Research Article

A Segmentation Analysis Utilizing Natural Language Processing Model with Interactive Data Analytics Dashboard for Research Management Platform

Raquel C. Adriano[1]*, Anthony U. Concepcion[2], Marian Minneli S. Cruz[1], Alaina Thea V. Concepcion[2]

[1]College of Information and Communications Technology, Bulacan State University, 3000, Philippines
[2]Information Technology Department – Bustos Campus, Bulacan State University, 3000, Philippines

Corresponding author:
E-mail:
raquel.adriano@bulsu.edu.ph

ABSTRACT

Research is a vital component of a university and, currently, unstructured big data is a significant issue in various ICT industries and institutions. To solve this modern problem, the researchers developed a system to streamline the manual operations and traditional research management system of the university through Natural Language Processing (NPL). This quantitative research utilizing descriptive-developmental design is about designing and evaluating A Segmentation Analysis Utilizing Natural Language Processing Model with Interactive Data Analytics Dashboard for Bulacan State University Research Management Platform utilizes the framework of progressive prototyping in the development process. Consultative meetings, interviews and the use of survey questionnaires were held to obtain data from ten (10) RDO/CDRU and staff, twenty (20) IT experts and twenty (20) academicians were chosen using random sampling. Results show that personalize learning management system is excellent in terms of functional suitability (M=4.66), performance efficiency (M=4.68), compatibility (M=4.67), usability (M=4.74), reliability (M=4.51), security (M=4.44), maintainability (M=4.72), and portability (M=4.65). Subsequently, the developed system recorded a grand mean of 4.63 interpreted as Excellent among all ISO/IEC 25010 criteria. This indicates that the system complies with end-user needs as well as software quality standards. It is therefore prepared for adoption. Along with its implementation, it is recommended to gather feedback regularly and conduct an impact analysis of the effectiveness of using the segmentation analysis utilizing natural language processing model with interactive data analytics dashboard for research management platform.

Introduction

As the information economy has become more globalized, research and development have become one of the most potent forces globally. The average number of academic research publications per year is between one and three, but greater numbers could contribute to a surge in academic research publications. The volume and speed of academic publishing have changed dramatically over the past century (Smith, 2015). The number of published papers has increased dramatically from one million in 1980 to greater than seven million in 2014. Furthermore, scientific and academic journals in the Philippines, like in other Asian nations, are confronted with an increasing number of academic research publications. Thus, the research management office considers it difficult to manage the vast majority of data manually.

Research is a critical component of a university's mission and plays a crucial role in advancing knowledge, fostering innovation, and solving complex problems. Therefore, an appropriate research management platform with accurate information regarding research proposals through interactive data analytics dashboards and segmentation analysis using Natural Language Processing (NLP) is necessary. Nowadays, big data is a significant focus of many industries, including finance, healthcare, retail, manufacturing, and even universities. As organizations seek to gain insights and make better decisions from the vast amounts of data they collect, big data can facilitate innovation by analyzing the interdependencies between humans, institutions, entities, and processes and determining new methods to apply these insights. According to the College of Arts and Technology (Philippines), big data analytics refers to the application of techniques to structured, semi-structured, and unstructured large data collections. Big data analytics tools and technologies have become increasingly sophisticated, allowing organizations to process and analyze large data sets in real-time, enabling faster and more accurate decision-making.

From the initial assessment of the processes of the research management office at Bulacan State University regarding research proposals, data banking, and processing, one of the most significant challenges is ensuring data quality. Organizations struggle to effectively maintain their data. To be valuable, data must be utilized, and this requires curation. Big data may contain errors, inconsistencies, and duplications, making analyzing and drawing meaningful insights challenging. Likewise, it is difficult to manually track each proposal's status and evaluation procedure when submitting many research proposals. Furthermore, delays in the review procedure may result in various problems, impacting the project's overall schedule. Today's researchers can publish in a growing number of traditional venues, like conferences and journals, as well as in electronic or online repositories and mega-journals with accelerated publication times (Jones & Kim, 2018).

Much to the desire of the researcher to enhance the processing of the research proposals, segmentation data analysis for big data is being proposed. Segmentation data analysis divides a large and complex data set into smaller, more manageable groups or segments based on specific criteria. Segmentation aims to identify patterns and trends within each segment, which can help develop targeted strategies, improve retention, and increase overall task performance. According to Brown (2020), text segmentation is a fundamentally significant task in NLP that has been tackled at various granularities. Additionally, text segmentation typically involves dividing a document into a succession of topically cohesive sections. Likewise, text segmentation, a common NLP task, separates text into its parts following predetermined criteria. It can also be used with documents, where the goal is to produce logically sound document units. These components, or segments, can be any structure, such as paragraphs or sections. Thus, segmentation analysis using NLP models can effectively categorize

large volumes of unstructured text data (Miller, 2021).

Natural language processing involves text preprocessing, part-of-speech tagging, parsing, and named entity recognition to extract meaningful information from text data. Deep-level grammatical and semantic analysis typically employs words as the basic unit, and text segmentation is typically the most important task of natural language processing (Garcia & Lee, 2019). The study aimed to explore the potential of using NLP models in segmentation analysis and demonstrate how it can uncover insights that may not be immediately obvious through traditional segmentation techniques. It also aims to contribute to the academic literature on data analytics, NLP, and dashboard development by proposing innovative methods and techniques to advance the field and inspire further research.

Hence, the study is expected to advance knowledge by exploring new and innovative ways of analyzing data by integrating NLP models and interactive data analytics dashboards. The proposal can uncover new insights that traditional data analysis methods might miss, leading to a better understanding of various phenomena and problems. Furthermore, using NLP models and interactive data analytics dashboards can help the Research Office make more informed decisions regarding research proposals submitted by faculty members. By analyzing data systematically and objectively, the Research Office can make decisions based on evidence rather than intuition or personal biases.

Data Visualization

According to the University of Waterloo, data visualization is the process of presenting vast quantities of data in a visual format that reveals meaningful patterns in the data. In addition, the incorporation of interactive visualization gave multimodal data such as audio snippets, video, photographs, and transcripts meaning (Chen, 2020). The effectiveness of data visualizations corresponds to the mechanism or method employed, with easily discernible, significant patterns (Smith, 2017). Moreover, as stated by IBM, these visual displays of information can communicate complex data relationships and data-driven insights in a manner that is simple to comprehend by utilizing common graphics such as charts, plots, infographics, and even animations. According to White (2021), when attempting to process text in a machine learning system, the conversion from qualitative to quantitative data is typically the first item to be observed. For images, this is straightforward because pixel values are already in numeric form; however, text must be processed and segmented prior to being fed into a neural network for NLP text analysis.

Research Management Platform

Research is an integral part of a university's mission and plays a vital role in advancing knowledge, fostering innovation, and resolving complex problems. The creation of comprehensive research management platforms is a challenge for institutions seeking to streamline processes and encourage innovation. Research provides a factual foundation for meeting society's needs. In other words, research is the foundation of all initiatives designed to promote global growth (Garcia & Lee, 2019). A study suggests that universities typically establish tenure or performance review criteria in terms of academic research publication. The number of published researchers has risen dramatically (Jones & Kim, 2018). To compare the impact of papers, researchers, journals, and universities, it is essential to use metrics such as the number of papers, number of citations, impact factor, and other sophisticated metrics such as the h-index and altmetrics.

Segmentation Analysis

Several academics have stressed the significance of text segmentation in NLP. Text segmentation is essential for several NLP applications, including information retrieval, machine translation, and text summarization. Research suggests that text segmentation is crucial for sentiment analysis because it can recognize and evaluate the sentiment of certain sentences or phrases (Brown, 2020; Garcia & Lee, 2019). Many techniques, including rule-based, statistical, and machine learning-based, can be used to segment text. The type of text data and the particular NLP task determine the approach to be used. Effective text segmentation allows

researchers and practitioners to extract insightful information from text data and enhance the functionality of NLP applications (Miller, 2021).

Table 1. The table represents the Conceptual Model of the Study.
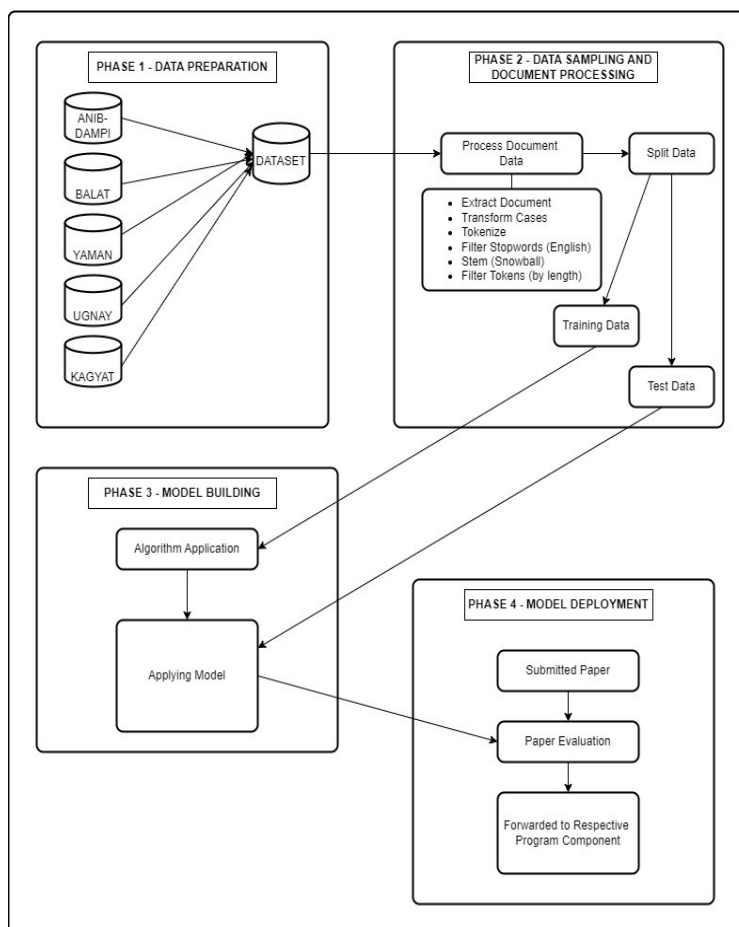
The researcher used the agile software development model in the process phase. It started with the planning wherein the problem is being identified.

Phase 1. In order to facilitate the research document review process, it is important to collect all existing research documents from all program components and compile them into a single dataset. This can involve gathering documents from various sources, such as databases, file systems, or websites, and aggregating them into a common format.

Phase 2. On this phase, the document processing and data splitting will be done. In Process Document Data stage, the dataset will undergo different sub-stages including Extract Content, Transform Cases, Tokenize, Filter Stopwords (English), System Snowball and Filter Tokens (by length).

Phase 3. A model was trained on a training dataset using an algorithm selected. The purpose of training a model is to learn a pattern or relationship between input features and the output variable. The algorithm tried to optimize certain parameters of the model to minimize the difference between predicted and actual values on the training dataset. Once the model was trained, it can be applied on the test dataset to evaluate its performance on unseen data.

Phase 4. It is common for new research documents to be submitted for review and evaluation. Traditionally, this process has been done manually by human reviewers, which can be time-consuming and resource intensive. To address this challenge, the developed system can be used to automate the categorization and evaluation of new research documents.

In the process phase also, the research management platform was evaluated according to some standardized criteria in determining the quality of the software as to the following indicators: Functional Suitability, Performance Efficiency, Compatibility, Usability, Reliability, Security, Maintainability, and Portability.

The output is a proposed Research Management Platform using the NLP model with a data analytics dashboard for BulSU Research Proposals.

Statement of the Problem

The general problem of the study was to develop a segmentation analysis utilizing natural language processing model with interactive data analytics dashboard for Bulacan State University research management platform.
Specifically, this study sought answers to the following questions:
1. What are the salient features of the proposed Research Management Platform?
2. How does the NLP model utilize in clustering the research components of the university?
3. How can the Research Management Platform visualize the status of the university research proposals?
4. How the developed system be evaluated in terms of ISO/IEC 25010 in terms of:
    1.1   Functional Suitability
    1.2   Performance Efficiency
    1.3   Compatibility
    1.4   Usability
    1.5   Reliability
    1.6   Security
    1.7   Maintainability; and
    1.8   Portability

Methods

This study was a quantitative type of research. This research method involves more than data collection and tabulation; it also necessitates accurate analysis, interpretation, comparisons, and trend and relationship detection.

Likewise, Agile as developmental research method was also utilized. The methods introduce new procedures, strategies, and instruments based on thoroughly examining specific circumstances. Furthermore, the study utilized a validated survey instrument to evaluate the developed research management platform.

Natural Language Process (NLP). The developed system involved using machine learning algorithms to categorize new research documents into one of five categories: Anib-Dampi, Balat, Yaman, Ugnay, or Kagyat. These categories may represent different fields of study, topics, or research areas, depending on the specific application. The machine learning algorithms are trained on a labeled dataset of research documents, where each document is associated with a category label.

Figure 2. The figure represents the Agile Development process as Research Methodology

Software Development Model. Users' input was used as the key control mechanism in agile models. It's referred to as a people-centered approach. Furthermore, because its testing cycle time is relatively short, it offers a workable version of the product at an early stage.

Table 1. The table represents the Respondents of the Study

Respondents	Frequency	Percentage
RMO/CRDU and staff	10	0.20
IT experts	20	0.40
Academician	20	0.40
Total	**50**	**100%**

As can be gleaned from the table, 10 or 20% of RMO (Research and Management Office) staff will evaluate the proposed RMO platform. Twenty or 40% will come from IT experts, and another 20 or 40% will come from academicians. The respondents of the study were RMO admin/CRDU and staff, IT experts and Academician.

Table 2. The table represents the 5-point Likert Scale as Research Instrument

Scale	Range	Descriptive Rating
5	4.21 -5.0	Excellent
4	3.41 - 4.20	Very Good
3	2.61 -3.40	Good
2	1.81 -2.60	Fair
1	1.00 -1.80	Poor

The study utilized a standardized software development instrument. A questionnaire was a critical tool used in the study. The researchers also utilized the ISO/IEC 25010 to standardize quality. A metric-based approach is used to assess a given system's quality, and an evolving system reliability model is used to predict quality (Budi et al., 2022).

Data Gathering Procedure

The researcher prepared a letter to conduct the study. The letter was noted by the adviser and approved by the dean of the Graduate School of La Consolacion Univerity-Philippines (LCUP). The researcher personally submitted the letter to the Office of the Vice President for Research and Extension of Bulacan State University for approval. Likewise, she requested permission to get the data from the Research Management Office and the Campus Research Development Unit for pilot testing of the developed platform.

Upon the approval of the Office of the Vice President for Research and Extension in the conducted study, a courtesy letter for the Director of Research Management Office through the Campus Research Development Unit was prepared and presented.

. As part of the study, the accomplished survey instruments were checked, classified, tallied, tabulated, analyzed, and processed based on the research design earlier stated in this chapter and in preparation for the upcoming oral/final presentation for dissertation evaluators/ panellists.

Data Processing and Statistical Treatment

The data gathered for the study was tallied and computed using Microsoft Excel. In contrast, for the statistical treatment of the data,

the researcher asked the assistance from a qualified statistician. The statistical treatment of the data was processed using Statistical Packages for Social Science (SPSS).

Descriptive statistics like frequency, percentage, and weighted mean were utilized in analyzing and interpreting the data.

Result and Discussion

The developed system was evaluated based on software and data quality of ISO/IEC 25010. This section presents the results of the system evaluation with the designed survey questionnaire adapted from ISO/IEC 25010 standard

software. The questionnaire was answered by twenty (20) information technology experts (programmer) as Alpha evaluators, twenty (20) Academician as Beta evaluators, and ten (10) RMO admin and staff Gamma evaluators that are concerned with the developed system. Hence, a total of fifty (50) respondents served as the basis for the system assessment.

The researcher used the ISO/IEC 25010 evaluation tool. The result from the tool determined whether the system is acceptable in terms of functionality suitability, performance efficiency, compatibility, usability, reliability, security, maintainability, and portability.

Table 3. The table shows the Summary of the Evaluation using ISO/IEC 9126-1:25010:2011 QUALITY STANDARDS of IT Experts, Academician, and RMO

Indicators	WM	Verbal Interpretation
Functional Suitability	4.72	Excellent
Performance Efficiency	4.78	Excellent
Compatibility	4.72	Excellent
Usability	4.80	Excellent
Reliability	4.67	Excellent
Security	4.58	Excellent
Maintainability	4.73	Excellent
Portability	4.77	Excellent
Grand Mean	**4.72**	**Excellent**

Table 3 presents the overall ratings of IT experts, academician, and RMO admin and staff using ISO/IEC 9126-1:25010: 2011 quality standards. It was revealed that the average grant mean was excellent (4.72). All indicators evaluated in terms of Functional Suitability, Performance efficiency, Compatibility, Usability, Reliability, Security, Maintainability, and Portability were rated excellent.

The salient features of the proposed Research Management Platform are the following:

The system may offer analytical tools to generate reports and charts summarizing the clustering of the research proposal based on the research components of the university. These reports can be useful for evaluating the institution's research productivity and identifying areas that may require improvement.
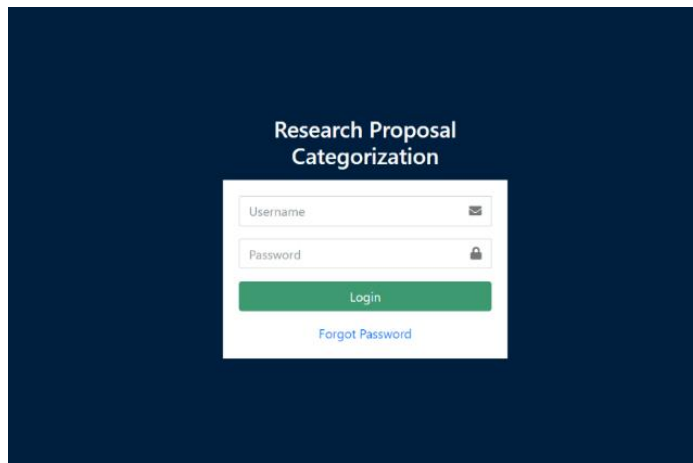
Figure 3. The figure shows one of the security features of the system, which is the Login Page. To track information, history of the users entering the system
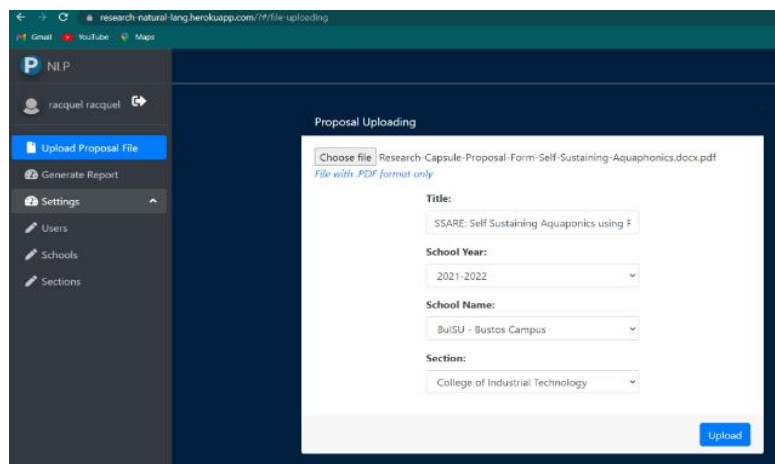

Figure 4. The figure presents File uploading, as inputted via a PDF file manually uploaded via the research proposals. This would include an upload of rationale of research proposal
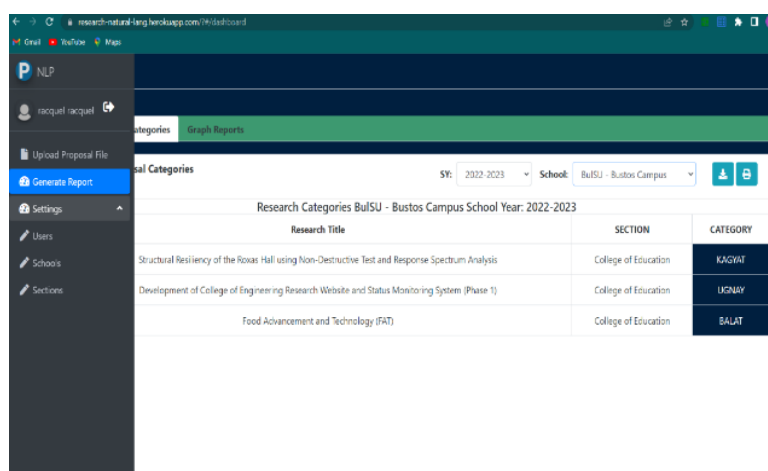

Figure 5. The figure shows the Generate reports page, which is a feature of the system that includes a report generation summary of research proposals per research title, section or department/college, category/ research components, school year and school/campuses
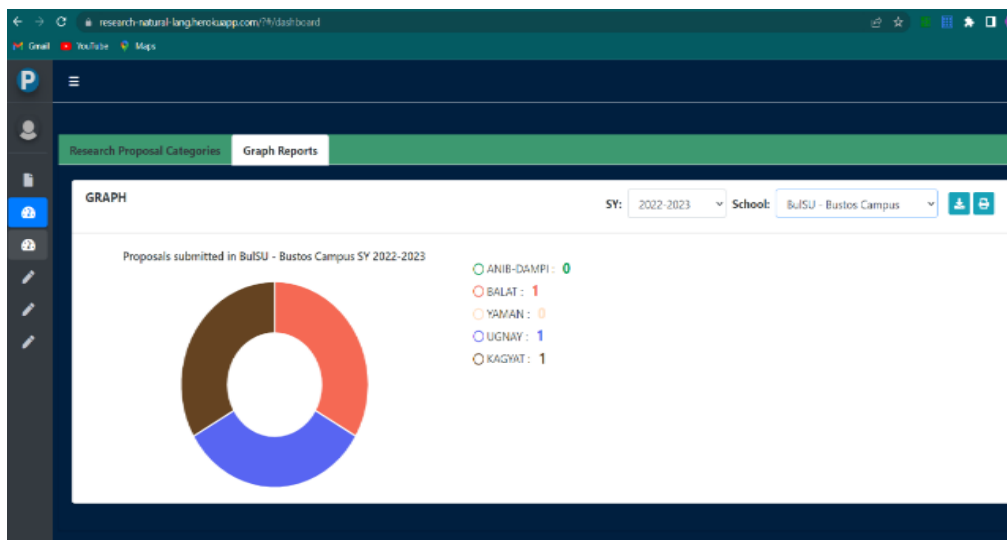
Figure 6. The figure shows the Graph Reports page, which is the part of the system that presents the summarized data and makes data presentable and easy to understand
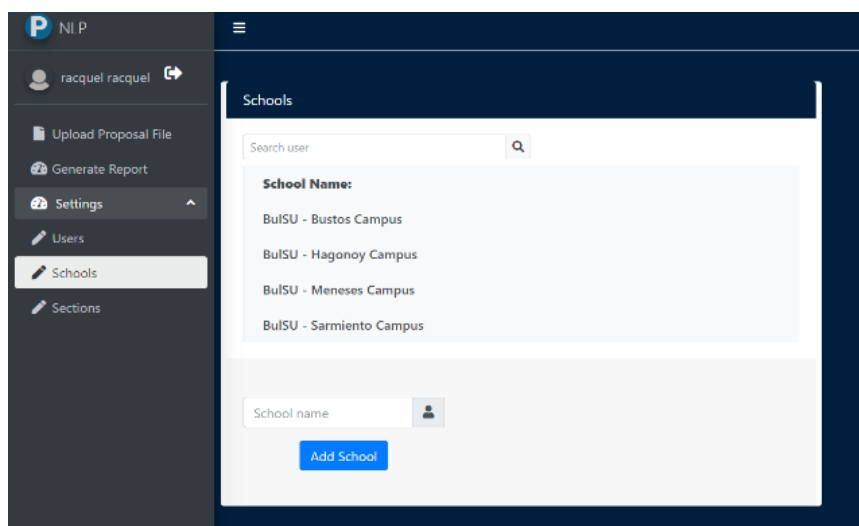

Figure 7. The figure shows the school page. Schools page is a part of the system was used to add school name/campuses for uploading the research proposals

The NLP model utilized in clustering the research components of the university.

The system involves using machine learning algorithms to categorize new research proposal documents into one of five categories: Anib-Dampi, Balat, Yaman, Ugnay, or Kagyat. These categories may represent different fields of study, topics, or research areas, depending on the specific application. The machine learning algorithms are trained on a labeled dataset of research documents.

A sample rationale was provided that was learned by the algorithm. A model was trained on a training dataset using an algorithm. The purpose of training a model was to learn a pattern or relationship between input features and the output variable. The algorithm tried to optimize certain parameters of the model to minimize the difference between predicted and actual values on the training dataset.
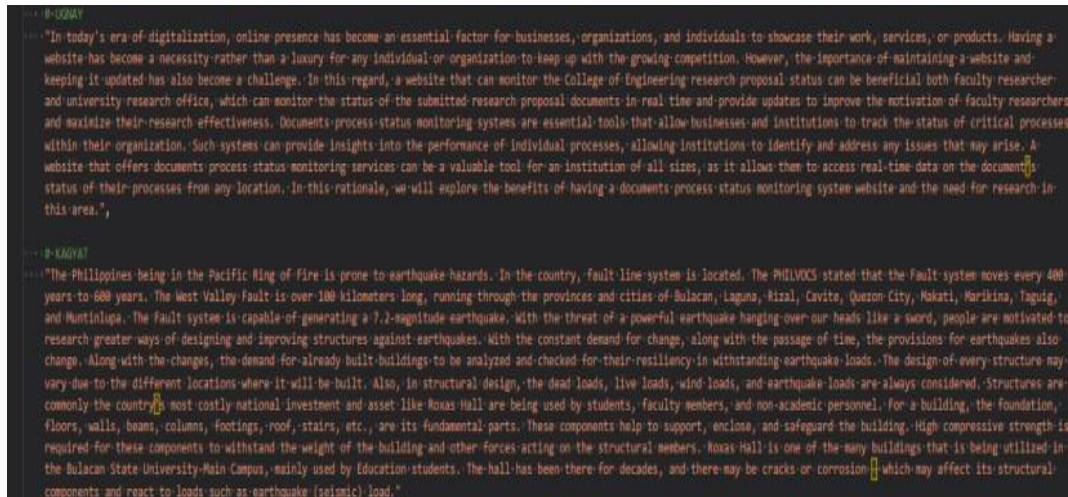
Figure 8. The figure shows a sample rationale as a test subject for the system

This are the sample rationale that was learned by the algorithm. A model was trained on a training dataset using an algorithm. The purpose of training a model was to learn a pattern or relationship between input features and the output variable. The algorithm tried to optimize certain parameters of the model to minimize the difference between predicted and actual values on the training dataset.

The "en_core_web_md" is like a dictionary where the words of rationale inside the system goes through the "en_core_web_md", then all the words of rationale inside the dictionary were compared.

```
nlp = spacy.load("en_core_web_md")
# print(train_x)
docs = [nlp(text) for text in train_x]
train_x_word_vectors = [x.vector for x in docs]
```

Figure 9. The figure shows a sample code named as en_core_web_md

The researcher used "mode_word_vector.fit()" to create model using the sample rationales and the words came from the "en_core_web_md" that serves as dictionary, then it will now group/cluster the words based on the given category.

```
test_x = [given_rationale]
test_docs = [nlp(text) for text in test_x]
text_x_word_vectors = [x.vector for x in test_docs]

    You, 2 months ago • done connecting php and python code

real_model = joblib.load(this_dir + '/raquel-training.joblib')

result = real_model.predict(text_x_word_vectors)

print(result[0])
```

Figure 10. The figure shows a sample code named as joblib file

This is the prediction part, the ones used was the "joblib file" (model that used to predict function in clustering research proposals).

The NLP algorithm will gather all the rationale in the research proposals for training and clustering, then algorithm will be proven

based on the words of the rationale to create NLP model.

Summary of Findings

After careful analysis of all the data gathered and evaluated, the summary of findings is presented as follows:

1. The development of the system was managed carefully to come up with the desired output. The comments given by the respondents comply with the result of the evaluation which corresponds that the system has met its objectives.
2. The study developed a system suitable for web platform that automates the processes involved in predicting program placement using python computer programming language that was often to use to build websites and software, automate tasks, and conduct data analysis.
3. The different features and functions that are provided in the system it includes major functions of segmentation analysis utilizing NLP model such as (a) File uploading (b) Generate reports and (c) Graph Reports.
4. The NLP model was successfully tested and used in the system and proved to be integral in the development of the system.
5. The system was evaluated using the ISO/IEC 25010. The following evaluators have been grouped into three: IT experts as alpha evaluator, Academician as beta evaluators, and RMO admin and staff as gamma evaluators. The results yielded are Excellent by the evaluators on the following attributes: Functional Suitability, Performance Efficiency, Compatibility, Usability, Reliability, Security, Maintainability, and Portability. In summary, this is a good indication that all respondents have seen the overall performance of the system as generally successful.

Conclusion

After the study was completed, demonstrated, and tested, the researcher concluded that all objectives in the study entitled A Segmentation Analysis Utilizing Natural Language Processing Model with Interactive Data Analytics Dashboard for Bulacan State University Research Management Platform were met.

Based on the NLP model used in the system, this model is useful in categorized BulSU research proposals because they provide a way to systematically analyze a large amount of data and identify the most important factors or variables to categorized the research proposals.

The developed system is suitable for web platform that helps to support the RMO admin and staff can more accurately know the categorize or classification of research proposals, which can help ensure to provide valuable tools and insights to streamline the research proposal process and improve research quality.

1. The system can now be accessed through internet with the use of any platform.
2. Based on the evaluation results yielded, the system has successfully met and achieved all the desired functionalities required of a Segmentation Analysis Utilizing NLP Model.
3. The results of the evaluation suggest the possibility of using the A Segmentation Analysis Utilizing Natural Language Processing Model with Interactive Data Analytics Dashboard for Bulacan State University Research Management.

Recommendations

The following are the deemed recommendations of the researcher based on the findings and conclusions that were derived:

1. Upgrade bandwidth to prevent network congestion and bottlenecks, which can cause latency problems, packet loss, and sluggish response times.
2. Include employees in the process of learning new skills, acquiring new knowledge, and training. Training employees means showing them how to do a certain task or method to help them do their jobs better.
3. Consider recommendations and gather feedback regularly. Create segmentation that satisfies customers by prioritizing the features that the majority of users was employ, based on their input.

References

Boston University. (n.d.). *Research Information Management System (RIMS).* Retrieved from https://www.bu.edu/tech/services/admin/research-systems/rims/

Bowdre, P. R. (2020). The use of predictive analytics to shift the culture of academic advising toward a focus on student success. *Journal of Educational and Social Research, 7(3),* 45–55. Retrieved from https://jespnet.com/journals/Vol_7_No_3_September_2020/3.pdf

Budi, S., Gata, W., Noor, M., Pangabean, S., & Rahayu, C. S. (2022). News portal website measurement analysis using ISO/IEC 25010 and McCall methods. *Journal of Applied Engineering and Technological Science (JAETS), 4(1),* 273–285.

Demmelmaier, G., & Westerberg, C. (2021). *Data segmentation using NLP: Gender and age.* Retrieved from https://www.diva-portal.org/smash/get/diva2:1527530/FULLTEXT01.pdf

Dhingra, P., Gayathri, N., Rakesh Kumar, S., Singanamalla, V., Ramesh, C., & Balamurugan, B. (2020). Internet of Things–based pharmaceutics data analysis. *Emergence of Pharmaceutical Industry Growth with Industrial IoT Approach,* 85–131. https://doi.org/10.1016/b978-0-12-819593-2.00004-2

Elia, G., Polimeno, G., Solazzo, G., & Passiante, G. (2020). A multi-dimension framework for value creation through big data. *Industrial Marketing Management, 90,* 617–632.

Elsevier. (n.d.). *Why you need a Research Information Management System (RIMS).* Retrieved from https://beta.elsevier.com/academic-and-government/why-you-need-cris?trial=true

Fire, M., & Guestrin, C. (2019). Over-optimization of academic publishing metrics: Observing Goodhart's Law in action. *GigaScience, 8.* https://doi.org/10.1093/gigascience/giz053

Gao, H., Xu, Y., Yin, Y., Zhang, W., Li, R., & Wang, X. (2019). Context-aware QoS prediction with neural collaborative filtering for Internet-of-Things services. *IEEE Internet of Things Journal,* 1–1. https://doi.org/10.1109/jiot.2019.2956827

Johnson, R., Watkinson, A., & Mabe, M. (2018). *The STM Report, 5th edition: An overview of scientific and scholarly publishing.* STM. Retrieved February 11, 2025, from https://coil-ink.org/20.500.12592/7qcc41

Joshi, A., & Unger, K. (2017). *Introduction to data visualization and visualizing your data effectively.* Retrieved from https://www2.cs.uh.edu/~ceick/UDM/COSC3337-DV1.pdf

Kelly, S. (2018). The continuing evolution of publishing in the biological sciences. *Biology Open, 7(8).* https://doi.org/10.1242/bio.037325

Li, J., Chiu, B., Shang, S., & Shao, L. (2020). Neural text segmentation and its application to sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering, 34(2),* 828–842.

Li, J., Sun, A., & Joty, S. R. (2018, July). SegBot: A generic neural text segmentation model with pointer network. *Proceedings of the 27th International Joint Conference on Artificial Intelligence,* 4166–4172. https://doi.org/10.5555/3304222.3304349

Love, DeMonner, & Teasley. (2021). *Show students their data: Using dashboards to support self-regulated learning.* Retrieved from https://er.educause.edu/articles/2021/7/show-students-their-data-using-dashboards-to-support-self-regulated-learning

Lukasik, M., Dadachev, B., Simoes, G., & Papineni, K. (2020). Text segmentation by cross-segment attention. *arXiv preprint arXiv:2004.14535.*

Nistor, N., & Hernández-García, Á. (2018). What types of data are used in learning analytics? An overview of six cases. *Computers in Human Behavior, 89,* 335–338.

Olivetti, E. A., Cole, J. M., Kim, E., Kononova, O., Ceder, G., Han, T. Y. J., & Hiszpanski, A. M. (2020). Data-driven materials research enabled by natural language processing and information extraction. *Applied Physics Reviews, 7(4),* 041317.

Shtekh, G., Kazakova, P., Nikitinsky, N., & Skachkov, N. (2018, October). Applying topic segmentation to document-level information retrieval. *Proceedings of the 14th Central and Eastern European Software Engineering Conference Russia,* 1–6.

Tyagi, A. (2021). A review study of natural language processing techniques for text mining. *International Journal of Engineering Research & Technology,* 10(9). https://doi.org/10.17577/IJERTV10IS090156

University of Waterloo. (n.d.). *Data visualization.* Retrieved from https://uwaterloo.ca/centre-for-teaching-excellence/catalogs/tip-sheets/data-visualization

Wang, D., Su, J., & Yu, H. (2020). Feature extraction and analysis of natural language processing for deep learning English language. *IEEE Access, 8,* 46335–46345.

Whitehead, D. (2019). *Re: What is the standard number of publications per year (especially journal papers) for a researcher?* Retrieved from https://www.researchgate.net/post/What-is-the-standard-number-of-publications-per-year-specially-journal-papers-for-a-researcher/5c42b0bcc7d8ab077805162a/citation/download

Wright, A., & Wiklicky, H. (2019). *Comparison of syntactic and semantic representations of programs in neural embeddings.* Retrieved from https://arxiv.org/pdf/2001.09201.pdf