

JGB 1753**Comparison of Filipino Presidents' SONAs using Cosine Similarity and
Latent Semantic Analysis**

*Heidi Batara, Raquel Bautista, Leonardo Gavino III, Karl Adrian Laroya,
Rodolfo Aaron Mercado, Rianne Justin Policarpio, Beverly Estephany Ferrer*

Saint Louis University

bpferrer@slu.edu.ph, 2200055@slu.edu.ph,

2200311@slu.edu.ph, 2200042@slu.edu.ph, 2200591@slu.edu.ph,

2202631@slu.edu.ph, 2200848@slu.edu.ph

Abstract

State of the Nation Addresses (SONA) is impactful speeches that reflect a president's plans and methods of governance for their terms. Despite the significance of these speeches, there still needs to be more literature that analyzes them in the context of similarity and machine learning. Hence, this study was conducted to find significant patterns from SONA using Cosine Similarity and Latent Semantic Analysis (LSA). To achieve this, the researchers scraped the Philippine presidents' transcripts from 1935 to 2022, then performed comparisons by utilizing the extracted topics and cosine distances. The analysis is categorized into presidents' family ties, party list affiliation, and historical era. Based on the findings, there are distinct recurring patterns for each category. For family ties, there is a pattern of influence in the Macapagal family, while Marcos and Aquino's families have none. Partylist affiliation does not influence where a president belongs, and Philippine party lists are mostly personality-focused instead of ideology-focused. Regardless of the era, agriculture, economy, labor, and employment are

recurrent issues, while technology issues are newly addressed. Similar SONAs based on cosine distance tend to have similar topics assigned through LSA. Thus, the LSA and cosine distance algorithms are practical and recommended document analysis methods.

Keywords: *State of the Nation Address, Machine Learning, Cosine Similarity, Latent Semantic Analysis.*

Introduction

State of the Nation Address, SONA, is delivered by the current president of the Philippines during their term. It is a speech informing people about the nation's current political and economic state and the current administration's plans (Bringula & Miranda, 2021). SONA and other presidential speeches play a significant role in the history of the Philippines' governance, which is why they are well documented along with their effects (Lineses, 2018). Also, presidential speeches have a more significant influence since they are designed to sway people and can contain events and developments of a country from different timelines (Bringula & Miranda, 2021). These characteristics of the SONA make it notable for applying data mining analytics for further analysis for these speeches to be better understood.

Although related literature analyzed the SONAs of Philippine Presidents using sentiment analysis and topic modeling through Latent Dirichlet Allocation, it only focused on 13 presidents. In addition, this study would use cosine similarity to identify the similarities and dissimilarities of the SONA. Lastly, the study would analyze the SONAs based on the different categories.

This study aims to determine the similarities and differences between all the SONAs of Philippine presidents until 2022, except Emilio Aguinaldo and José P. Laurel, as they did not

deliver a SONA during their term. Specifically, it aims to extract significant connections and patterns between the presidential SONAs in terms of (1) Family Ties, (2) Partylist, (3) Eras, and (4) the Pandemic Period.

The reason for choosing these categories is to address the problems that stem from them. Political dynasties are widely recognized as a source of corruption in the Philippines, but despite this, there is still a lack of solutions to fix it (Purdey et al., 2016). Philippine party lists also focus more on personality than pursuing their proclaimed ideologies. Their ideologies are almost identical and are another source of corruption (Quimpo, 2007a). The Philippines is also heading towards a pattern of increasing authoritarian tendencies, threatening its democracy, hence why different eras are considered (Quimpo, 2007b).

By extracting patterns and determining similarities and dissimilarities between the presidential SONAs, the paper can provide insights into the political patterns of political dynasties, party list corruption, and increasing authoritarian tendencies while determining what issues have prevailed or have been solved by succeeding presidents. This information can then be used as a reference for future literature in Philippine politics.

Considering that SONAs have various types of listeners, including motley groups: the investors and business people (Sicat, 2012), they can gain information on the current and future focus of the country, which helps to understand the business direction of the country. The study could also be a starting point for the analysis of documents through the use of automation instead of manual examination. This can speed up the process of extracting useful information, especially patterns, from numerous documents, which is especially useful with the prevalence of big data that is often underutilized by public government sectors despite the possible benefits (Gouri & Uddin, 2019). Moreover, the study acknowledges the significance of embodying

competitiveness and business-related topics. The SONA's economic section covers several business sectors, all of which are critical to the country's economic progress. With that, the paper can contribute to the analysis of the economic aspect within the SONA.

Review of Related Literature

This section presents the related literature and studies that centers around the two machine learning methods in analyzing documents.

Cosine Similarity

Numerous studies have attempted to use proximity measures, specifically cosine similarity when comparing documents or responses. Their purpose is either to develop an alternative to facilitate their operations or to find patterns in what they are investigating.

A study focusing on the educational field in Indonesia utilized cosine similarity to enhance the Automated Essay Scoring system. It found that using this principle will hasten the correction process and provide more objective results regarding the relevant text on documents (Lahitani et al., 2016), which was supported by another study that used cosine similarity in calculating text relevance, which concluded that the closer the value to 1, the more relevant it is (Gunawan et al., 2018). These researches point out the advantages of using cosine similarity in measuring documents.

Another study on health analyzes the similarities between countries on how residents respond to government policies during the COVID-19 pandemic. They identified that countries in the same region have similar responses, emphasizing the importance of analyzing patterns to assist in planning for future emergencies (McKenzie & Adams, 2020).

One source of issues is the president's speech, as they involve what is happening in the country. A study examined the distance between the US presidents' speeches concerning

economic content. This results in a network pattern, wherein it was divided into two components: one strongly connected speech and lighter edges. They discovered that during the president's political career, they mentioned financial and economic sectors at least once (Cinelli et al., 2021).

Most of these studies tend to concentrate on other countries, and there still needs to be more study that analyzes Philippine-related documents. No research has focused on the similarities or dissimilarities between the Filipino Presidents' SONAs.

Latent Semantic Analysis

Topic modeling has helped model textual data; two popular approaches are Latent Dirichlet Allocation (LDA) and Latent Semantic Analysis (LSA). Despite their commonalities, this paper focused on LSA. A study using TopicView found that LSA provided good summarizations among broad groups of documents compared to LDA (Rocamora, 2002). It applies to this study as the documents vary based on presidents, focus, and periods. Other studies have also yielded LSA as a promising approach in their studies (Hoblos, 2020; Bergamaschi & Po, 2015).

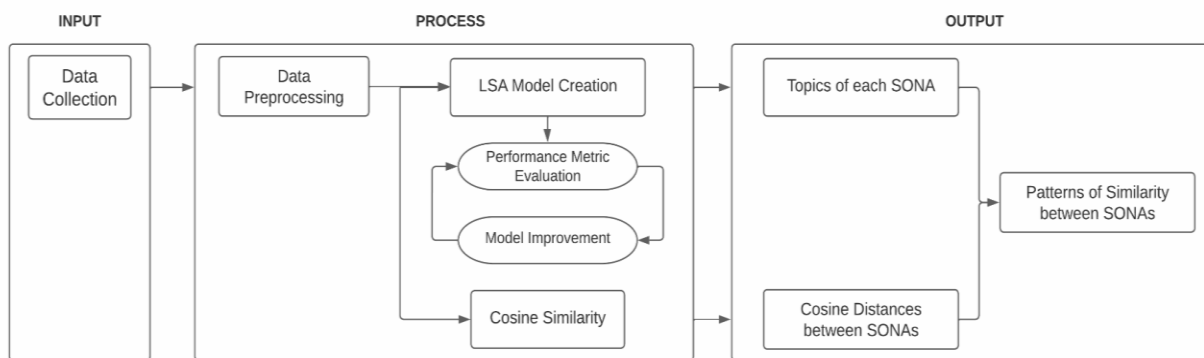
Moreover, a study analyzed the transcripts of the 2016 US presidential debates between Hillary Clinton and Donald Trump using LSA (Valdez et al., 2018). It emphasized the relevance of topic modeling when it comes to social issues. In the Philippines, a study utilized LDA and sentiment analysis to analyze president SONAs (Bringula & Miranda, 2021). Currently, no studies have applied LSA to study the similarity of SONA concerning cosine similarity. Therefore, this study will utilize LSA and Cosine Similarity to broaden further topic modeling's contribution to social issues in the Philippines.

Framework

As presented in Figure 1, the research framework uses the input-process-output model. The input covers the data collection, the process covers the Cosine Similarity and LSA model, and the output covers the analysis process to find the patterns of similarity between the SONA.

Figure 1

Research Framework



Methodology

The study utilized the quantitative approach, which is used to quantify a concept based on numerical data and mathematical methods (Idowu, 2015). The researchers mainly corroborated the results of the trained LSA Model and Cosine Similarity score to produce the patterns of similarity.

Data Collection

The raw SONA database was obtained by retrieving all the available presidents’ transcripts from 1935 to 2022 using the BeautifulSoup Python library to scrape from the Philippines’ Official Gazette website.

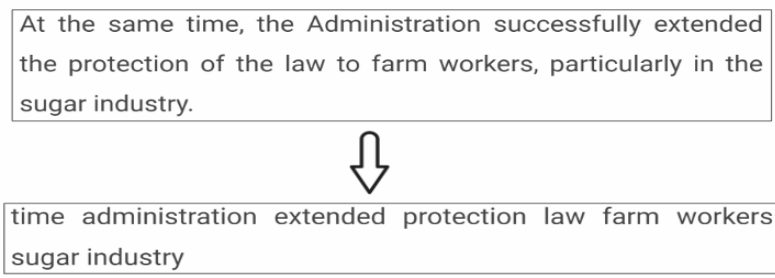
Data Preprocessing

Upon exploring, the speeches are already well constructed and prepared, considering that it is a national address. Thus, the researchers only removed English stop words and special

characters such as punctuations, apostrophes, and periods using the NLTK Python library (see Figure 2).

Figure 2

Data Preprocessing Snippet

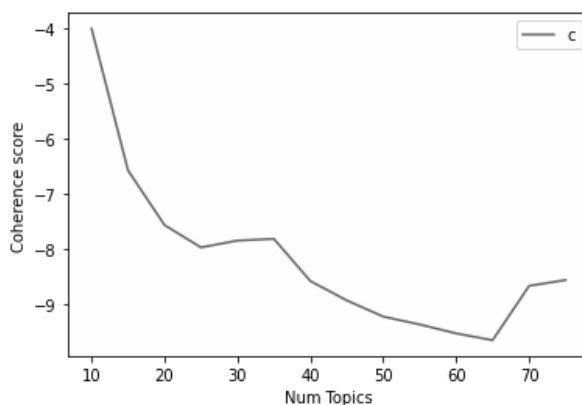


Latent Semantic Analysis

The researchers first created a TF-IDF model from the corpus and fitted its output into the LSA model. The number of topics used in the LSA model is 65, the lowest number in the Coherence Score in Figure 3. The coherence score measures the similarity of words, which is obtainable using a coherence model. The lower the score, the better the performance and coherence of the model (Röder et al., 2015).

Figure 3

U_{mass} Coherence Score



The top 10 words with negative and positive weights of each topic are used for the interpretation, with 20 words used in total. The model will then assign the top 3 topics of each SONA, further corroborating with related articles and journals to validate the assigned topics.

Cosine Similarity

Cosine Similarity is a measurement that quantifies the similarity between two or more vectors. It is measured by the cosine of the angle between two vectors and determines whether two vectors are pointing in roughly the same direction. This is an ideal method to utilize for measuring document similarity. All SONAs were vectorized using the TF-IDF vectorizer and then subjected to the formula for Cosine Similarity.

Data Analysis

For LSA, the researchers considered the weights of the words per topic before coming up with the interpretation. Words with a positive weight signify that it is highly representative of the topic; otherwise, it does not represent the topic (Aquino & Chavez, 2018).

For cosine similarity, the researchers considered the distance. A distance near 0 means high similarity, while near 1 is the opposite. Heat maps are used to visualize the comparisons for the categories. The darker color indicates a high degree of similarity, and the lighter color indicates the opposite.

The results are then further corroborated through other literature, such as news articles, journals, and historical accounts, and checking the contents of the SONA to validate the findings.

Discussion of Results

This section presents the topics derived from the SONA and cosine similarity scores.

Topic Modeling

There are 65 extracted topics from the SONAs collected, and 28 main topics presented in Table 1, which the presidents commonly address. Sixty-five is the best number of topics that yielded the best coherence value of 0.56. With this, it signifies that from the second president of the Philippines to the latest president, the presidents focused on and discussed different concerns essential to the economy's growth.

Table 1

Extracted Topics from the SONA

No.	Topic Interpretation
0	Agricultural production; Proposed solutions such as export, import, exchange
1	Filipino people's living situation; Employment situation
2	Veteran official's programs; Reparation of damages
5	Taxation and poverty
8	Veteran official programs;
9	Economic status report; Labor and employment
11	Societal and geographic organizations; Resignation of officials
13	COVID pandemic and online mode of transactions
14	Sales tax; Drug dealings, kidnapping, and corruption
15	Anniversary of Ninoy Aquino's Death
17	Survivors of Payatas tragedy; National/Economic Recovery
18	Announcement of the resignation of high-ranking officials
23	Martial law, National Defense, Political normalization
24	National socio-economic development; Value-Added Tax

26	Rescuing Angelo dela Cruz
30	Economic growth over the years
31	Historical events discussion
34	Plans for government officials
43	Kilusang Kabuhayan; Dictatorship
44	Resignation of political leaders; ASEAN
45	Kilusang Kabuhayan at Kaunlaran
46	Value-Added Tax
47	Ratification for a political change
54	Tobacco
56	Technology and modernization
58	Illegal mining and logging
62	The assistance from the US

The extracted topics are interpreted using their top 20 words with positive and negative weights. For example, the top words with positive weights for topic 2 are "veteran," "reparation," "recommend," and "commonwealth," hence the interpretation is "Veteran official's programs" or "Reparation of damages." Lastly, the topic is verified by checking the contents of the SONA assigned with that topic and other related literature, such as news and journal articles.

These topics can provide information to different individuals and organizations, especially the citizens, about the issues or concerns of the country from the past until the present. Also, this will be used as a basis for further discussion, which will be presented in the following subsections of this paper.

Comparison of SONAs based on Family Relationship

During the election, it is evident that political dynasties are present. A political dynasty is when several family members are involved in politics or occupy a political position (Mendoza et al., 2013). With that, it is vital to analyze if there are existing patterns, as some studies found that the increase in political dynasties can be related to bad governance (Mendoza et al., 2020; McCoy, 1994). Three notable families are the Macapagal, Marcos, and Aquino, as two family members occupied the highest position in the Philippine government.

Macapagal Family

Diosdado Macapagal (Father) and Gloria Macapagal-Arroyo (Daughter) tackled social and economic development, Topic 24. During their term, Diosdado Macapagal presented a five-year program concerning socio-economic (Starnier, 1963). Gloria presented different agendas for socio-economic reforms that became the foundation during her term. Upon exploring these SONAs, both discussed issues concerning taxes belonging to Topics 5, 2, and 46 and labor and employment belonging to Topics 1 and 9. She was able to discuss it as well since one aspiration of the residents of the Payatas was a livelihood, which she provided (see Table 2). Hence, both concentrated on employment to achieve economic growth.

The cosine distance of the SONA can support the similarities between both presidents. The three of the SONAs of Gloria were similar to the four SONAs of Diosdado, as the scores range from 0.18 to 0.24, as shown in Figure 4(a).

As for their dissimilarities, some of Gloria's SONAs were highly dissimilar to her father, ranging from 0.37 to 0.48. Diosdado discussed the international relations of the Philippines with other Western countries and the United States and tobacco issues. In contrast, Gloria tackled

illegal drug dealings and kidnapping, which she discussed in her 2003 SONA and during the Angelo de la Cruz kidnapping incident.

Table 2

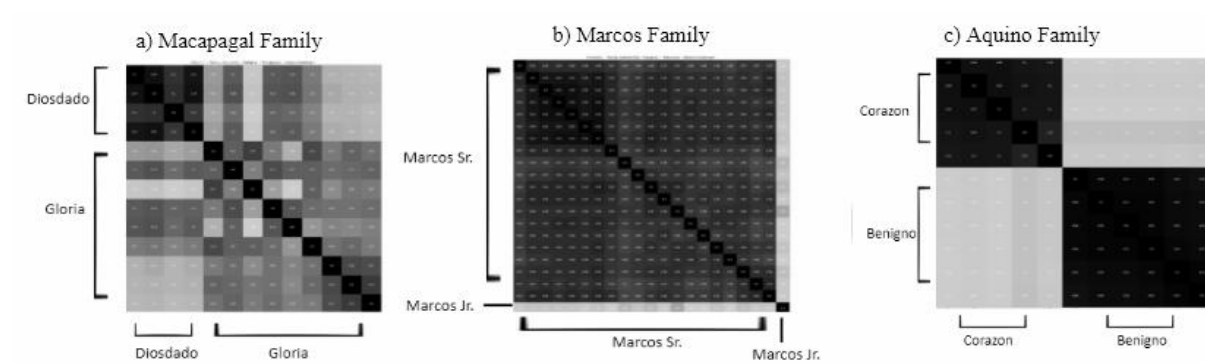
Similarities of the presidents based on family relationship

Macapagal Family	Marcos Family	Aquino Family
Agricultural Production Social and Economic factors Labor and Employment	Labor and Employment	Agricultural Production

Marcos Family

The SONAs of Ferdinand Marcos Sr. (Father) are dissimilar to Ferdinand Marcos Jr. (Son), as Marcos Sr. addressed more on the economic and moral crisis. Also, he discussed the proclamation of martial law, the creation of Kilusang Kabuhayan at Kaunlaran, and the cooperation with the Philippines’ ASEAN neighbors. From his first term to his last term, he talked about agriculture. In comparison, his son focused more on the COVID Pandemic, Topic 13, and the application of technology and modernization in different industries. He signed an agreement with other countries to improve prominent areas such as digital and education, creating greater collaboration in the digital economy (*Joint Press Statement, 2022*). Still, both addressed labor and employment.

The dissimilarities of their topics were shown using the cosine similarity, as the SONAs of Marcos Sr. are slightly dissimilar to Marcos Jr. since the values fall between 0.22 and 0.29, as presented in Figure 4(b). Moreover, when comparing the SONA delivered by Marcos Sr. to his other SONAs, it can be seen that these are highly similar. This means that throughout his presidential terms, he addressed almost similar issues.

Figure 4*Cosine Similarity of the three families*

Note. The darker the color, the lower the distance between the SONAs, indicating the more similar the documents are.

Aquino Family

The topics Corazon Aquino (Mother) addressed were highly dissimilar to her son, Benigno Aquino III. After the Marcos Sr. administration, Corazon talked about the reformation of the internal structure, Topic 43, and the death of his husband, Topic 15. She achieved reform, though not all, and the political environment changed (Dohner & Intal, 1989). In comparison, Benigno focused more on the affected Filipinos by the Typhoons Pepeng and Ondoy, belonging to Topic 1, in which he provided housing units to the victims, which was considered in Topic 2 (Gaje, 2010).

The cosine distance between their SONAs can support these. Based on Figure 4(c), Corazon's SONAs were highly dissimilar to her son, as unlike the first two families, the scores range from 0.89 to 0.98.

Comparison of SONAs based on Partylists

Political parties or party lists are parties composed of political elites and are distinguished by their unique ideologies. In the Philippines, however, party lists consist of the elite and non-

elite, and their ideologies are nearly indistinguishable (Rocamora, 2002). What sets apart Philippine party lists instead are the top brass's personalities, which are treated like celebrities and have become a source of corruption (Quimpo, 2007b).

Nacionalista Party

The Nacionalista party is the earliest political party in the Philippines, founded by Manuel Quezon and Sergio Osmeña. Their main ideology centers on Philippine independence by focusing on developmental aspects such as education, social justice, and democracy (*Platform of the Nacionalista Party*, 1941).

The most discussed topics under the Nacionalista party list are Agriculture and Veteran official programs and reparation, as seen in Table 3. While these two topics broadly fall under Nacionalista’s theme of independence, they must be more prevalent in other SONAs to be considered unique to the party list. The president most faithful to the Nacionalista ideal is Marcos Sr., who often focused on the current state of economics, manpower, and labor. His known contribution is the development of numerous public infrastructures, which are still utilized today (*Edifice Complex*, n.d).

Table 3

Similarities of the presidents based on partylist

Nacionalista	Liberal
Agricultural Production	Agricultural Production
Veteran Officials Programs and Reparation	Tobacco
The current state of Economics, Manpower, and Labor	US Assistance
	Economic Report

The results of the Cosine Similarity corroborate with the LSA topic assignments (see Figure 5(a)), as the similarity scores indicate that most SONAs are similar. Most of the values in

the heatmap are below 0.10, except Quezon's fourth SONA and Magsaysay's first SONA, whose distance values fall mostly above 0.10. These can be justified as the topics assigned to each SONA are similar.

Liberal Party

The Liberal Party of the Philippines is a political party founded by Manuel Roxas; it originally stemmed from the "Liberal Wing" of the Nacionalista party and focused on the Filipino people. Specifically, they stand by the values of justice, freedom, and solidarity or "Bayanihan" (*Values Charter, 2017*).

Like the Nacionalista party, the most talked about topic is still Agriculture. Other topics include tobacco, US assistance, and economic reports. None of these topics align with the party list's ideology of solidarity or "Bayanihan." They are instead more in line with the Nacionalista party's ideas for development.

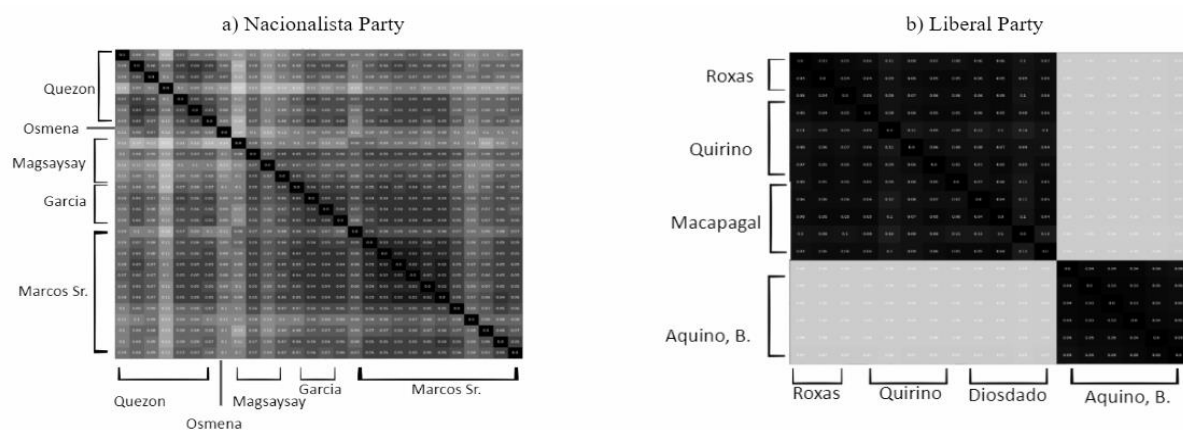
Benigno stood out from the rest of the group as he focused on employment and veteran programs instead of agriculture like the rest of the SONA under the party list, which is consistent with the focus of his administration: the economic development, and education (Mirasol, 2021).

The results of the Cosine Similarity once again corroborate with the LSA assignments, as shown in Figure 5(b). The similarity scores between the SONAs of earlier presidents Roxas, Quirino, and Macapagal are low, meaning they are similar. On the other hand, all of Benigno's SONAs are similar to the rest of the group despite having the same party affiliation. Despite their SONAs being similar to each other, except for Benigno Aquino III, the topics discussed in each do not reflect the values of the Liberal party, which once again corroborates the finding that party lists do not represent a group's ideals. Presidents who are more appropriate for the Nacionalista are Quirino's administration which focused on economic reconstruction and

restoring the faith of the people in the government (*Philippine Presidents*, 2010), and Macapagal's administration, which focused on the economy with the addition of fighting corruption.

Figure 5

Cosine Similarity of the two-party lists with multiple members



Comparison of SONAs based on Eras

Looking into the SONAs from the different eras may show the differences or similarities between the pre-authoritarian eras, the authoritarian era, and the post-authoritarian eras.

Commonwealth

This is when the Philippines is securing its independence from American sovereignty (*The Commonwealth of the Philippines*, n.d.). This era's SONA started with the first SONA of President Quezon (1935) and ended with the first SONA of President Manuel Roxas (1946).

Similar topics between the presidents are presented in Table 4. For Quezon and Roxas, they focused on agriculture and programs for veterans and compensation for damages. Quezon contributed to Philippine agriculture through significant laws like the National Rice and Corn Corporation (NARIC) in 1936, which established rice and corn prices, helping poor consumers (*Agrarian Reform History*, 2021).

The SONAs with the highest similarity are Quezon’s second, fifth, sixth, and seventh SONAs, ranging from 0.04 to 0.05. The results of the cosine similarity can support the derived topics. The SONAs with the highest similarity have Topic 0, which explains their similarity (see Figure 6(a)).

Table 4

Similarities of the presidents based on eras

Commonwealth; Third Republic; Martial Law; Fourth Republic; Fifth Republic	Pandemic Period
Agriculture	Covid Pandemic and Employment

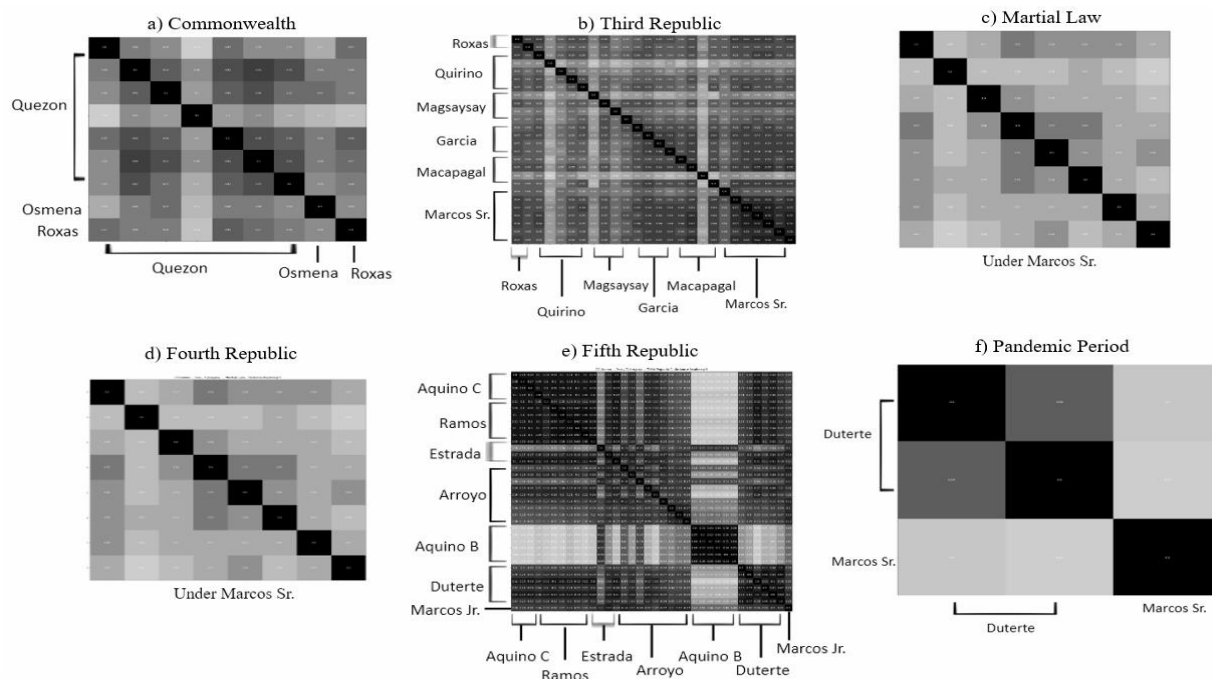
Third Republic

The Third Republic is when the global community officially recognizes the Philippines as an independent nation. The SONAs of this era started with the second SONA of Roxas (1947) and ended with the seventh SONA of Marcos Sr. (1972) (*Declaration of Martial Law*, n.d.).

Most presidents contributed to Philippine agriculture since Topic 0 was the most prevalent. Some of the agricultural contributions of the presidents are: For Roxas, he created a message about the proposed amendments by the Agrarian Commission, which studied the conditions of rice regions (*Message of President Roxas on Agrarian Reforms*, 1946). Quirino helped agricultural workers by creating a minimum wage law that could help improve their living conditions (*Foreign Relations of the United States*, 1950). Magsaysay created laws providing lands to farmers and corporations, providing housing lots and farmlands to rebel returnees. Garcia continued Magsaysay’s agrarian programs. Macapagal created the Agricultural Land Reform Code, which ended shared tenancy in the Philippines. Lastly, Marcos Sr. created laws that created the Department of Agrarian Reform (DAR) (*Agrarian Reform History*, 2021).

Figure 6

Cosine Similarity between Presidents in different eras



Based on Figure 6(b), the SONA with the highest similarity belongs to Marcos Sr., with a score ranging from 0.02 to 0.04. Regarding dissimilarity, Macapagal’s third and Quirino’s second SONA are dissimilar, with a score of 0.14. There are only slight differences when comparing all the SONA of the presidents in the Third Republic era. For the SONAs with the highest dissimilarity, Macapagal’s third SONA and Quirino’s second SONA have different main topics. Macapagal focused on Plans for government officials, Topic 34, while Quirino is about Veteran official programs, Topic 8.

Martial Law

This era started when Marcos Sr. proclaimed that the Philippines was under Martial Law upon signing Proclamation No. 1081 on September 21, 1972, which allowed him absolute authority over the nation (*Declaration of Martial Law, n.d.; The Makings of a Constitutional*

Dictator, n.d.). The SONAs of this era start with Marcos' eighth SONA (1973) and ends with his fifteenth SONA (1980).

The main topic of his SONAs always includes Topic 0 or Agriculture. This is prevalent in his SONAs because he launched many projects about agriculture, like Masagana 99, which addresses the food shortages and enhances the rice production of the Philippines (*The Marcos Agrarian Reform Program*, n.d.). The SONAs with the highest similarity are Marcos Sr.'s eleventh, twelfth, and thirteenth SONAs, with a score of 0.05, which talked about Agriculture and Martial Law. While his other SONAs tackled the Societal and geographical groups and the resignation of officials (see Figure 6(c)).

Fourth Republic

During Marcos Sr.'s term, the 1973 Constitution suspended the 1935 Constitution. Upon doing so, he announced the "New Republic" in 1981, also called the Fourth Republic (*The Executive Branch*, n.d.). The SONAs of this era start with Marcos' sixteenth SONA (1981) and ends with his last SONA (1985).

The main extracted topic is still Agriculture. This could be explained by the number of agricultural projects launched by Marcos Sr (*The Marcos Agrarian Reform Program*, n.d.). Also, Topics 8, 43, and 45 may be connected as they discussed Kilusang Kabuhayan at Kaunlaran livelihood program, intended to benefit the Filipino people (*Official Week in Review*, 1981).

As presented in Figure 6(d), the SONAs with the highest similarity are Marcos Sr.'s nineteenth and twentieth SONAs, with a score of 0.05. Both had the same topic: Topic 0. When it comes to the dissimilar SONAs, which are Marcos Sr.'s seventeenth and twentieth SONAs, they have different main topics. The seventeenth SONA discussed Topic 47, while the twentieth SONA on Topic 0.

Fifth Republic

This era started when the 1987 constitution was put into effect. It marked the end of Marcos Sr.'s reign and Corazon's inauguration. The SONAs of this era start from Corazon (1987) up to Marcos Jr., who is the current president and has just given his first SONA (2022) (*Declaration of Martial Law*, n.d.).

Similar to previous periods, the most discussed topic still involves Agriculture. Some of the agricultural contributions of the presidents are: For Corazon, she signed Republic Act No. 6657, which is the Comprehensive Agrarian Reform Law (CARL) that helps landless farmers to have their lands (Official Gazette of the Republic of the Philippines, 1988). Ramos was committed to creating a better Agrarian Reform Program by enacting laws like Republic Act No. 7905, which strengthened the Comprehensive Agrarian Reform Program. Estrada launched the Magkabalikat Para sa Kaunlarang Agraryo (MAGKASAKA), where the DAR will work with private investors to find ventures in the agrarian sector. Arroyo also had her agrarian reform program, where a part of the program was the Land Tenure Improvement. Benigno created the Agrarian Reform Community Connectivity and Economic Support Services (ACCESS) project to reduce rural poverty. Lastly, Duterte sought an aggressive approach to the land reform program by prioritizing support services and land distribution. Looking at the topics, the presidents also had some similarities, as they addressed different topics.

Moreover, regarding cosine similarity, as presented in Figure 6(e), the SONAs with the highest similarity are all of Benigno's SONAs, with scores ranging from 0.03 to 0.08. Corazon's SONAs are similar to Ramos's SONAs, with scores ranging from 0.06 to 0.17 since both discussed Agriculture. While Benigno's SONAs have dissimilarities with almost all of the other presidents' SONAs.

Pandemic Period

The pandemic period changed the course of the presidents' focus, primarily observed in their SONAs. It is beneficial to compare the president's SONA during this period, considering that this is the current period in which this paper is written. It is necessary to analyze issues the presidents have addressed and are addressing because it may reveal the shift from different presidencies and how the pandemic affected their focus.

Duterte's SONA focused on the COVID pandemic and online mode, employment, and agriculture. During Duterte's term, he signed an employment recovery plan to recuperate from shifts in the labor market brought upon by the COVID pandemic (Philippine News Agency, 2021). The government needed to prepare for these changes since the pandemic suddenly started during his term, thus needing to adapt to the workforce changes.

In contrast, Marcos Jr. has focused on technology and modernization of the rice industry (Cueto, 2022; Unite, 2022). However, Marcos Jr.'s term started when the pandemic started to subside; hence he focused on other areas, such as technology and modernization, due to globalization. The traditional sectors, such as the agricultural sectors, can also benefit from modernization. The topics mentioned are in sync with the results of the cosine similarity. The most similar SONAs are the two SONAs of Duterte with a cosine distance of 0.09, as seen in Figure 6(f). Duterte's SONAs are also considerably similar to Marcos Jr.'s first SONA since they tackled the COVID pandemic and employment.

Conclusions

It is deduced that the different Philippines presidents have distinct personalities, plans, and goals regarding how they address their SONAs. Not to mention that different factors must be

considered when formulating their address, such as the current issues and their future perspective for the country.

Grounded by the result of the model, the objective questions are now answerable. On the comparisons of SONA based on categories starting on relationships, the similarities and dissimilarities depend on each family. Since even though they had similar connections, they had different outlooks or views during their time as president.

As for partylists, the results corroborate that they have indistinguishable ideologies as the SONAs in both the Nacionalista and Liberal parties focus on similar topics. Additionally, none of the topics in the SONAs represent their aligned ideology. This means that regardless of a president's party, their agendas for the country are similar, further supporting the personality-focused claim of Philippine partylists.

Furthermore, regardless of the era, the topic mentioned repeatedly is agriculture. The Philippines is an agricultural country, and the presidents tackle this as one of the best assets. The pandemic period explains a highly similar result due to the COVID pandemic and employment as there is a transition from the physical to online, shifting the workforce. However, Marcos Jr. also concentrated on technology.

The results prove that the LSA model and Cosine Similarity effectively analyze the similarities and differences of documents such as SONAs. Also, determining similarities and dissimilarities between the presidential SONAs can provide insights into the patterns while identifying what issues have been solved or recurring.

Limitations and Recommendations for Future Researchers

The study's main limitation is that it only focused on analyzing the transcripts of SONAs, which could be more descriptive of a president's intentions. Future researchers should

incorporate other sources, such as the reactions of the Senate, to get a broader perspective. Less personal documents can also be considered.

Future researchers should improve the LSA model since some topics are difficult to distinguish using their top words for better interpretations. Lastly, it would be beneficial to seek the insights of experts such as political analysts, historians, and other professionals who possess specialized knowledge in the SONAs to validate and gain further understanding of the context surrounding it, leading to more comprehensive interpretations.

References

Agrarian Reform History. (2021). Department of Agrarian Reform.

<https://www.dar.gov.ph/about-us/agrarian-reform-history/>

Aquino, A. & Chavez, E. (2018). Analysis of the use of Latent Semantic Indexing (LSI) for document classification and retrieval system of PNP files. *MATEC Web of Conferences*.

<https://doi.org/10.1051/matecconf/201818903009>

Bergamaschi, S. & Po, L. (2015). Comparing LDA and LSA Topic Models for Content-Based Movie Recommendation Systems. *Lecture Notes in Business Information Processing*

226. https://doi.org/10.1007/978-3-319-27030-2_16

Bringula, R. & Miranda, J. (2021). Exploring Philippine Presidents' speeches: A sentiment analysis and topic modeling approach. *Cogent Social Sciences*, 7(1).

<https://doi.org/10.1080/23311886.2021.1932030>

Cinelli, M., Ficcadenti, V., & Riccioni, J. (2021). The interconnectedness of the economic content in the speeches of the US Presidents. *Annals of Operations Research*.

<https://doi.org/10.1007/s10479-019-03372-2>

- Cueto, F. (2022). *Marcos expected to back the AFP modernization plan*. The Manila Times.
<https://www.manilatimes.net/2022/07/23/news/national/marcos-expected-to-back-afp-modernization-plan/1851908>
- Declaration of Martial Law | GOVPH*. (n.d.). Official Gazette of the Republic of the Philippines.
<https://www.officialgazette.gov.ph/featured/declaration-of-martial-law/>
- Dohner, R. & Intal, P. Jr., (1989). The Aquino Government and Prospects for the Economy. *In National Bureau of Economic Research*. <https://www.nber.org/books-and-chapters/developing-country-debt-and-economic-performance-volume-3-country-studies-indonesia-korea/aquino-government-and-prospects-economy>
- Edifice Complex: Building on the Backs of the Filipino People*. (n.d.). Martial Law Museum.
<https://martiallawmuseum.ph/magalar/edifice-complex-building-on-the-backs-of-the-filipino-people/>
- Elpidio R. Quirino (1948-1953)*. (2010). Philippine Presidents.
<https://bloomspressidents.wordpress.com/elpido-r-quirino/>
- Foreign Relations of the United States, 1950, East Asia and the Pacific, Volume VI - Office of the Historian*. (n.d.). <https://history.state.gov/historicaldocuments/frus1950v06/d852>
- Gaje, V. (2010). *Aquino turns over housing units to Ondoy, Pepeng victims*.
<http://archives.pia.gov.ph/?m=12&sec=reader&rp=10&fi=p101015.htm&no=98&date=10/15/2010>
- Gouri, V. & Uddin, M. (2019). An empirical study of big data analytics in the rural arena. *International Journal of Social Sciences & Educational Studies*, 5(3), 192–206.
<http://eprints.tiu.edu.iq/22/1/ijsses.v5i3p192.pdf>

- Gunawan, D., Sembiring, C., & Budiman, M. (2018). The Implementation of Cosine Similarity to Calculate Text Relevance between Two Documents. *Journal of Physics*, 978, 012120. <https://doi.org/10.1088/1742-6596/978/1/012120>
- Hoblos, J. (2020). Experimenting with Latent Semantic Analysis and Latent Dirichlet Allocation on Automated Essay Grading. *2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS)*. <https://doi.org/10.1109/snams52053.2020.9336533>
- Joint Press Statement between the Republic of the Philippines and the Republic of Singapore on Strengthening the Philippines-Singapore Partnership.* (2022). <https://www.mfa.gov.sg/Newsroom/Press-Statements-Transcripts-and-Photos/2022/09/20220907jpcpsp>
- Idowu, T. (2015). *Qualitative and quantitative research approaches*. Academia. https://www.academia.edu/15209462/Qualitative_and_Quantitative_research_approaches
- Lahitani, A., Permanasari, A. , & Nugroho, H. (2016). Cosine similarity to determine similarity measure: A study case in online essay assessment. *2016 4th International Conference on Cyber and IT Service Management*. <https://doi.org/10.1109/citsm.2016.7577578>
- Lineses, E. (2018). *Policy Types and Democratic Governance Values in Presidential Speeches: Content Analysis of Philippine State of the Nation Addresses from 1973 to 2015*. <https://space.snu.ac.kr/handle/10371/144045?mode=full>
- McCoy, A. (1994). An Anarchy of Families State and Family in the Philippines. *Philippine Political Science Journal*, 18(1–2), pp. 149–153. <https://doi.org/10.1163/2165025x-0180102007>

- McKenzie, G., & Adams, B. (2020). A country comparison of place-based activity response to COVID-19 policies. *Applied Geography*, p. 125, 102363.
<https://doi.org/10.1016/j.apgeog.2020.102363>
- Mendoza, R., Banaag, M., Hiwatig, J., Yusingco, M. & Yap, J. (2020). *Asia-Pacific Social Science Review*, 20(4), 88–99. <https://www.dlsu.edu.ph/wp-content/uploads/pdf/research/journals/apssr/2020-December-vol20-4/8-Term-Limits-and-Political-Dynasties-in-the-Philippines-Unpacking-the-Links.pdf>
- Mendoza, R., Beja, E., Venida, V., & Yap, D. (2013). Political dynasties and poverty: Resolving the “chicken or the egg” question. *SSRN Electronic Journal*.
<https://archium.ateneo.edu/asog-pubs/101>
- The message of President Roxas on Agrarian Reforms, August 8, 1946, | GOVPH.* (1946). Official Gazette of the Republic of the Philippines.
<https://www.officialgazette.gov.ph/1946/08/08/message-of-president-roxas-on-agrarian-reforms-august-8-1946/>
- Mirasol, P. (2021). *A look back at the PNoy administration*. BusinessWorld Online.
<https://www.bworldonline.com/the-nation/2021/06/24/378119/a-look-back-at-the-pnoy-administration/>
- Official Week in Review: September 21 – September 27, 1981, | GOVPH.* (1981). Official Gazette of the Republic of the Philippines.
<https://www.officialgazette.gov.ph/1981/09/28/official-week-in-review-september-21-september-27-1981/>
- Philippine News Agency. (2021). *Duterte signs employment recovery agenda*.
<https://www.pna.gov.ph/articles/1145306>

- The platform of the Nacionalista Party, August 16, 1941, | GOVPH.* (1941). Official Gazette of the Republic of the Philippines. <https://www.officialgazette.gov.ph/1941/08/16/platform-of-the-nacionalista-party-august-16-1941/>
- Purdey, J., Tadem, T., & Tadem, E. (2016). Political dynasties in the Philippines. *South East Asia Research*, 24(3), 328–340. <https://doi.org/10.1177/0967828x16659730>
- Quimpo, N. (2007a). The Philippines: Political Parties and Corruption. *Southeast Asian Affairs*, pp. 277–294. <http://www.jstor.org/stable/27913337>
- Quimpo, N. (2007b). Trapo Parties and Corruption. *KASAMA*, 21(1). <https://cpcabrisbane.org/Kasama/2007/V21n1/TrapoPartiesAndCorruption.htm>
- Republic Act No. 6657 | GOVPH.* (1988). Official Gazette of the Republic of the Philippines. <https://www.officialgazette.gov.ph/1988/06/10/republic-act-no-6657/>
- Rocamora, J. (2002). *Philippine Political Parties, Electoral System and Political Reform.* <https://www.semanticscholar.org/paper/Philippine-Political-Parties-%2C-Electoral-System-and-Rocamora/7914077e4c905cb3398b6f970d0c4baff1aa4387>
- Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the Space of Topic Coherence Measures. *Web Search and Data Mining*. <https://doi.org/10.1145/2684822.2685324>
- Sicat, G. (2012). *The Sona and its messages.* Per Se. <https://econ.upd.edu.ph/perse/?p=1254>
- Starner, F. (1963). The Philippines: Politics of the “New Era.” *Asian Survey*, 3(1). <https://doi.org/10.2307/3024649>
- The Commonwealth of the Philippines | GOVPH.* (n.d.). Official Gazette of the Republic of the Philippines. <https://www.officialgazette.gov.ph/the-commonwealth-of-the-philippines/>
- The Executive Branch | GOVPH.* (n.d.). Official Gazette of the Republic of the Philippines. <https://www.officialgazette.gov.ph/about/gov/exec/>

The Makings of a Constitutional Dictator. (n.d.) Martial Law Museum.

<https://martiallawmuseum.ph/magalar/the-makings-of-a-constitutional-dictator/>

The Marcos Agrarian Reform Program: Promises and Contradictions. (n.d.). Martial Law

Museum.<https://martiallawmuseum.ph/magalar/the-marcos-agrarian-reform-program-promises-and-contradictions/>

Third Republic | GOVPH. (n.d.). Official Gazette of the Republic of the Philippines.

<https://www.officialgazette.gov.ph/featured/third-republic/>

Unite, B. (2022). *Marcos bats for rice industry modernization.* Manila Bulletin.

<https://mb.com.ph/2022/11/29/marcos-bats-for-rice-industry-modernization/>

Valdez, D., Pickett, A., & Goodson, P. (2018). Topic Modeling: Latent Semantic Analysis for the Social Sciences. *Social Science Quarterly*, 99(5), 1665–1679.

<https://doi.org/10.1111/ssqu.12528>

Values Charter. (2017). Liberal Party of the Philippines. <https://liberal.ph/values-charter/>