

# A Provision System for making Presentation Slides form documents

**E.Poonguzhali<sup>1</sup>, M.Vasanth<sup>2</sup>, S.Brindha<sup>3</sup>**

<sup>1</sup>*Research Scholar, St. Peter's University, Chennai.*

<sup>2,3</sup>*Asst. Prof, Dept. of Computer Science & Applications, St. Peter's University, Chennai.*

*kanmani93ek@gmail.com*

**Abstract:** *In this paper, analyze a real fascinating task of always producing presentation slides for scholastic papers. The produced presentation slides can be utilized as drafts to offer the moderators some assistance with preparing their formal slides quickly. A novel framework called PPSGen is proposed to address this assignment. It first utilizes the relapse technique to take in the significance scores of the sentences in a scholarly paper, and after that endeavors the whole number straight programming (ILP) strategy to create all around organized slides by selecting and adjusting key expressions and sentences. Assessment results on a test set of 200 sets of papers and slides gathered on the web exhibit that our proposed PPSGen framework can produce slides with better quality. A client study is additionally shown to demonstrate that PPSGen has a couple of clear preferences over pattern techniques. And also we are enhancing the graphical elements into slides.*

**Keywords:** *Authentication; access control, ILP*

## **I. INTRODUCTION**

Presentation slides have been a prevalent and compelling intends to present and exchange data, particularly in scholarly gatherings. The scientists dependably make utilization of slides to show their work pictorially on the gatherings. There are numerous virtual products, for example, Microsoft Power- Point and Open Office to offer analysts some assistance with preparing their slides. Be that as it may, these instruments just help them in the arranging of the slides, yet not in the substance. Regardless it takes moderators much time to compose the slides starting with no outside help. In this work, we propose a strategy for consequently creating presentation slides for scholastic papers. We expect to naturally create very much organized slides and give such draft slides as a premise to decrease the moderators' chance and exertion while setting up their last presentation slides.

Scholastic papers dependably have a comparative structure. They for the most part contain a few areas like dynamic, presentation, related work, proposed strategy, examinations and conclusions. Every area is adjusted to one or more slides and one slide for the most part has a title and a few sentences. In this study, we propose a novel framework called PPSGen to create very much organized presentation slides for scholastic papers. In our framework, the significance of every sentence in a

paper is found out by utilizing the bolster vector relapse (SVR) model with various valuable elements, and after that the presentation slides for the paper are produced by utilizing the whole number straight programming (ILP) model with extravagantly planned target capacity and imperatives to choose and adjust key expressions and sentences. Whatever remains of this paper is sorted out as takes after. Related work is presented in Section 2. We depict our technique in subtle element in Section 3. We demonstrate the examination results in Area 4 and finish up our work in Section 5.

## II. MATERIAL AND METHODS

### Slides Generation

Programmed slides era for scholastic papers stays far under-examined these days. Few concentrates specifically inquire about on the theme of programmed slides era. Utiyama and Hasida [1] endeavored to naturally produce slides from information archives explained with the GDA tagset. GDA labeling can be utilized to encode semantic structure. The semantic relations incorporate linguistic relations. Yasumura et al. [2] presented an emotionally supportive network for making slides from specialized papers. The inputs of the framework are scholarly papers in LATEX design. The framework computes the weights of the terms in the paper utilizing TFIDF scores.

Shibata and Kurohashi [3] proposed a technique to naturally produce slides from crude writings. Conditions and sentences are considered as talk units and soundness relations between the units, for example, list, contrast, topicchaining also, cause are distinguished. Hayama et al. [4], Kan [5] and concentrated on the issue of adjusting specialized papers and presentation slides. Hayama et al. utilized a variety of the Hidden Markov Model (HMM) to adjust the content in the slides to the doubtlessly segment in the paper, which likewise utilized the extra data of titles and position holes. Kan [5] connected an altered most extreme likeness strategy to do the monotonic arrangements also, prepared a classifier to distinguish slides which ought not be adjusted.

#### *INPUT*

1. TRAININGSET  $\{ x_i, y_i, i=1..l \}$
2. WHEIGHTS  $q_i, i=1..l$
3. BIAS  $b$
4. TRAININGSET PARTITION INTO SUPPOTSET(S) , ERRORSET(E)  
AND REMAININGSET(R)
5. PARAMS:  $e, C$ , KERNELTYPE AND KERNELPARAM
6. R MATRIX
7. SAMPLE TO REMOVE INDEX ( $c$ )

Pseudo Code.1 SVR

A inquiry particular extractive summarizer QueSTS is utilized to remove sentences from the content in the paper to produce slides. Journeys exchanges the information content to an incorporated chart (IG) where a sentence speaks to a hub and edges exist between the hubs that the sentences comparing to them are comparable. The weights of the edges are computed as cosine likeness between the sentences. More subtle elements can be found in [10].lines from the end of the content piece.



**Fig.1 System Design**

### **Scientific Article Summarization**

The objective of exploratory article rundown is to create a short outline for a given logical article or article set. Early works including [11], [12], [13] attempted to utilize different elements particular to exploratory content (e.g., logical intimations highlights). Reference data has as of now demonstrated its viability for synopsis of the experimental articles..

### **Document Summarization**

The errand of report synopsis means to create a short synopsis for a given report or record set. Different techniques have been proposed for report synopsis, including guideline based strategies diagram based techniques learning-based strategies ILP-based techniques and so forth. As of late bolster vector relapse and ILP have been utilized generally as a part of the errand of rundown. Ouyang et al. [32] and Galanis and Malakasiotis [33] utilized SVR to prepare also, take in the sentence significance score. McDonald [34] proposed the primary ILP strategy for outline. It developed outlines by expanding the significance of the chosen sentences and minimizing their pair wise comparability.

### **CORPUS**

In our work, we plan to naturally create presentation slides for scholastic papers. We have to eate all around organized slides as the draft slides for a moderator to set up the last slides. There are different sorts of slides which are made by Microsoft Power Point and Open Office. They can be much diverse in styles and we clearly can't consider all sorts of styles.

A fledgling for the most part gets ready slides which are successively adjusted to the paper. One segment in the paper is for the most part adjusted to one or more slides. One slide for the most part incorporates a few visual cues and sentences that clarify the comparing visual cues. From Fig. 1, we can have a look at the style of the slides we produce. Here, key expressions "Specialists Behaviors" what's more, "Requirement Approach" are set as the slug focuses.

In this work, we just consider the content components in the paper. Different components, for example, tables and figures are definitely not incorporated into the created slides. In spite of the fact that tables and figures are helpful in the slides, we disregard them to rearrange the issue also, better concentrate on the era of the content components.

### **III.OUR PROPOSED METHOD**

#### **Overview**

In this paper, we propose a framework to naturally create slides that have great structure and substance quality from scholastic papers. The design of our framework is appeared in Fig. 2. We utilize the SVR-based sentence scoring model to dole out a significance score for every sentence in the given paper, where the SVR model is prepared on a corpus gathered on the web. At that point, we produce slides from the given paper by utilizing ILP.

#### **Sentence Importance Assessment**

In our proposed PPSGen framework, sentence significance evaluation is one of the two key steps, which intends to dole out a significance score to every sentence in the given paper. The score of every sentence will be utilized as a part of the slides era process. In this study, we present a couple of valuable elements what's more, propose to utilize the bolster vector relapse model to accomplish this objective.

#### **Support Vector Regression Model**

For our situation,  $u$  is the element vector of the sentence and  $y$  is the significance score of the sentence. We utilize LIBSVM [4] with the RBF portion to actualize the SVR model. We have to anticipate the significance score of every sentence for sentence choice in slides era. The motivation behind why we utilize the SVR model rather than the arrangement model is that the relapse score is better to be utilized for sentence choice than the coarse parallel classification.

#### **Training Data Construction and Model Learning**

To build preparing information in view of the paper-slides sets, we apply a likeness scoring strategy to dole out the significance scores to the sentences in a paper. The primary speculation is that the sentences in the slides ought to speak to the substance of the comparing paper.

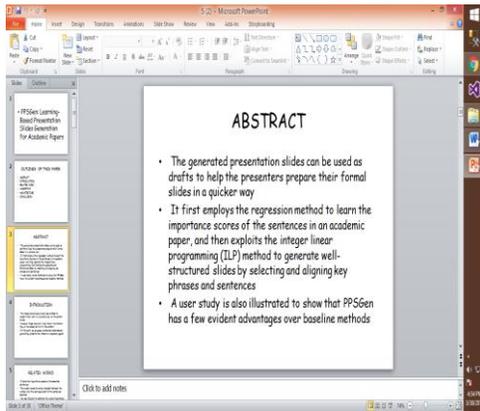
A sentence with a higher most extreme similitude is more like one sentence in the writer composed slides. Since the writer composed slides contain the sentences that human creators

considered most essential, a sentence with a higher score is well on the way to be critical, as well. We embrace the most extreme closeness rather than the generally comparability with every one of the sentences in the slides or the normal similitude with every sentence in the slides. The sentences in a particular area ought to be more like the relating part in the slides and less like the other parts. Hence, it is more sensible to utilize the most extreme similitude to dole out the significance scores of the sentences. Where  $s$  is a sentence in the paper, the arrangement of the sentences in the comparing. The standard cosine measure is utilized as the comparability capacity so the sentence's significance score is set as the most extreme comparability between the sentence and any sentence in the relating slides.

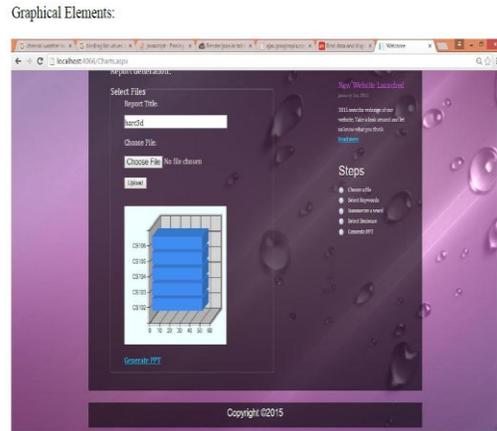
### **Slides Generation**

In the wake of getting the anticipated significance score for every sentence in the given paper, we misuse the whole number straight programming strategy to produce all around organized slides by selecting and adjusting key expressions and sentences. Not at all like those techniques [1], [2], that produce slides by essentially selecting vital sentences and setting sentences on the slides, we select both key expressions and sentences to build very much organized slides. We utilize key expressions as the visual cues and sentences applicable to the expressions are put underneath the visual cues. We characterize two sorts of expressions: worldwide expressions and neighborhood phrases. Any remarkable expression in an article is a worldwide expression; furthermore, a neighborhood expression implies a worldwide expression in a specific segment. For instance, "SVR" is a worldwide expression of this paper, while its appearances in changed areas are considered distinctive neighborhood phrases. "SVR {Introduction}" and "SVR {Our Proposed Method}" indicate diverse nearby expressions, also, they speak to the appearances of "SVR" in Sections 1 also, 4 of this paper, separately. So a worldwide expression that shows up in various segments can compare to a couple of nearby phrases. Since an essential expression is constantly utilized as a part of numerous distinctive segments, a worldwide expression that compares to additional nearby expressions ought to be respected to be more vital and more inclined to be chosen. Subsequently, we utilize the neighborhood phrases to create the visual cues straightforwardly for various segments also, utilize the worldwide expressions to address the significance contrasts between various interesting expressions. .

The aggregate length of the sentences chose is altered. So if the model has a tendency to choose the more drawn out sentences, the less sentences can be chosen. The model needs to make a tradeoff between the number and the normal length of the sentences chose. This issue is tended to in [39]. Galanis et al. [39] demonstrate the adaptation of the ILP model that does not consider the length of the sentences deteriorates execution. So we consider the length of the sentences in our model. We utilize this part to appraise the general significance of the produced slides. The second part augments the aggregate numbers of the bigrams in the paper which likewise show up in the slides. At the point when more exceptional bigrams are displayed in the slides, the sentences in the slides are less repetitive what's more, the slides can be more assorted. Furthermore, when more critical bigrams are incorporated into the slides, the sentences chose are more critical and slides can improve quality.



**Fig.2 Slides**



**Fig.3. Graphical Elements**

The last part means to amplify the weighted scope of the key expressions chose and figure out which expressions ought to be chosen. We say a sentence is secured by an expression when this sentence contains the phrase. Fantastic slides ought to cover the substance in the paper however much as could reasonably be expected. As indicated by our characterized kind of slides which incorporate expressions and sentences beneath the expressions, we ought to choose the phrases that are significant to more sentences. We consider not just the tally of the sentences secured by the chose expresses additionally the general significance score of the secured sentences.

### **Experimental Results**

Additionally, a client study is likewise performed to subjectively assess the slides created by various strategies. We arbitrarily select 20 papers in the test set and utilize TF-IDF strategy, MEAD and our technique to produce slides. Arbitrary Walk and C-Lexrank are skipped on the grounds that they are additionally rundown techniques and get lower or comparable exhibitions as MEAD.

### **IV. CONCLUSIONS AND FUTUREWORK**

This paper proposes a novel framework called PPSGen to produce presentation slides from scholarly papers. We prepare a sentence scoring model taking into account SVR and utilize the ILP strategy to adjust and extricate key expressions and sentences for creating the slides. Test results demonstrate that our strategy can produce much preferable slides over conventional strategies.

In future work, we will enhance our framework by utilizing both content and graphical components in the paper and make slides more understandable and striking. At the point when managing the graphical components, we have to recognize the graphical components in the paper first. The relationship between the content components and the graphical components additionally should be recognized. We have to know which sentences are most applicable to a graphical component and which graphical components ought to be chosen to create the slides. We can utilize run based

strategies or machine learning based techniques to fathom the above issues. At that point we can essentially append the tables and figures we select to the most important sentences in the slides.

## REFERENCES

- [1]M. Utiyama and K. Hasida, "Automatic slide presentation from semantically annotated documents," in Proc. ACL Workshop Conf. Its Appl., 1999, pp. 25–30.
- [2]Y. Yasumura, M. Takeichi, and K. Nitta, "A support system for making presentation slides," Trans. Japanese Soc. Artif. Intell., vol. 18, pp. 212–220, 2003.
- [3]T. Shibata and S. Kurohashi, "Automatic slide generation based on discourse structure analysis," in Proc. Int. Joint Conf. Natural Lang. Process., 2005, pp. 754–766.
- [4]T. Hayama, H. Nanba, and S. Kunifuji, "Alignment between a technical paper and presentation sheets using hidden Markov model," in Proc. Int. Conf. Active Media Technol., 2005, pp. 102–106.
- [5]M.Y. Kan, "SlideSeer: A digital library of aligned document and presentation pairs," in Proc. 7th ACM/IEEE-CS Joint Conf. Digit. Libraries, Jun. 2006, pp. 81–90.
- [6]B. Beamer and R. Girju, "Investigating automatic alignment methods for slide generation from academic papers," in Proc. 13th Conf. Comput. Natural Lang. Learn., Jun. 2009, pp. 111–119.
- [7]S. M. A. Masum, M. Ishizuka, and M. T. Islam, "Auto-presentation: A multi-agent system for building automatic multi-modal presentation of a topic from world wide web information," in Proc. IEEE/WIC/ACM Int. Conf. Intell. Agent Technol., 2005, pp. 246–249.