY

In paper author describes comparison between techniques for automatic generation of Presentation slides. It depend on knowledge of information extraction. Extraction summarizes can build a summary and pick up the important sentences from an articles or data. So that important sentences should be necessary accurate. The summarization strategy can be fully work on statistical method .The summarization data exploit by using citation information. This extraction summery build an important modification article. Machine learning focuses on the development of computer programs that can teach themselves to grow and change when exposed to new data. [1]

In paper author proposed method for automatic generation at slide GDA (global document annotated) target as XML target allow the semantic input a raw document under the current state of the art, GDA tagging can be only calls for manual correction. Annotated document is generic information content from presentation, summarization, information retrieval and information extraction. [2]

Describes a method of TF-IDF in which it describe a simple scoring mechanism. Which is totally based on the number of matched terms in generating the slides automatically. This methodology assigns bullet points to the paragraph or lines and one main disadvantages for this

technique is that it focuses only on either alignment of the sentences or the paragraphs. [3]

Proposes a method by M. Sravanthi, C. R. Chowdary and P. S. Kumar that provides automatic generation of slides from given latex document. In this methodology it first converts latex document to XML format which is further used to generate a slides using query specific extractive summarizer. This method is efficient for generating important points along with necessary figures, tables and graphs from input paper. [4]

Elaborates a method by Ektaa G. Meshram, D. A. Phalke that provides automatic PPT generation by data mining technique, which will help in scoring sentences and as well as generating slides with graphical elements. Natural Language Processing is used for scoring sentences for extracting of key phrases Open NLP Library is used to extract images from whole document with the help of PDF box and Documents. The main advantage is that it not only focuses on text only but graphical elements also. [5]

Describes a novel system called PPs Gen to generate presentation slides from academic papers. Presentation slides are made by bullet points and corresponding sentence selection of sentence done by summarization. The most important sentences are selected using ROUGE tool and Pyramid appraisal. Natural Linear Processing model is used to generate well organized slides by selecting and aligning key phrases and sentences. [6]

Proposes a method by Amritha S Dev and Mrs. Amina Beevi A which uses scoring model of sentences. It is based on SVM. For extracting key phrases and sentences it uses integer language programming. Local and Global phrases are identified by ILP whereas SVM finds the score of the sentence. Summarization is a process similar to data mining where important sentences are detected. It is a process where important information is extracted from text. There are two ways of summarization 1.Extractive Summarization system 2.Abstractive summarization system. [7]

Proposes a method by Tomohide Shibata and Sadao Kurohashi which refers summarization based on citation, to create summarized scientific

papers. This method helps in optimizing the data such as removal of unimportant sentences. The citation of sentences of the target paper are solved by address. Efficient summaries are produced compared to other systems. In citation the important parts are highlighted from the cited paper such as methods, advantages, limitations, drawbacks. Many sentences are cited and by collecting these cited sentences together we get a huge amount of information. [8]

IV. CONCLUSION

Preparation of power point presentation from given document takes much more manual efforts. To ease this process many automatic power point presentation systems are existed as analyzed by this this paper. After analyzing many prior works this paper comes to a conclusion that most of the input documents are narrated in many different ways. So this paper feels natural language processing is the good solution for dealing with the document text to extract power point presentation slide contents.

References

- [1] Biju P. Dais, Smitha C.S., "A Comparison on Techniques for Automatic Generation of Presentation Slides", International Journal of Computer Applications (0975 - 8887) Volume 131 - No.17, December 2015.
- [2] UTIYAMA Masao, HASIDA Koiti, "Automatic Slide Presentation from Semantically Annotated Documents".
- [3] Parvez Shaikh, "PPS Gen: Learning-Based Presentation Slides Generation for Academic Papers", IOSR Journal of Computer Engineering (IOSR-JCE) E-ISSN: 2278-0661, p-ISSN: 2278-8727PP 62-64.
- [4] M. Sravanthi, C. Ravindranath Chowdary, P. Sreenivasa Kumar, "SlidesGen: Automatic Generation of Presentation Slides for a Technical Paper Using Summarization", Proceedings of the Twenty-Second International FLAIRS Conference (2009).
- [5] Ektaa G. Meshram, D. A. Phalke, "Technique for Generating Automatic Slides on the basis of Paper Structure Analysis", International Journal of Innovative Research in Science, Engineering and Technology (An ISO 3297:2007 Certified Organization) Vol. 5, Issue 6, June 2016.
- [6] Ektaa G. Meshram, Mrs. D. A. Phalke, "Survey on Presentation Slides Generation for Academic Papers", Ektaa G Meshram et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (1) , 2016, 22-25. Ektaa G. Meshram, Mrs. D. A. Phalke, "Survey on Presentation Slides Generation for Academic Papers", Ektaa G Meshram et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (1) , 2016, 22-25.
- [7] Amritha S. Dev, Mrs. Amina Beevi A," Survey: A Learning Based Slide Generation from Academic Paper", International Journal of Engineering Research ISSN: 2319-6890 (online), 2347-5013(print) Volume No.5, Issue No.6, pp.: 467-468, 1 June 2016.
- [8] Tomohide Shibata and Sadao Kurohashi, "Automatic Slide Generation Based on Discourse Structure Analysis", R. Dale et al. (Eds.): IJCNLP 2005, LNAI 3651