Identifying Keyword Predictors in Lecture Video Screen Text

Farah Naz Chowdhury Raga Shalini Koka Mohammad Rajiur Rahman Thamar Solorio Jaspal Subhlok
Department of Computer Science, University of Houston, Houston, TX 77204
Email: fchowdhury4@uh.edu, ragasalini@gmail.com, mrahman13@uh.edu, tsolorio@uh.edu, jaspal@uh.edu

2021 IEEE International Symposium on Multimedia (ISM) | 978-1-6654-3734-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISM52913.2021.00055

***Abstract*—Automatic discovery of keywords for lecture video segments is an important component of advanced navigation systems for lecture videos. The suitability of a word or a short phrase to be a keyword depends on various factors, including the frequency in a segment, relative frequency in reference to the full video, font size, time on screen, and the existence in domain and language dictionaries. The research presented in this paper provides a refined understanding of how various factors contribute to predicting keywords based on logistic regression analysis. The analysis employs a real-world dataset consisting of lecture videos from Biology, Computer Science, and Chemistry, hosted on Videopoints, a lecture video management portal. Term frequency, maximum font size, and presence in a domain dictionary were identified as the most important predictors of keywords. The results provide a scientific foundation and valuable insights into the design of future keyword prediction systems.**

I. Introduction

Lecture videos are widely employed as the core medium for online learning, and as a supplementary tool for traditional face-to-face learning. Students value videos for allowing them to review at their own pace and studies report a positive impact on grades and overall course satisfaction [1], [2].

Quickly navigating to the content of interest can be challenging with the lecture video format. An innovative approach to improving navigation is to automatically divide the lecture video into semantically cohesive segments, and then automatically generate keywords (or tags) and visual summaries for the segments. This approach has been implemented in Videopoints [2]–[6], a lecture video player developed to support research in lecture video content retrieval. Content-based indexing and summarization in Videopoints is illustrated in Figure 1.

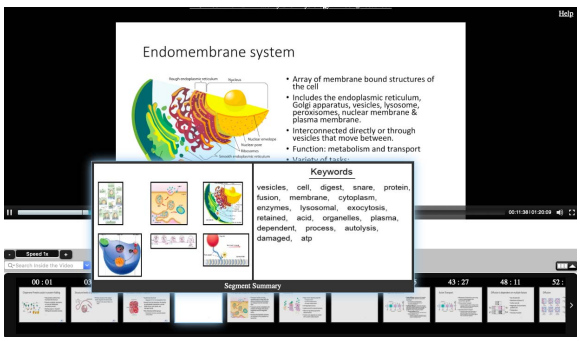

Fig. 1: Videopoints player showing topical indexing, keywords, and visual summaries

Keywords extracted from lecture video segments can provide an overview of the content discussed in each segment and improve navigation. The central goal of the research presented in this paper is to develop a foundation for automatically generating keywords for a lecture video segment based on the screen text in the lecture video. Traditional approaches to keyword selection for text documents are generally not suitable due to the unique nature of text in lecture videos; text is unstructured, contains variable size and style fonts, is displayed for varying duration, has terms with meaning only in a local context, and typographical errors are common.

This work contrasts with recent work on keyword selection in lecture videos based on analyzing transcripts [7], [8]. Screen text is typically based on the instructor's viewgraphs, and hence, is well prepared and focused. Transcripts, on the other hand, are based on improvised speech and not as focused, but the amount of text is plentiful. Related work on lecture video indexing indicates that screen text is likely to lead to greater accuracy than speech transcripts [3].

Automatic generation of keywords by heuristically combining screen text features on lecture videos was presented in [9]. The features considered include term frequency, TF-IDF (product of term frequency and inverse document frequency), font size, time on screen, and presence in domain and English dictionaries. This paper systematically explores the relationship between these features of lecture video screen text and the keywords for that lecture video segment. Ground truth keywords from a set of STEM lecture videos used in [9] were also employed in this analysis. The toolset employed is based on logistic regression analysis which is known to be an effective approach to modeling such scenarios [10].

The steps involved in identifying the factors that influence keyword selection are as follows. Text is extracted from lecture video frames with OCR (Optical Character Recognition) tools. It is then pre-processed by removing irrelevant background text, error correction, stop word removal, and stemming. Next, all potential keywords are identified. Finally, logistic regression analysis is applied to this data set to identify the relationships between input variables and the ground truth keywords.

The scientific contribution of this paper is to determine the extent to which screen text may be a predictor of keywords in a lecture video segment and quantify the importance of different properties of screen text in keyword prediction.

978-1-6654-3734-9/21/$31.00 ©2021 IEEE
DOI 10.1109/ISM52913.2021.00055

Authorized licensed use limited to: University of Houston. Downloaded on January 26,2022 at 21:57:32 UTC from IEEE Xplore. Restrictions apply.

II. Related Work

Video is widely and increasingly used in higher education. Popular commercial lecture video hosting platforms include Echo360 [11], Kaltura [12], and Panopto [13]. Opencast [14] is a community-driven open source software project for producing, managing, and distributing academic video. Keyword search is supported by several lecture video players, with some also supporting semantic search [15], [16]. The concept of content overviews of lecture videos is introduced in virtPresenter [17].

Movie style video summarization is a long-studied problem in multimedia content analysis but not directly relevant to lecture videos. In recent years, a body of research has been developed to analyze lecture videos visually [5], [18]–[21]. The focus of this paper is on textual summarization.

Several projects have addressed the extraction of keywords from long text documents such as news articles using statistical, machine learning, and graph-based approaches [22]–[29]. The most commonly used statistical measure to extract significant words in an unsupervised way is TF-IDF (Term Frequency-Inverse Document Frequency), a measure of relative frequency of a keyword candidate, and its variants [30], [31]. Though a substantial research has been done on extracting keywords, these techniques were applied either on documents or news articles that are well-phrased. For lecture videos, we are utilizing the text extracted from video frames using OCR, and the output given by OCR does not follow the standard structure of the text documents. These approaches do not suffice for educational video lectures as the domain is not fixed and data is mostly in slide format where duration, fonts, and positional information play a vital role.

Recent research has focused on methods for efficient access to educational videos with automatic identification of keywords from videos or their segments [7], [8], [32], [33]. These approaches are based on extracting keywords from the audio transcripts of video lectures. Our experience in a related project indicates that screen text is likely to be a better guide to keywords than a speech transcript [3]. Possible reasons are the colloquial nature of a classroom interaction as well as errors in ASR (Automatic Speech Recognition) transcripts due to a variety of factors like accents, technical vocabulary, and poor recording quality.

This paper builds on a heuristic approach to extracting keywords developed in [9]. The main factors in keyword prediction that we employ, including frequency, font size, and duration of N-grams, were introduced in this work. This paper introduces a rigorous statistical analysis to analyze the significance of different parameters, in contrast to a heuristic approach. Finally, this paper employs binary logistic regression [34], [35] to infer the relationship between a set of predictors and a binary response variable.

III. Predictor Variables

In this Section, we list the factors that, in our experience, are the most likely potential predictors of a word or a phrase to be a keyword. We will refer to every potential keyword as an N-gram, which is simply a phrase with N consecutive words.

Frequency: Clearly, the frequency of occurrence of an N-gram is a potential predictor. Additionally, the relative frequency of a keyword in a specific segment of a lecture video is also potentially relevant. For illustration, in a course on computer networks, the word "network" is likely to occur often throughout. However, it is unlikely to be an interesting keyword for a particular segment. We consider two specific factors:

- *Term frequency(TF):* The number of times an N-gram occurs in a segment.
- *TF-IDF:* where

$$IDF = log(TotalSegments/NgramSegments)$$

where $TotalSegments$ is the total number of segments in a lecture video and $NgramSegments$ is the number of segments where the relevant N-gram occurs.

Font size: Text can occur in varying font sizes on lecture video slides. We hypothesize that font size impacts the likelihood of an N-gram being a keyword. This work considers:

- *Average font size:* To evaluate if the text in larger fonts is more likely to be among the keywords.
- *Maximum font size:* Content of slide titles, typically in the largest font, may be more likely to include keywords. Hence, maximum font size may have a predictive value independent of the average font size.

For the analysis in this paper, font size value is squared and then normalized. The squaring is motivated by the fact that fonts typically vary in a small range in a document, and squared value may be a better indicator of the impact of a larger font in a document.

Time on screen: The amount of time for which text is displayed in a video segment is another consideration. The hypothesis is that the longer an N-gram is displayed, the more likely it is to be a keyword.

External corpus: This study explores if the presence of text in the following external corpora is a factor in its likelihood of being among keywords.

- *Domain dictionary:* Important keywords often have a unique meaning or relevance in a scientific domain. For example, the term *solution* may be worthy of additional consideration in a chemistry lecture.
- *English dictionary:* Most, but not all, words in a lecture video segment are part of the English dictionary. The reasons range from the use of proper nouns, acronyms, and spelling errors in text, or in text extraction.

If any component of an N-gram is in a dictionary, the entire N-gram is considered to be a part of that dictionary.

IV. Experimental Design

The objective of the research presented in this paper is to measure the predictive value of screen text features listed in

Authorized licensed use limited to: University of Houston. Downloaded on January 26,2022 at 21:57:32 UTC from IEEE Xplore. Restrictions apply.

Section III for automatic keyword extraction. Text is extracted from lecture video frames with OCR; specifically MODI (Microsoft Office Document Imaging) tool set was employed. This is followed by pre-processing of this raw input. Logistical regression analysis is then conducted to derive the relationship between the variables and the available ground truth keywords from human experts. The key steps are explained in this Section.

A. Text preprocessing

The text extracted by OCR contains noise and errors. The text contains a large number of common words called stop words. Also, the same root typically occurs in multiple words, such as "network", "networking", and "networks". An analysis that relates frequency to importance often provides nonsensical results without consideration of these factors. Hence, a number of steps including stemming and stop word removal were taken as pre-processing steps to prepare input for keyword analysis. This is common in text analysis. We briefly discuss these pre-processing steps as they are central to meaningful regression analysis. More details are available in [6].

The input to keyword prediction analysis consists of all words identified by OCR on the lecture video frames in the relevant video segment, along with their font size, location on the screen, and duration on the screen. This raw input is processed to generate a set of valid N-grams. Only 1-gram, 2-gram, and 3-grams are considered as virtually all ground truth keywords provided by users consisted of 1 to 3 words. Following is a summary of the key pre-processing steps.

- *Removal of background text:* Most lecture videos have some fixed screen content that is not relevant to the topic of discussion. Examples include a course title, the name of an organization, or the taskbar of software in use. An algorithm was employed to eliminate such "background" text.
- *Error correction:* The text discovered with OCR from video frames often contains errors. An automated context-based error correction was applied using *Google Spell Check* API; an approach suggested in [36].
- *Generation of valid N-grams:* Discovering N-grams from OCR extracted slide text poses some unique challenges as i) OCR scans an image left to right and top to bottom and does not guarantee contextual relation between consecutive words, and ii) content on slides usually does not include proper sentences as it lacks punctuation and sentence boundaries. Any N-gram that can be constructed from consecutive words is validated with the *Phrase Finder* API.
- *Stop word removal*: Stop words are frequently occurring and trivial words which help frame sentences but do not represent meaningful content. Examples of stop words are: "a", "and", "at", "the", "it", "with", "what", and "how". The list of stop words is obtained from an external source [37]. Stop words are not removed if they are part of a valid N-gram, such as "cost of living".
- *Stemming:* A stem or root is the part of a word retained after removing its suffixes and prefixes. The stemming process groups all versions of the inflected word to a canonical form. For instance, nouns 'Computer', 'Computers', and 'Computing' are all associated with the stem 'Compute'. This work applied the Snowball stemmer [38]. The stemming process proceeds as follows: i) All occurrences of words are replaced by their stems, and ii) occurrence of each N-gram in the original text is replaced by the most commonly occurring N-gram with a matching stem.

B. Ground truth

This paper uses the ground truth keywords collected by surveying instructors and students familiar with the content of the corresponding lecture videos [9]. Here we briefly describe the nature of the ground truth to provide a context for the research presented in this paper.

Ground truth was available for 121 video segments. The subject areas, in decreasing order of the number of selected segments, are Biology, Computer Science and Chemistry. 16 students and instructors in the relevant subject areas participated in the survey. Keywords were provided by one survey participant for 22 segments, 2 participants for 80 segments, and 3 or more participants for the remaining 19 segments.

C. Logistic regression analysis

The objective of the analysis in this work is to measure the importance of different factors in predicting keywords for lecture videos. After reviewing the data and information available to us, we selected binary logistic regression analysis [34], [35] as the most suitable for this problem. We developed a binary logistic regression model with variables listed in Section III as predictor variables, and ground truth, consisting of N-grams labeled as keywords or not keywords, as the response variable. Our primary goal is quantifying the role of the predictor variables with logistic regression in a manner discussed in [10]. The standardized coefficients resulting from this analysis reflect the strength of the association between the corresponding predictor variable and the outcome. The research questions we address are i) How relevant is a particular variable in predicting if an N-gram is a keyword based on the ground truth, and ii) whether the association between each variable and the ground truth is statistically significant. For our analysis, we have used Minitab software [39].

V. Results

We present a set of results obtained with a linear logistic regression model that capture the relationship between the potential predictor variables listed in Section III and the likelihood of the corresponding N-gram being a keyword, which is the outcome variable.

We observe from Table I that term frequency, screen time, maximum font size, presence in domain dictionary and presence in English dictionary have a significant positive relationship with the outcome, based on the coefficients. They are all

Authorized licensed use limited to: University of Houston. Downloaded on January 26,2022 at 21:57:32 UTC from IEEE Xplore. Restrictions apply.

statistically significant based on the p-values. The model has a AUC (Area under the ROC Curve) value of 0.79 implying a 79% chance that the model will be able to distinguish between keywords and non-keywords.

Term	Coefficient	p-value
Term frequency	0.5131	0
Screen time	0.3335	0
Average font size	0.0125	0.852
Maximum font size	0.5568	0
Domain dictionary	0.5422	0
English dictionary	0.2237	0

TABLE I: Relationship between the set of predictors and the likelihood of an N-gram being a keyword

For a more compact model, we remove average font size from consideration as it was found to have no meaningful relationship. This results in Table II that has modestly changed coefficients. The AUC value for this compact model was nearly identical to that of the original model.

Term	Coefficient	p-value
Term frequency	0.5129	0
Screen time	0.3341	0
Maximum font size	0.5672	0
Domain dictionary	0.5417	0
English dictionary	0.2237	0

TABLE II: Compact predictor variable relationships

Impact of Inverse Document Frequency: It is common in information retrieval systems to employ TF-IDF, that is, the product of term frequency and inverse document frequency, as a predictor variable. Logistic regression results with term frequency replaced by TF-IDF are presented in Table III.

Term	Coefficient	p-value
TF-IDF	0.3301	0
Screen time	0.3752	0
Maximum font size	0.6051	0
Domain dictionary	0.5317	0
English dictionary	0.1999	0

TABLE III: Analysis with TF-IDF in place of Term Frequency

Comparing Table II with Table III, the key observation is that TF-IDF has a weaker relationship with the outcome than simple term frequency.

Analysis for Unigrams: In our analysis a keyword can be a unigram, 2-gram or a 3-gram. In order to form meaningful N-grams from the unstructured text on the screen, a number of heuristic measures were taken, which may impact this analysis. To gain more insight, we conducted an analysis only with unigrams, with the result presented in Table IV.

Term	Coefficient	p-value
Term frequency	0.6041	0
Screen time	0.1291	0.026
Maximum font size	0.6073	0
Domain dictionary	0.4841	0
English dictionary	0.2688	0

TABLE IV: Analysis limited to unigrams

Table IV shows that the relationships are significantly different for unigrams, with the most striking difference being a much reduced degree of positive relationship for screen time. For this model the AUC value increased to 0.81 indicating better prediction performance for unigrams.

Analysis for Specific Domains: The lecture video analyzed in this research span multiple STEM fields with the largest set from Biology followed by Computer Science. In order to gain an understanding whether the relationships of predictor variables may be different for different domains, we analyzed the lecture video segments from Biology and Computer Science. The results are presented in Figure V and VI.

Term	Coefficient	p-value
Term frequency	0.5866	0
Screen time	0.4386	0
Maximum font size	0.6262	0
Domain dictionary	0.4802	0
English dictionary	0.1802	0.002

TABLE V: Analysis with Biology lecture videos

Term	Coefficient	p-value
Term frequency	0.717	0
Screen time	0.515	0
Maximum font size	0.215	0.033
Domain dictionary	0.566	0
English dictionary	0.267	0.018

TABLE VI: Analysis with Computer Science lecture videos

We observe substantial differences, most notably that maximum font size was a much more important keyword predictor for Biology lectures, while term frequency was much more relevant in Computer Science.

We now list some of the more interesting observations with possible explanations:

1) Maximum font size was a strong keyword predictor, while the average font size was not significant. We speculate that the reason is that the maximum font size is much more strongly correlated with being in a title than the average font size. And being in a title/header is the true feature that impacts the likelihood of being a keyword.
2) The raw frequency count is a significantly better predictor than TF-IDF for lecture videos. A possible reason is that the IDF was calculated across segments of a single lecture which is a short span. It is possible that computing IDF across all videos in a course may be more valuable.
3) Presence in English dictionary is a weak predictor. We hypothesize that some context specific non-English words may be keywords among the relatively few words that are not in the dictionary.
4) Significant differences between the two fields that are the main sources of video segments indicate that it may be important to build different models for different fields.

Authorized licensed use limited to: University of Houston. Downloaded on January 26,2022 at 21:57:32 UTC from IEEE Xplore. Restrictions apply.

VI. Limitations

This is an exploratory study that is limited in a number of ways.

- *Limited ground truth:* The ground truth data size is limited to 121 segments in different STEM areas. While we believe this is sufficient to draw a basic understanding, the results are likely to change with a larger body of ground truth. In particular, the relationship of the subject area to the importance of various parameters is not well understood.
- *Set of parameters:* The potential predictor parameters were selected based on our experience and the ease of measuring and including them in the analysis. Several factors were not included, such as i) text color, ii) text boldfacing or italicizing, and iii) location of text on the screen. It is plausible that these factors also are important for keyword analysis.
- *No sentence structure:* The analysis is limited to the set of words and does not include sentence structure.
- *No consideration of transcripts:* The analysis presented only considers screen text.
- *Alternate analysis methods:* Binary logistic regression analysis was selected as the most suited for this problem based on our experience, and informal analysis of the available data. However, it is plausible that an alternate analysis method might provide better results.

We plan to address some of these issues in ongoing and future work.

VII. Conclusions

Identifying keywords in a text document is a challenging task that is complicated further when the text is obtained from frames in a lecture video. This paper builds a foundation for automatic keyword generation systems for lecture videos by quantifying the predictive power of various features of screen text from lecture videos. Results indicate that maximum font size and presence in domain dictionaries are the most important features in predicting keywords.

Future work will address some of the limitations discussed. In particular, we plan to include a larger body of ground truth and explore other properties like text color and text boldfacing. We plan to use speech text and screen text jointly for improved keyword identification as both contain useful information about lecture content. We also plan to incorporate these findings towards building an accurate lecture video keyword prediction system and evaluate it in a real-world lecture video portal.

Acknowledgement

Research was conducted in collaboration with Videopoints LLC. with support, in part, from the National Science Foundation under award NSF-SBIR-1820045.

References

[1] S. D. Sorden and J. L. Ramírez-Romero, "Collaborative learning, social presence and student satisfaction in a blended learning environment," *IEEE 12th International Conference on Advanced Learning Technologies*, pp. 129–133, 2012.

[2] L. Barker, C. L. Hovey, J. Subhlok, and T. Tuna, "Student perceptions of indexed, searchable videos of faculty lectures," in *Proceedings of the 44th Annual Frontiers in Education Conference(FIE)*, Madrid, Spain, Oct 2014.

[3] T. Tuna, M. Joshi, V. Varghese, R. Deshpande, J. Subhlok, and R. Verma, "Topic based segmentation of classroom videos," in *Proceedings of the 45th Annual Frontiers in Education Conference(FIE)*, El Paso, Texas, Oct 2015, pp. 1–9.

[4] T. Tuna, J. Subhlok, L. Barker, S. Shah, O. Johnson, and C. Hovey, "Indexed captioned searchable videos: A learning companion for STEM coursework," *Journal of Science Education and Technology*, vol. 26, no. 1, pp. 82–99, 2017.

[5] M. R. Rahman, S. Shah, and J. Subhlok, "Visual summarization of lecture video segments for enhanced navigation," in *22nd IEEE International Symposium on Multimedia*. Virtual: IEEE, Dec 2020.

[6] R. S. Koka, "Automatic keyword detection for text summarization," Master's thesis, University of Houston, Houston, TX, May 2019.

[7] A. Albahr, D. Che, and M. Albahar, "Semkeyphrase: An unsupervised approach to keyphrase extraction from mooc video lectures," in *2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, 2019, pp. 303–307.

[8] H. Shukla and M. Kakkar, "Keyword extraction from educational video transcripts using NLP techniques," in *6th International Conference - Cloud System and Big Data Engineering (Confluence)*, 2016, pp. 105–108.

[9] R. S. Koka, F. N. Chowdhury, M. R. Rahman, T. Solorio, and J. Subhlok, "Automatic identification of keywords in lecture video segments," in *2020 IEEE International Symposium on Multimedia (ISM)*, 2020, pp. 162–165.

[10] D. Thompson, "Ranking predictors in logistic regression," in *The 2009 Annual MidWest SAS Users Group Conference (MWSUG), Paper D10-2009*, Cleveland, Ohio, 2009.

[11] Echo360, "Echo360," Echo360, 2020. [Online]. Available: https://echo360.com/

[12] Kaltura, "Kaltura," Kaltura, 2020. [Online]. Available: https://corp.kaltura.com//

[13] Panopto, "Panopto," Panopto, 2020. [Online]. Available: https://www.panopto.com/

[14] M. Ketterl, O. A. Schulte, and A. Hochman, "Opencast matterhorn: A community-driven open source software project for producing, managing, and distributing academic video," *Interactive Technology and Smart Education*, 2010.

[15] J. Waitelonis, N. Ludwig, and H. Sack, "Use what you have: Yovisto video search engine takes a semantic turn," in *International Conference on Semantic and Digital Media Technologies*. Springer, 2010, pp. 173–185.

[16] J. Osterhoff, J. Waitelonis, J. Jäger, and H. Sack, "Sneak preview? instantly know what to expect in faceted video searching." in *GI-Jahrestagung*. Citeseer, 2011, p. 261.

[17] R. Mertens, M. Ketterl, and O. Vornberger, "The virtpresenter lecture recording system: Automated production of web lectures with interactive content overviews," *Interactive Technology and Smart Education*, 2007.

[18] C. Choudary and T. Liu, "Summarization of visual content in instructional videos," *IEEE Transactions on Multimedia*, vol. 9, pp. 1443 – 1455, 12 2007.

[19] T. Liu and J. Kender, "Rule-based semantic summarization of instructional videos," in *International Conference on Image Processing*. NY, USA: IEEE, 02 2002, pp. I–601.

[20] T. Liu and J. R. Kender, "Semantic mosaic for indexing and compressing instructional videos," in *International Conference on Image Processing*, vol. 1. Barcelona, Spain: IEEE, 10 2003, pp. I–921.

[21] T. Liu and C. Choudary, "Content extraction and summarization of instructional videos," in *International Conference on Image Processing*. Atlanta, GA: IEEE, Oct 2006, pp. 149–152.

[22] S. Rose, D. Engel, N. Cramer, and W. Cowley, "Automatic keyword extraction from individual documents," *Text mining: applications and theory*, vol. 1, pp. 1–20, 2010.

Authorized licensed use limited to: University of Houston. Downloaded on January 26,2022 at 21:57:32 UTC from IEEE Xplore. Restrictions apply.

[23] Y. Matsuo and M. Ishizuka, "Keyword extraction from a single document using word co-occurrence statistical information," *International Journal on Artificial Intelligence Tools*, vol. 13, no. 01, pp. 157–169, 2004.
[24] M. Litvak and M. Last, "Graph-based keyword extraction for single-document summarization," in *Coling 2008: Proceedings of the workshop Multi-source Multilingual Information Extraction and Summarization*, 2008, pp. 17–24.
[25] A. Hulth, "Improved automatic keyword extraction given more linguistic knowledge," in *Proceedings of the 2003 conference on Empirical methods in natural language processing*, 2003, pp. 216–223.
[26] G. Ercan and I. Cicekli, "Using lexical chains for keyword extraction," *Information Processing & Management*, vol. 43, no. 6, pp. 1705–1714, 2007.
[27] C. Zhang, "Automatic keyword extraction from documents using conditional random fields," *Journal of Computational Information Systems*, vol. 4, no. 3, pp. 1169–1180, 2008.
[28] K. Zhang, H. Xu, J. Tang, and J. Li, "Keyword extraction using support vector machine," in *international conference on web-age information management*. Springer, 2006, pp. 85–96.
[29] R. Chakraborty, "Domain keyword extraction technique: Anew weighting method," *Computer Science & Information Technology*, vol. 109, 2013.
[30] A. Mishra and S. K. Vishwakarma, "Analysis of tf-idf model and its variant for document retrieval," *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 772–776, 2015.
[31] S. Lee and H.-j. Kim, "News keyword extraction for topic tracking," in *2008 Fourth International Conference on Networked Computing and Advanced Information Management*, vol. 2. IEEE, 2008, pp. 554–559.
[32] A. S. Imran, L. Rahadianti, F. A. Cheikh, and S. Y. Yayilgan, "Semantic tags for lecture videos," in *IEEE Sixth International Conference on Semantic Computing*, 2012, pp. 117–120.
[33] A. Balagopalan, L. L. Balasubramanian, V. Balasubramanian, N. Chandrasekharan, and A. Damodar, "Automatic keyphrase extraction and segmentation of video lectures," in *IEEE International conference on technology enhanced education (ICTEE)*. IEEE, 2012, pp. 1–10.
[34] R. E. Wright, "Logistic regression." *American Psychological Association*, 1995.
[35] M. Tranmer and M. Elliot, "Binary logistic regression," *Cathie Marsh for Census and Survey Research, Paper*, vol. 20, 2008.
[36] X. Tong and D. A. Evans, "A statistical approach to automatic OCR error correction in context," in *Fourth Workshop on Very Large Corpora*, 1996.
[37] S. PowerSuite, "https://www.link-assistant.com/seo-stop-words.html," 2018.
[38] M. F. Porter, "Snowball: A language for stemming algorithms," 2001.
[39] B. F. Ryan, T. A. Ryan, and J. B. L. Joiner, "Minitab," Minitab, LLC, 2021. [Online]. Available: www.minitab.com/

Authorized licensed use limited to: University of Houston. Downloaded on January 26,2022 at 21:57:32 UTC from IEEE Xplore. Restrictions apply.