Article

The North Korean Economy: Sustainable or Muddling-Through?

Soohyon Kim [1] , Wook Sohn [2,]* and Changyong Choi [2]

[1]   Section for New Digital Technology, Bank of Korea, Seoul 04514, Korea; soohyonkim@bok.or.kr
[2]   KDI School of Public Policy and Management, Sejong 30149, Korea; cychoi@kdischool.ac.kr
*   Correspondence: wooksohn@kdischool.ac.kr

check for updates

Abstract: Confronting the continuing economic sanctions, the flood disaster, and the COVID 19, North Korea is now in a critical phase whether the North Korean economy will shift to a sustainable economy or, as it has done in the past three decades, barely manage to maintain at a survival level. To explore the question, this study investigates North Korea's economic policy changes by extracting the keywords from the North Korean economic journal *Economic Research*, which is a representative state published document. We use a text mining approach to overcome limitations, such as the lack of objectivity and accuracy of the classification of topics in manual text analysis for North Korean studies. Contrary to the conventional assumption that the North Korean economic policies are driven by dogmatic *Juche* (self-reliance) ideology, we find that the strategy of economic policies differs depending on the leader and political and economic circumstances at the time. In particular, since Kim Jong-un assumed office in 2011, the topics have diversified, and now cover a wide range of subjects on capitalism. Nevertheless, without North Korea's decisive reform measures and international cooperation, it will be hard to observe that the North Korean economy can go beyond 'muddling-through' for sustainable development.

1. Introduction

North Korea's leader Kim Jong-un held the 5th plenary session of the 7th Central Committee in December 2019 and presented a 'new path' that calls for self-reliance on the economy, political and military stability, and party control as key elements. However, at the plenary session of the 7th Party Central Committee, which was held only half a year later, Kim Jon-un officially recognized the 'poor economic performance' and to 'adjust and change' some of the policy tasks decided at the 5th meeting, and to hold another evaluation meeting on economic performance at the 8th Central Committee meeting in January 2021. The statement of the party's plenary session pointed out that "(the) internal and external situation has continued and unexpected challenges have failed to improve the economic situation, failed to meet the planned growth target of the national economy, and failed to improve the lives of residents."

As Kim Jong-un himself mentioned, the internal moves in the North confirm the challenges faced by the North Korean authorities in triple trouble: The continuing economic sanctions, the flood disaster, and the COVID19. It is an important turning point that shows whether the North Korean economy will shift beyond recovery to a sustainable economy or, as it has done in the past three decades, barely manage to maintain at a minimum level of survival.

To be sure, the North's economic difficulties are not new, and they are the result of the accumulation of structural problems accumulated over the past decades. Both the death of Kim Il-sung in 1994 and

the series of events that led to the collapse of the socialist bloc in the 1990s brought unprecedented economic hardships to the Democratic People's Republic of Korea (DPRK. Hereafter, North Korea); the North Korean economic situation at that time was likened to that of the *Arduous March* during Japanese colonialism. Kim Jong-il, then leader, used military-first policies to stimulate a military-first economy to overcome the unprecedented economic crisis, but the North Korean economy failed to recover, and in 2011, Kim Jong-un became the next leader [1–4]. In his 2012 New Year's joint editorial, shortly after his succession to power, Kim Jong-un declared his future course of economic rehabilitation and nuclear weapons development, stressing that, '(We) will continue to walk our own path of self-reliance, strong military power, and socialism' and that he will further consolidate the position of the political thought of revolution by 'building a strong socialist nation'.

However, the economic sanctions imposed by the international community after it conducted nuclear tests, and the adherence to the socialist economy, pose respective external and internal constraints on North Korea's economic regeneration. It is not likely to be easy for the North to achieve the 'five-year national economic development strategy (2016–2020)' (hereafter, New Strategy) announced by Kim Jong-un in 2016. Nevertheless, the Kim Jong-un regime is also pursuing a series of changes that were promoted by previous leaders Kim Il-sung and Kim Jong-il in the event of a national crisis. The new economic policy direction and discourse that has emerged under the Kim Jong-un regime, in particular, are well worth noting in the study of North Korea's economic policy.

Against this backdrop, this present paper attempts to explore the following research questions. First, what are the specificities of a change in economic strategy and economic policy within North Korea? This allows us to gauge the content, level, and speed of the North Korean government's 'change' or 'reform'. Second, if the outline of North Korea's current development strategy is understood, what are the specific policies for achieving such development strategies or challenges ahead for North Korean for its sustainable development beyond muddling-though or survival from the current economic hardships? What, in particular, are the characteristics and changes of sectoral policies? This highlights that policy changes based on the leader, and circumstances are also significant. Third, to what extent can a deep analysis of first-hand literature be produced in North Korea in relation to the research methodology? As the literature produced in socialist countries is often contaminated with excessive political rhetoric and revolutionary slogans that are alienated from reality, there is a limit to what can be used in a meaningful analysis. North Korea's official publications are not free from such content. Despite these restrictions, an in-depth analysis of primary data produced in North Korea, particularly the economic journal *Economic Research*, is essential for tracking the North Korean economic authorities' perception of reality, limitations, and alternatives.

Since the North Korean economy is inevitably influenced by a single leader, the articles published in *Economic Research* can be regarded as reflecting the political interests of the leader. Thus, the articles can be classified, and subjects can be extracted based on the characteristics of these leaders. In particular, we find that, since Kim Jong-un assumed office in December 2011, the number of publications has increased, and the writers and topics covered by the journal have diversified. By covering a wide range of subjects related to capitalism, the 'internal learning' of the capitalist economy has been widely conducted. In addition, keywords were expected at the end of each article from 2017. Another notable change is the increase in papers covering a variety of topics using econometric models, as well as ideological descriptions. Specifically, recent issues require very specific knowledge and techniques, for example, factors that affect hedge accounting for derivatives transactions, exchange rate forecasting and the management of statistics, methods to create a national foreign exchange balance of payments, or methods to determine expected patient tax returns. After the New Strategy was declared in 2016, along with the end of the Nuclear Economic Pathway, Kim Jong-un's economic strategy is believed to foster economic development in the face of the international economic sanctions. This is evidenced by such phrases as 'the knowledge economy', 'an industrial revolution in the new century', 'streamlining business management' and 'entering international trade markets' in the literature.

This paper is organized as follows. Section 2 outlines the text mining techniques, and how they are employed as research methodologies for this study. In Section 3, we attempt a context analysis of the changes in North Korea's economic policies, which were detected by analyzing similarities and correlations between the keywords extracted through morphological analysis. In Section 4, we discuss the Kim Jong-un regime's economic strategy and economic policy, and sustainability of the North Korean economy. Section 5 concludes with key findings and implications of this study.

2. Analysis of North Korean Literature

In this study, the title and keywords of a paper in *Economic Research* is an analytical unit that can track key agenda and policy changes that North Korea has emphasized for its economy. In particular, this research investigates whether any changes in economic policies introduced and implemented by Kim Il-sung, Kim Jong-il, and Kim Jong-un can be classified by the characteristics of the keywords used in the titles at the time of their respective leadership. This investigation expects to help us comprehend North Korea's strategic directions, as well as its challenging issues for the country.

Artificial intelligence text mining and machine learning, which can analyze meaningful results from a limited number of primary data, have been actively utilized recently. However, North Korea-related research and national statistics are severely limited in terms of accessibility and credibility, thus, being the biggest obstacle in this study. The main significance of this study is that, despite the limited data, results can be found with advanced text mining techniques. Additionally, the significance of this study is clear in the comparative study of different governance systems. Big data that is composed of text mining can also be analyzed objectively by using machine learning techniques.

2.1. North Korean Studies on Economic Policy

Outsiders can use a three-stage process to observe the performance of North Korean policy. The first stage is referred to as the planning phase, in which official propaganda and state media released by North Korean authorities, such as *Rodong Sinmun*, a daily newspaper in North Korea, and *Economic Research*, confirm that North Korean officials have reshaped or revised their policies. The second stage is the implementation of the policy, and is reflected in the actual application of the new policy. Due to the limitations of intelligence on North Korea, the actual execution is not clear. However, it is possible to verify the actual policies through the testimonies of North Korean defectors. The third stage is a performance realization step at which outsiders identify whether or not new policies result in different outcomes. It is quite challenging for outsiders to measure the actual policy effects, but partial confirmation is possible through evaluating the data compiled by international organizations, such as the Food and Agricultural Organization, the World Food Program, or the United Nations Development Program. Taking agricultural reform as an example, the first stage can be conjectured by reports and papers in *Rodong Sinmun* and *Economic Research*, respectively, on the agricultural reform. The second stage can be reflected in the agricultural reform being applied to cooperative farms in North Korea, and the third stage can be assessed using the data compiled by international organizations, whether or not actual agricultural production increased, due to the agricultural reforms. This study focuses on the first stage using economic papers officially published in Pyongyang.

Concerns about data shortage, accessibility, and credibility have always been raised in studies about the North Korean economy. Due to the nature of economics, the lack of statistics, in particular, inhibits the ability to analyze reforms and clearly serves as a limitation to any research. Despite these limitations, however, some scholars have attempted to estimate macroeconomic statistics on North Korea. Kim et al. [5], for example, calculated North Korea's gross national product and gross domestic product growth rates using official statistics from North Korea and South Korea. Some other studies explored changes in economic policies of the North Korean regime through state publications—although, these studies were conducted mainly by South Korean scholars who are able to access and read North Korean documents [6–10]. Choi and Lecy [6] divided their research into four periods (1986–1993, 1994–2001, 2002–2005, and 2006–2009), dependent on the policy direction, and presented a map of

the inter-topic correlations shown in *Economic Research*. Park Jong-hui et al. [7] examined the subject changes of North Korea's New Year's speeches with automated text analysis techniques. Lee [8], Lim [9], and Na [10] also conducted in-depth text analyses of the primary literature, and publications and reports by the state media. In addition, Choi et al. [11] compared the industrial policies presented under the Kim Il-sung and Kim Jong-il administrations with the industrial policies emphasized by Kim Jong-un by conducting economic research on the *Rodong Sinmun*. Admittedly, these studies expanded upon the academic approaches and methodologies on North Korean studies, but they had limitations, such as the lack of objectivity and accuracy of the classification of topics, as well as the small number of documents that researchers could directly access and categorize.

2.2. Text Mining Approach

Though text analysis can be done manually, a more sophisticated analysis is possible using text mining techniques with computer algorithms. Text mining includes converting unstructured data and text into structured data for a computer to automatically analyze [12]. The translated text takes the same form as conventional big data, and by using machine learning techniques, researchers are able to analyze those reconstructed texts in various ways. There are limited cases in which text analysis utilizing in-depth text mining techniques are applied to North Korean literature. Park et al. [7], for instance, divided North Korea's New Year's messages by year to investigate annual changes in the direction of North Korea's policy by classifying similarities among the subjects that were extracted from keywords. Rich [13,14] analyzed the attitude of the six-party talks and U.S. foreign policy on the North Korean denuclearization issue in an English version of news reports released by the Korean Central News Agency. Whang et al. [15] compared patterns in the reports using a method of supervised learning, and examined military offenses to predict when North Korea would launch a small-scale military provocation.

There are many difficulties in analyzing North Korean texts. First of all, after more than 70 years of separation, the Korean language spoken and written in North Korea has become very different from that in South Korea. For example, there are different spellings, spacing standards, and double-tone rules. Second of all, certain words that are no longer used in South Korea are still used in North Korea, such as *Gurapa* for Europe or *Rossia* for Russia. Thus, the difference between the two Koreas makes morphological analysis—the most fundamental form of text analysis—difficult. Therefore, words that are not recognized by the morpheme analyzer tend to be treated as missing data. (A morpheme analyzer is built by an automated algorithm to learn a large volume of existing literature. A morphological analyzer for the North Korean language has not yet been established because of the lack of electronic data from North Korea).

To solve this problem, this study replaced the unique spelling and words of North Korea with similar words that can be analyzed. This is a method that has not been introduced in previous studies. In the absence of a morpheme analyzer for North Korean, replacing words that are not recognized by the morpheme analyzer with the closest alternative words can prevent any bias in the data, due to missing data. In fact, automated document analysis often results in a completely different outcome, even if only a few morphemes or words are missing. In particular, the accuracy of the analysis results can be greatly reduced if an unrecognized word is one that carries a significant amount of information instead of being one that is used habitually.

2.3. Economic Research, an Economics Journal in North Korea

The main texts for this study were extracted from North Korea's economics journal *Economic Research*. The period covered is between 1988 and 2018, which spans the ruling periods of Kim Il-sung, Kim Jung-il, and Kim Jong-un. The journal is a quarterly publication focusing on related economic issues, and was first published in 1956. The papers in *Economic Research* are written to discuss and disseminate North Korea's economic policies and future goals, as well as summarize international economic situations, not make academic contributions. *Economic Research* provides insight

into the changes in North Korea's policies, institutions, and practices regarding economic policies, trade, and international financial issues. (Numerous academic journals are published in North Korea. They include *Kyongje Yongu* (Economic Research), *Kimilseng Jonghaptayhak Hakpo* (Academic Journal of Kim Il-sung University), *Yeogsa Gwahag* (Study of History), *Chosun Uihak* (Journal of Chosun Medicine), and *Cheolhag Yeongu* (Study of Philosophy). Among these journals, *Economic Research* is the only source of updated information about North Korea's economic policies and institutions).

While the New Year's address reveals the leader's intentions with regard to economic policies and goals, *Economic Research* represents a specific roadmap that is presented by experts across various sectors of the economy and written under the North Korean guidelines. Media outlets, such as the *Rodong Sinmun* and the *Chosun Shinbo*, share the intentions of the highest-ranking government officials to the public, but are not regarded as being as professional or as specific to economic issues as is *Economic Research*. In addition, the value and reputation of *Economic Research* as a professional journal in Pyongyang stems from its ability to disseminate various policy alternatives and economic policies of North Korea. To analyze the changes in economic policy topics in *Economic Research*, this paper divided the analysis period into three time periods: Kim Il-sung (1988–1994), Kim Jong-il (1995–2011), and Kim Jong-un (2012–2018). The reason for dividing the target period by the 'supreme leader' is that, firstly, there is no consensus in the academic community as to the period of division. Secondly, the distinction between periods under a specific policy agenda is expected to show whether a new regime enjoys any autonomy from its previous leader(s)' regimes, as North Korean people and government officials, who live under totalitarianism, have no choice but to follow the leader's instructions and guidance. Therefore, it is deemed appropriate to analyze the changes in policy subject matter by dividing the period by the leader.

Kim Jong-il appointed Kim Jong-un as his successor in 2009. Kim Jong-un was then appointed the Commanding General of the Korean People's Army in 2010, showing that a process of actual power inheritance started to take place. Table 1 shows that the number of papers published every year since 2009 increased by 10, and reached over 130 in 2012, reflecting a change in the composition of *Economic Research*.

Table 1. Number of papers published in *Economic Research* by year.

Periods	Year	No. of Papers
Kim Il-sung	1988	27
	1989	49
	1990	64
	1991	58
	1992	64
	1993	60
	1994	47
	subtotal	338
Kim Jong-il	1995	61
	1996	59
	1997	57
	1998	58
	1999	63
	2000	66
	2001	69
	2002	71
	2003	85
	2004	79
	2005	81
	2006	85
	2007	84

Table 1. *Cont.*

Periods	Year	No. of Papers
	2008	83
	2009	94
	2010	103
	2011	117
	subtotal	1346
Kim Jong-un	2012	133
	2013	139
	2014	139
	2015	139
	2016	126
	2017	180
	2018	172
	subtotal	1073
Total		2757

There is a kinship between the three leaders, who each inherited his predecessor's political and ideological ideas. For example, authentic socialist principles, such as 'self-reliance', based on the *Juche* ideology, and the 'strategic management by the state of the construction of a socialist economic powerhouse' are consistently emphasized. Nevertheless, clear distinctions between the leaders are observed from time to time. For instance, Kim Il-sung's main policy goals were to strengthen the rural economy under the socialist ideological framework, and establish an industrial base centered on small- and medium-sized businesses. Kim Jong-il introduced the capitalist system in North Korea during a rivalry with the capitalist South. Capitalism still exists under Kim Jong-un's leadership, but the views are less critical than they were under previous leadership. Remarkably, Kim Jong-un emphasized partial acceptance of the capitalist system for the sake of economic revival, including science, information, a knowledge-based economy and trade, and the so-called industrial revolution in the new century.

Furthermore, academic journals reflect different tones depending on the leader at the time of publication. For example, during the leadership of Kim Jong-il in 1995, a paper containing critical details of the capitalist system was published at the end of every quarter. In addition, the column entitled *Sangsik* ('Common Sense' in English), which has been included at the end of each issue since 2010, is a structure designed to introduce basic economic concepts, such as knowledge, labor, the balance of payments, interest rates and depreciation. Consequently, some papers contain introductions, as well as criticisms about the market economy and liberal economics.

3. Main Results from *Economic Research* Using Text Mining Approach

In this section, the titles of the articles in *Economic Research* are analyzed to verify whether the characteristics shown in the titles of the paper can be classified by leader or time. After the text is converted into structured data, machine learning is used to analyze and validate text as the dimensions in the dataset increase, and it, thus, becomes big data. This process represents the probabilistic or non-probabilistic word representation of a document or word as a vector. In this section, we mobilize a term frequency-inverse document frequency (tf-idf) process, a method of non-probabilistic representation, and a machine learning classification technique to break down the main themes of economic policy in the title by time, subject, and keyword [12].

First, we look at how year-by-year article titles are clustered by subject by using unsupervised learning without prior information on the characteristics of the title. If a cluster of titles is consistent within the same time span, it can be implied that the focus and policy direction of the economy depends on the leader at that time. Second, it predicts when individual papers were written by using supervised learning, and reviews them to see if individual titles can be classified according to the characteristics.

Knowing when a paper was published could provide further evidence of the distinct characteristics of the economic policy direction in the concerned time period. Third, we provide a visualized frequency of the keywords of the thesis titles, depending on each leader's period. We also look at the major topic changes between each period using a latent Dirichlet allocation (LDA), which extracts the subject from the Bayesian probability process. Figure 1 displays our research process in a flowchart.
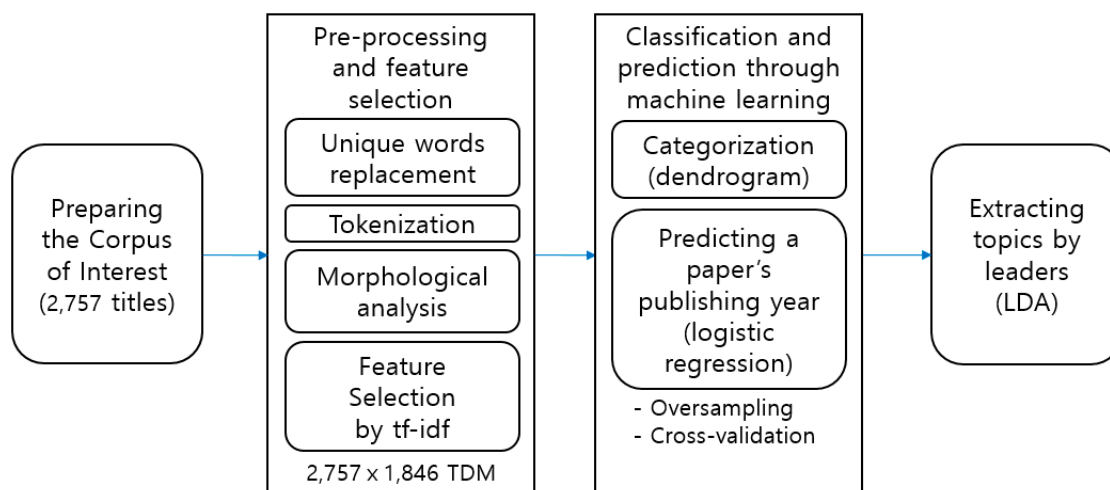

Figure 1. Research process.

3.1. Keyword Correlation Coefficient for Each Year

The information characteristics of the data are tested before using tools, such as machine learning. Firstly, the title words in *Economic Research* for each year are represented in a Term-Document Matrix (TDM), and secondly, the correlation coefficient of the title of the document by year is calculated (see Figure 2). If the value of the coefficient in the title for each year is 0.6 or higher, the topics covered in *Economic Research* can be considered to have a high level of similarity. However, there are limitations to only using the value of the correlation coefficient to analyze how the topics of interest in economic policies change over time. For example, the large number of keywords commonly used in the title of the papers in the journal make using only this method inefficient. Therefore, it would be a more efficient method of analysis to determine and classify the weights of common words using a tf-idf process.

3.2. Categorization of the Title of Each Year's Thesis

Applying unsupervised learning to draw inferences from titles that are not randomly classified in advance is the first step in determining whether information can be extracted from the characteristics of the year-by-year thesis. There are many different methods of unsupervised learning taxonomy, but the algorithm used in this study is simple and intuitive; the hierarchical clustering method, which can group data that have not been classified according to interconnections, is used. In summary, it can be inferred that, when the title of the annual paper is clustered according to shared attributes, and the cluster coincides with the leadership period of a leader, the characteristics are shown in the title represent a change in the economic policy made by the leader.

Figure 3 is a dendrogram that classifies the title of the paper for each year from 1988 to 2018, generated using a hierarchical clustering method. The paper titles by year are clustered by the time periods of each leader: Kim Il-sung (1988–1994), Kim Jong-il (1995–2011), and Kim Jong-un (2012–2018). (In the hierarchical clustering method used in this study, the titles of articles published in *Economic Research* are classified by year, but no information about the leader of the country is added. Nevertheless, hierarchical clustering is performed relatively precisely against the years the respective leaders were in power). Considering the economic conditions and changes in the economic policies in North Korea, the results of the clustering show a distinct pattern.
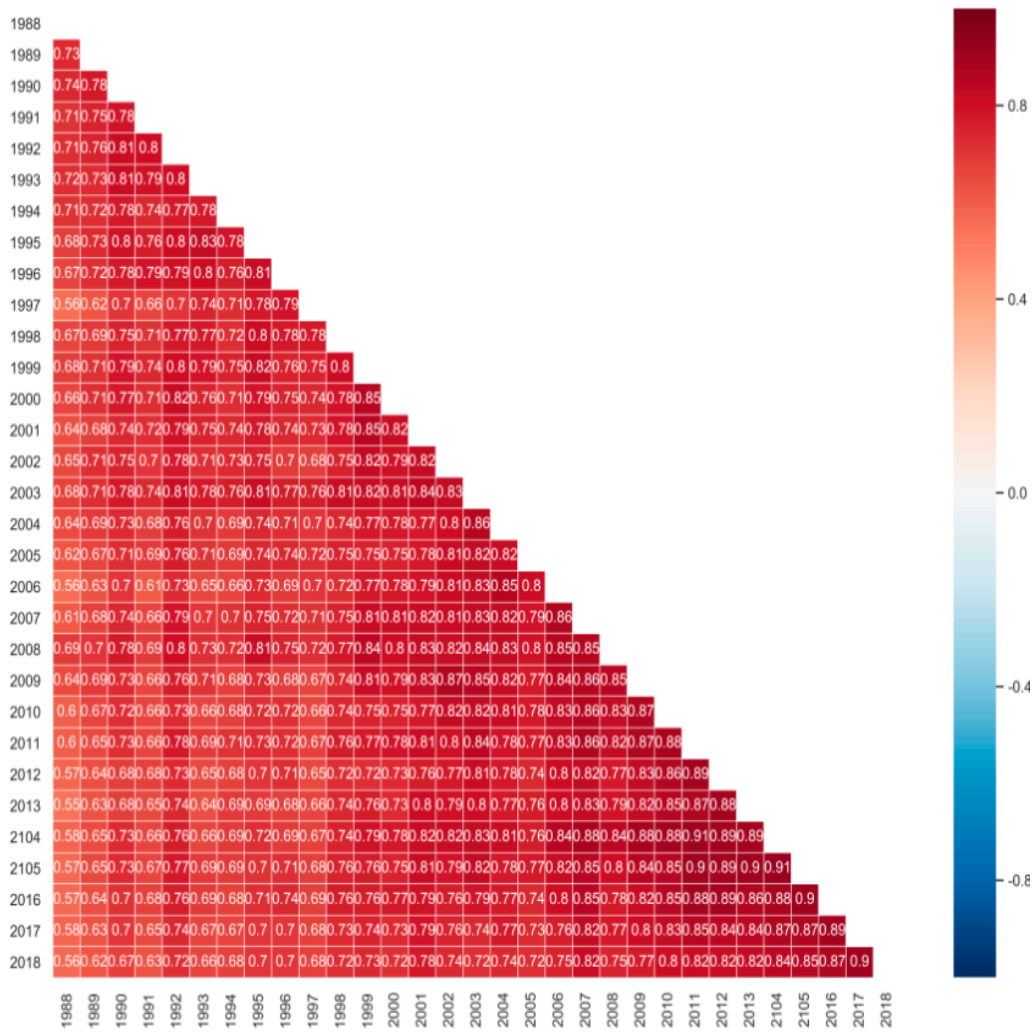
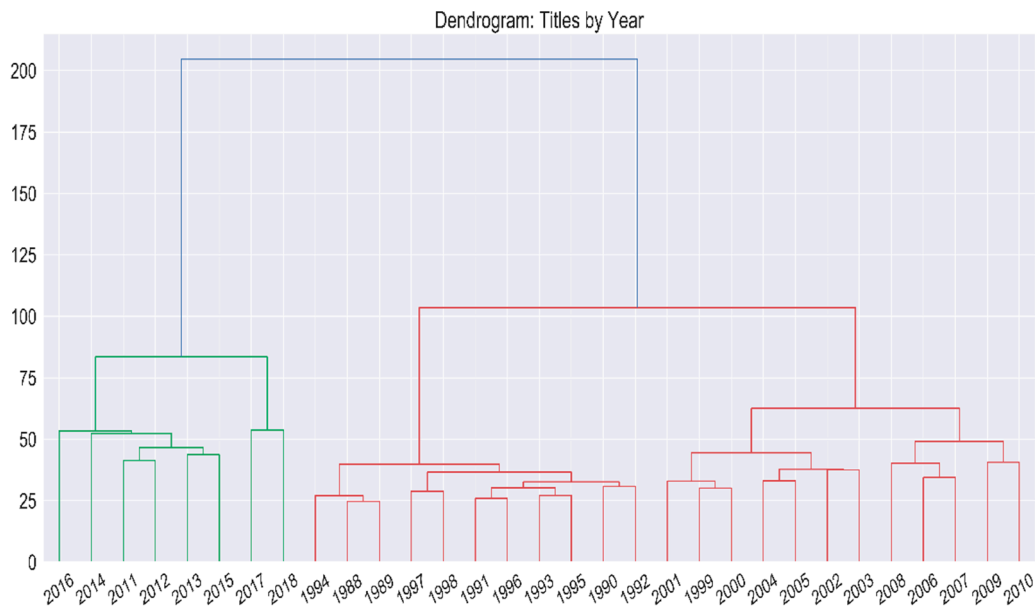**Figure 2.** Number of correlations of topic words in the annual report.

Figure 3. Clustering of the paper title by year.

The first subcluster of the second cluster in the middle of the dendrogram (1988–1998) covers the period of Kim Il-sung's leadership (1988–1994). In this subcluster, the similarity between annual titles is very high. First of all, compared with North Korea's economic growth rate, as shown in Figure 4, it can be seen that the first cluster is within the period that the North Korean regime was in crisis, during which the economy was designated as a default country in 1987 and experienced negative growth. Second of all, it can be inferred that the topic of interest in the economic policies during the Kim Il-sung era was maintained with no major changes. That is, early in his leadership (1995–1997), Kim Jong-il followed his father's policy direction without making any major changes.
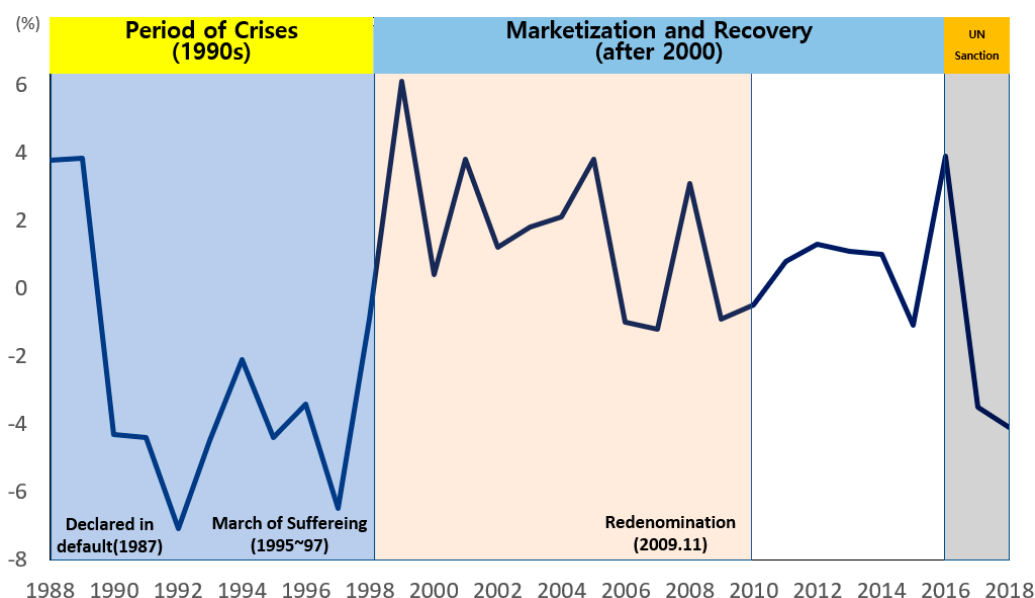
Figure 4. Economic growth rate of North Korea. source: Kim (2017), Bank of Korea economic statistical system.

The second subcluster in the second cluster of the dendrogram (1999–2010) corresponds to the period of Kim Jong-il's leadership (1995–2011), which includes the first stage, 1999–2005, and the second stage, 2006–2010. The second subcluster reflects a change in policy direction, which was triggered by the North Korean government, in a state of crisis, allowing some marketization in its economic system from 1999. It is also speculated that this period was followed by South Korea's policy of reconciliation and cooperation with the North, known as the Sunshine Policy, which contributed to the recovery of the North Korean economy, despite much criticism that the policy led to the North developing nuclear weapons and subsequently worsening the national security issue [16,17]. Growth started to rise sharply, but the North Korean authorities began to curb the expansion of marketization from the bottom in 2006. These socioeconomic changes in North Korean society are reflected in the second stage of the second subcluster of the dendrogram (2006–2010) in Figure 3.

The left-hand cluster of the dendrogram, the first cluster, is the Kim Jong-un period (2011–2018), which is divided into two subclusters, 2011–2016 and 2017–2018. The results of the clustering show that the Kim Jong-un period is not similar to the previous periods. (The vertical height of the stem in the dendrogram indicates the cosine similarity between titles by year. The higher the height, the lower the similarity. The Kim Il-sung and Kim Jong-il periods are also clustering, but in terms of similarity, they are closer to each other than they are to the Kim Jong-un period). This suggests that there have been notable changes in policy direction and issues of interest compared to the earlier eras. In fact, Kim Jong-un, who has been in power since 2011, began to take full advantage of the market when exports and imports greatly increased. As a result, the growth rate shifted upward, but the North Korean economy once again experienced negative growth in the second subcluster (2017–2018) after

the United Nation Security Council imposed tougher sanctions on North Korea. Figure 3 shows that the topics in *Economic Research* are clearly distinct from those in any other period.

3.3. Predicting a Paper's Publishing Year Using Logistic Regression

By utilizing a supervised learning approach, one of the machine learning classification techniques, this section verifies whether the timing of the appearance of keywords can be predicted by the characteristics found in the title of papers. Supervised learning techniques include logistic regression, support vector machines, and naïve Bayesian classifiers. This study predicts when each paper published in *Economic Research* was written based on the logistic regression, which shows the highest probability of classification among the various classification techniques (see Appendix A.1).

First, the documents that have undergone pre-processing and non-probability word representation are a set of multidimensional vectors. The total number of main words used in this study is 1846, so each title (2757 in total) is represented as a vector of 1×1846 dimension (nouns that appear more than once in every article title are called 'vocabulary' or 'dictionary'). The dimension of the DTM or tf-idf matrix is 2757×1846. Since the sample size is relatively small (this has a relatively small dimension compared to the sample size of so-called big data, which has been actively researched recently), the characteristics of the titles of the papers are distinguished by the leader's period of leadership when the accuracy of the classification through the machine learning algorithm is high.

Based on the results of hierarchical clustering, the categories were divided into four scenarios: 1. Kim Il-sung, Kim Jong-il, Kim Jong-un; 2. Kim Il-sung, Kim Jong-il; 3. Kim Il-sung, Kim Jong-un; and 4. Kim Jong-il and Kim Jong-un. Since the number of subject words is 1846, there are too many characteristics or variables for each title. We, therefore, reduce the dimension of the characteristics through a principal component analysis (PCA). We set the dimension to obtain an optimal number of characteristics up to 300. The logistic regression with L2 regularization selects the most relevant features of the 300. In addition, the k-overlapping cross-checking described earlier evaluated the performance of the parent model, selected the optimal model, and checked the classification performance using selected hyperparameters and rating indicators at a level that minimizes the overfitting and underfitting dilemmas. The hyperparameter used for cross-checking is:

$$c = \left\{ 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{1}, 10^{2}, 10^{3} \right\}. \tag{1}$$

Classifications of the eras of Kim Il-sung, Kim Jong-il, and Kim Jong-un. Figure 5 shows the results of the cross-examination from classifying the titles published during the leadership of Kim Il-sung, Kim Jong-il, and Kim Jong-un. To classify the titles by logistic regression, 1846 dimensions of the rare tf-idf matrix were reduced to 300 dimensions by PCA. (The number of features to be classified corresponds to a hyperparameter. In this case, the classification was classified into various numbers ranging from 3 to 300, but the accuracy of the classification was inferior to the characteristics of a number that is too low. Therefore, 300 characteristics or dimensions are implied to reflect the most abundant characteristics). The cross-validation test in Figure 5 confirms that there is little overfitting, despite it mitigating the constraints of the model. While the risk of overfitting is not significant, the appropriate modeling constraint is $c = 100$, at a level that minimizes underfitting. Here, the accuracy of the training data prediction is 0.77, and the subject of the paper published in *Economic Research* throughout the three leaders' periods can be seen as having relatively distinct characteristics. The accuracy of the prediction (if there is more than one category, evaluation indicators other than precision are omitted because their meaning is unknown) is 0.64, and each published title in *Economic Research* alone can predict, to some extent, when the paper was written. Therefore, over the three time periods, the main concerns and policy directions of the economy are reflected in the titles of the papers.
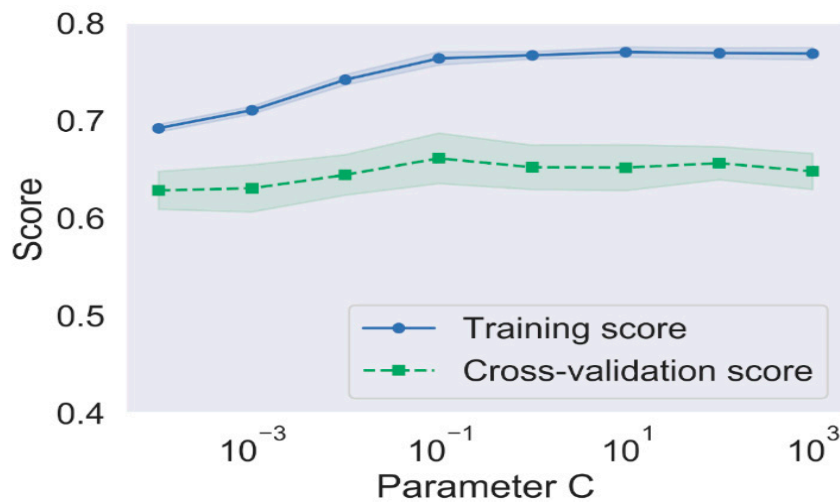
Figure 5. Accuracy test of classification of the title of the paper by period (Kim Il-sung, Kim Jong-il, and Kim Jong-un).

Figure 6 shows the accuracy ratio of the results predicted by the title of the paper and the confusion matrix. The diagonal elements in the confusion matrix are the probability that the predicted results are true, and the closer we get to one, the more distinct the characteristics are. When the Kim Il-sung, Kim Jong-il, and Kim Jong-un periods are grouped together, the results of the test data predict (this means the out of sample forecast discussed in econometrics) that a thesis title in the Kim Il-sung time period (class 0) is written in the Kim Il-sung period with 80% accuracy. The results also predict that a thesis title in the Kim Jong-il time period (class 1) is written in the Kim Jong-il period with 53% accuracy, and a thesis title in the Kim Jong-un time period (class 2) is written in the Kim Jong-un period with 61% accuracy. These accuracies were lower than those in the Kim Il-sung period, implying that the farther from the diagonal element of the confusion matrix, the lower the level of error. Therefore, it can be said that the characteristics of a paper title are clearly distinguished during the leadership periods of Kim Il-sung and Kim Jong-un. However, there is a 15–29% discrepancy between the titles of the papers between the Kim Il-sung and Kim Jong-il periods, and between the Kim Jong-il and Kim Jong-un periods.
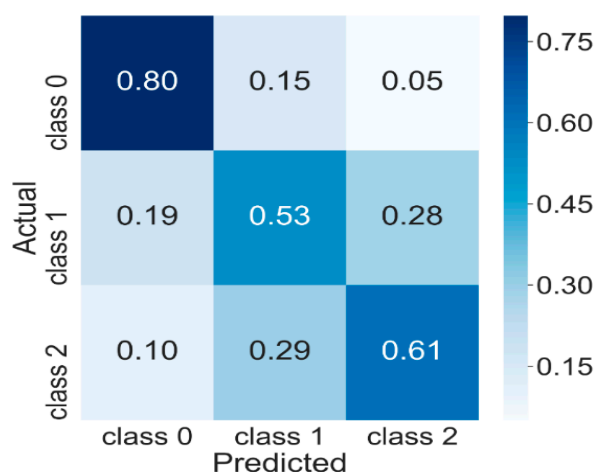

Figure 6. Result of classification of the titles of papers by period.

Kim Il-sung and Kim Jong-il time period comparisons. The accuracy of predictions will be greatly improved if the characteristics of the titles classified as only Kim Il-sung and Kim Jong-il are compared. The cross-validation test in Figure 7 shows no overfitting, while the modal constraint of underfit improves to a low level, at $c = 10^4$. (This shows that many of the characteristics of the 300 dimensions

contribute to the classification). The accuracy of the classification of training data at the constraint level of c $= 10^4$ is 0.92, and the subjects of the papers published in *Economic Research* during Kim Il-sung and Kim Jong-il's leadership periods have distinct characteristics. As shown in year-by-year clusters, Kim Jong-il initially inherited the policies of his father Kim Il-sung, but pursued his own, due to changes in the economic environment and potential national security threats at home and abroad.
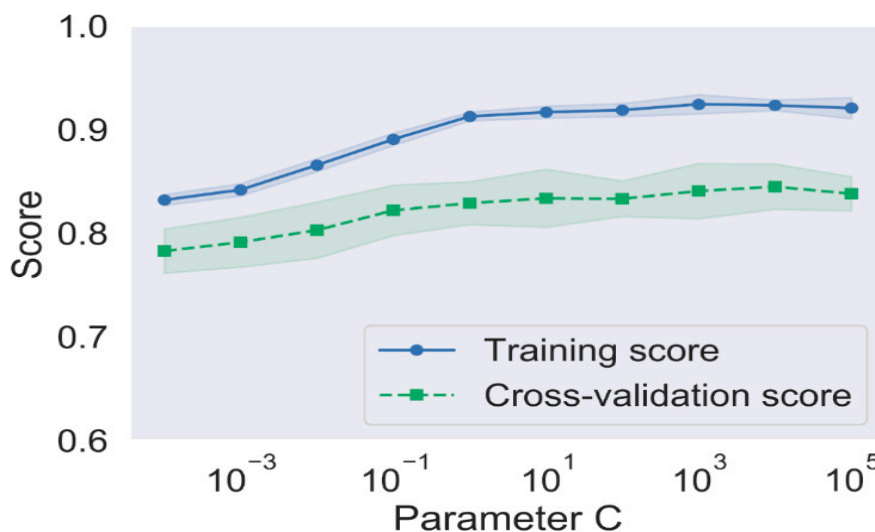

Figure 7. Accuracy test of classification of the title of the paper by period (Kim Il-sung, Kim Jong-il).

The confusion matrix between Kim Il-sung and Kim Jong-un is the same as in Figure 8. The accuracy of predictions for Kim Il-sung (class 0) and Kim Jong-il (class 1) periods in the paper titles is 82% for both. The recall, precision, and F1 scores are also seen to be very decent at predicting in both periods, and are 0.82, 0.81, and 0.82, respectively.
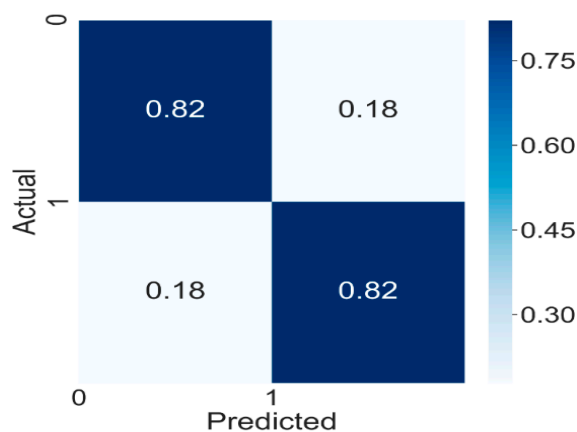

Figure 8. Result of classification 'Kim Il-sung, Kim Jong-il'.

Kim Il-sung and Kim Jong-un time comparisons. The accuracy of predictions is improved even further if the characteristics of the titles classified as only Kim Il-sung and Kim Jong-un are compared. The results show that there has been a major shift in economic policy and the economic environment between the two generations. The results of cross-checking are shown in Figure 9. Even the regulatory parameter at the c $= 10^5$ level is significantly improved, and the predicted accuracy of training data is close to one. Accordingly, the papers published in *Economic Research* during the Kim Il-sung and Kim Jong-un periods can be predicted with very high accuracy.
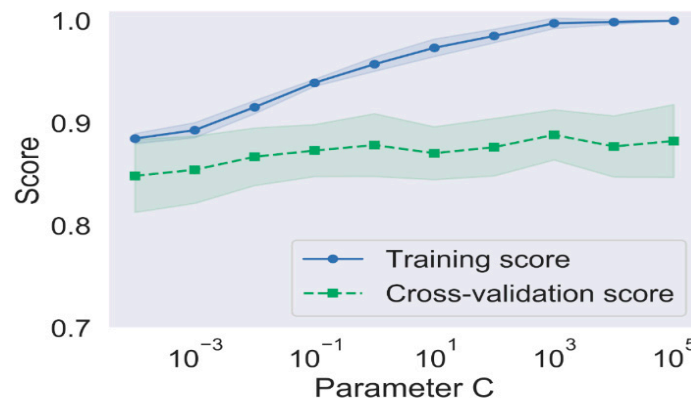
Figure 9. Accuracy test of classification of the title of the paper by period (Kim Il-sung, Kim Jong-un).

The confusion matrix between Kim Il-sung and Kim Jong-un are shown in Figure 10. The prediction accuracy of the Kim Il-sung period (class 0) and the Kim Jong-un period (class 1) for the test data are 97% and 85%, respectively, and the overall accuracy is 0.91. The recall, precision, and F1 scores are also seen to be very decent at predicting the model, and are 0.85, 0.96, and 0.90, respectively.
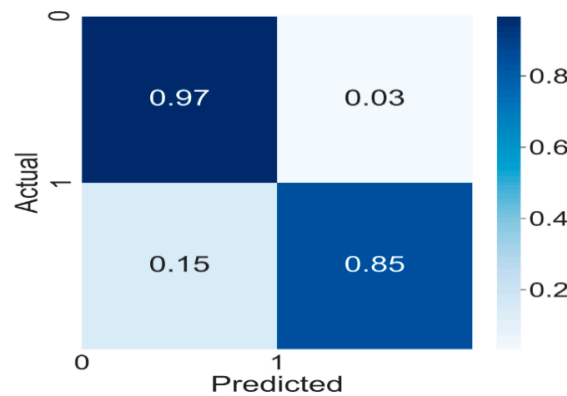

Figure 10. Result of the classification 'Kim Il-sung, Kim Jong-un'.

Kim Jong-il and Kim Jong-un time period comparisons. The prediction accuracies of the Kim Jong-il and Kim Jong-un comparisons is not much higher than that of the Kim Il-sung and Kim Jong-il comparisons, or the Kim Il-sung and Kim Jong-un comparisons. The results of the cross-checking are shown in Figure 11. The level of the regulation parameter at $c = 10^5$ is improving, but there is no indication of overfitting. The precision of the training data prediction at $c = 10^5$ is 0.82.
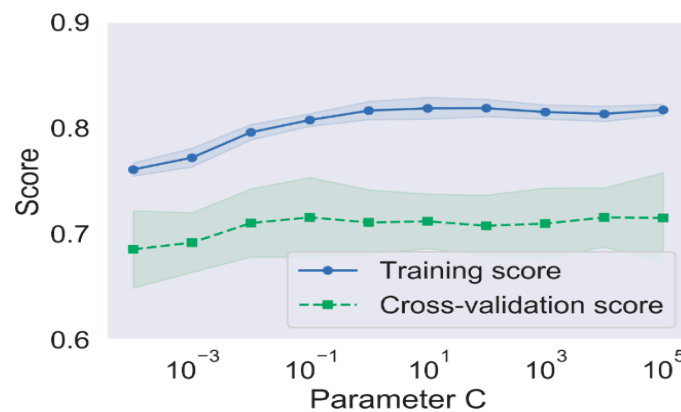

Figure 11. Accuracy test of classification of the title of the paper by period (Kim Jong-il, Kim Jong-un).

The confusion matrix between Kim Jong-il and Kim Jong-un is shown in Figure 12. The prediction accuracy of the paper titles being in the respective Kim Jong-il (class 0) and Kim Jong-un (class 1) periods are 68% and 71%, respectively. The recall, precision, and F1 scores are 0.71, 0.69, and 0.70, respectively, and the overall prediction accuracy for the model is 0.70. Therefore, the prediction is relatively decent.
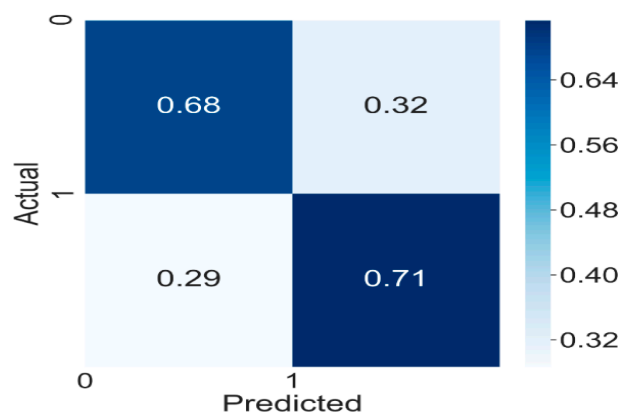

Figure 12. Result of classification 'Kim Jong-il, Kim Jong-un'.

3.4. Changes in Economic Environment and Policy Topics by Leader

Some keywords regarding socialist systems, criticisms of capitalism, and encouragement for the people frequently appeared throughout the Kim Il-sung, Kim Jong-il, and Kim Jong-un periods. As Figure 2 shows, there is a high correlation coefficient of paper topics. Nevertheless, words related to topics of particular interest in a period are expected to appear more frequently in the titles in that period than in other periods. Taking these aspects into account, we look at the big picture of North Korea's economic policy changes by analyzing the differences between each period based on the appearance and frequency of certain keywords.

Frequency of keywords by time period. Figure 13 shows the frequency of keywords during Kim Il-sung's leadership period (1988–1994). The most commonly appeared words or clichés, such as 'critical tasks', 'essential requirements', 'reactions', 'problems', 'preconditions', 'important principles', and 'struggle tasks', are excluded for being stopwords (When calculating the sample frequency, the common words may be treated as stopwords, but if the common words are also used as meaningful words in the title, they cannot be treated as stopwords. In the case of the tf-idf process described above, as the weight for any particular common word approaches zero, that common word is naturally removed). However, common words that were significant in a given context were not deleted.

If the frequency of keywords, such as 'technology', 'labor', 'rural areas', 'agriculture', 'industry and capitalism', is high, we assume that, at the time, the regime was shifting the direction of economic policy toward industrialization. Even though it criticized the capitalist system, it was concerned about labor productivity, technology, and capitalism, and was eager to strengthen its agricultural production.

Figure 14 shows the frequency of certain keywords during Kim Jong-il's leadership period (1995–2011). First, we can see that the frequency of 'technology' is as high as during his predecessor's leadership period, and some of the frequently-appearing words are 'capitalism', 'science', 'information', 'management', 'strong state', 'modernization', 'industry', 'international relations' and related words. During the Kim Jong-il period, there are many papers criticizing the capitalist system, while propagandizing the socialist system amid the rivalry with the South. In particular, at the end of each issue, one or two papers are classified as making criticisms of the capitalist system. At the same time, however, modern business management and economic management methodologies, such as international capital markets, market development, and increases in trade volume, are actively introduced amid economic hardships, when revitalizing the socialist economy is a goal.
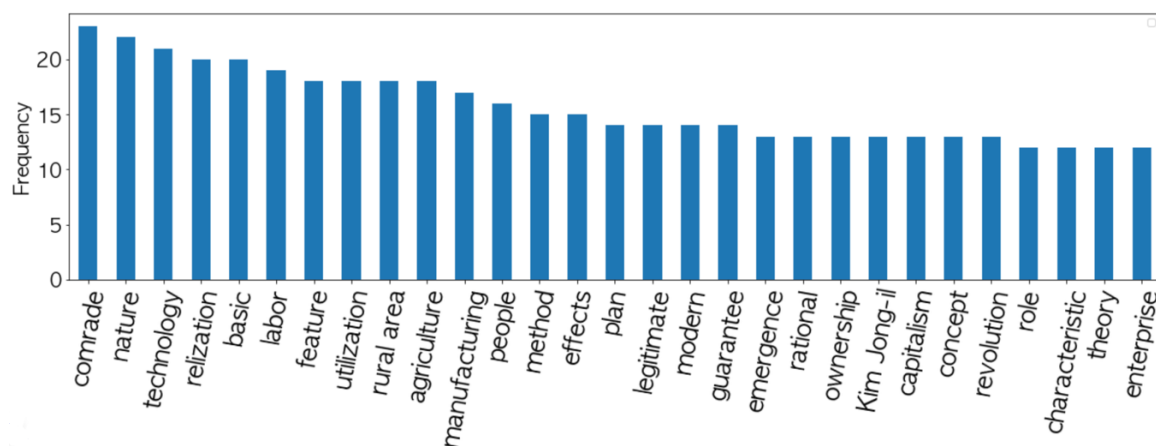
Figure 13. Frequency of keywords during Kim Il-sung period.

During the Kim Jong-un leadership period (2012–2018), there is a high frequency of words, such as 'production', 'knowledge', 'technology', 'power', 'information', and 'science'. Under the Kim Jong-un regime, the weight of titles related to criticisms of the capitalist system decreased significantly. However, the emergence of words like 'improving productivity', 'knowledge', and 'technology' indicates that economic policies are turning to economic growth based on pragmatism.
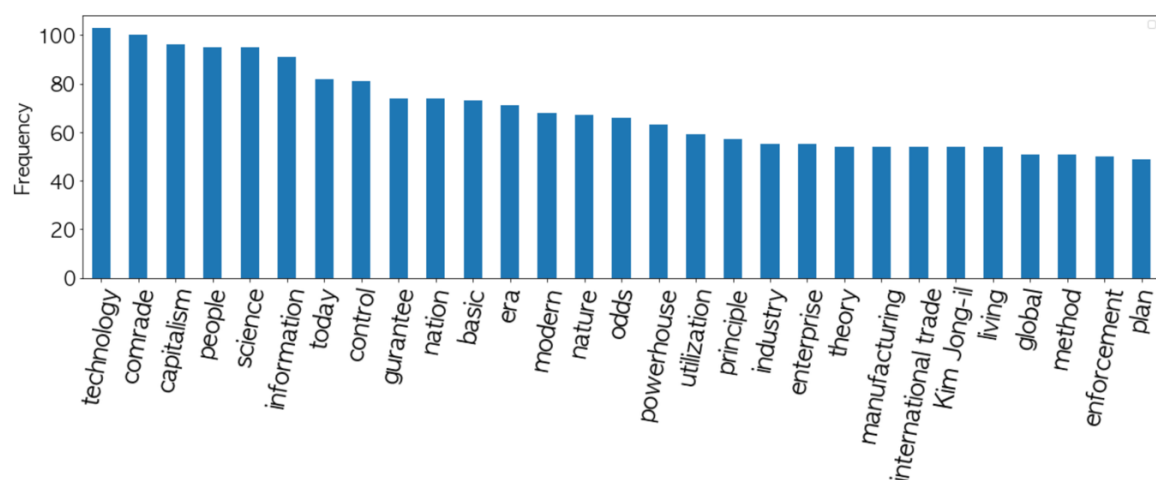

Figure 14. Frequency of keywords during Kim Jong-il period.

Topic models for economic policy by time period. In this section, the LDA method is used to conjecture keywords and the topic of interest for each leader's leadership period. An LDA [18] method is a probability model that classifies the topics of a document based on the probability that a particular document is a mixture of a small number of topics, and that each word is attributable to one of the document's topic [12]. The probabilistic model of the LDA method shows the most appropriate topic and probability according to the document or document set by estimating the maximum probability of there being a combination of the words and topics. In this paper, the hyperparameter representing the maximum number of topics is limited to 25. As a result of the estimation, similar topics are represented as one topic, and so the number of topics for each period can be determined as such.

Table 2 is the result of selecting a topic based on the keywords of each topic as a result of the LDA estimation. As can be seen from the table, the distribution of topics varies as the Kim Jong-un period progresses. Meanwhile, there are many topics on criticizing the capitalist system and promoting the construction of a communist society in the Kim Il-sung period.

Table 2. Results of topic extraction through latent Dirichlet allocation (LDA).

Period	Subject
Kim Il-sung	Criticism of the Defects of Capitalism and Propaganda of the Socialist System Role of Rural Areas and Life Revolution Modern imperialist products and Neo-colonialism Building a Communist Society with People, Labor and Agriculture as Fundamental Realization of a Socialist Enterprise System
Kim Jong-il	In praise of Kim Il-sung and Kim Jong-il's Achievement Bourgeois System Criticism International Economy and Capital Movement Achievement of the People's Revolution through Export Features of the Capitalist Market and its Criticism Criticism of Japan's Economic Invasion Review and Criticism of the Global Financial Crisis from the U.S. Agriculture and Finance, the Growth of Productivity across All Industries Achievement based on Kim Il-sung's Economic Theory Strategies for the Strategic Development of Socialist Companies against Imperialism Economic Development Based on Scientific Research and Information Criticism of Capitalist System, such as Monopoly Structure of Trade in Goods
Kim Jong-un	Achievement of Kim Jong-un and Predecessor Bourgeois System Criticism Light Industry Production System, Trading Principle Foundation of Commercial Activities and Income Revolution Subjective and Practical Use of Resources Improving Corporate Management in terms of Contracts, Transportation, and Accounting Foreign Banking System and Foreign Credit Strategy Recent Financial Markets and Money flows Currency Circulation and Exchange Rate Balance and Functional Integration of the National Economy Modern Insurance and Regulation Realization of Socialism based on the Working Class Capitalist Industrial System Product Development Investment Rational Management Method using Statistics Business and Policy, Society Economic Modeling, Trade Theory, Technical Knowledge, Mathematics, Chemistry Knowledge power science power Competitiveness in the era of internationalization Overseas Banking System

There was a growing interest in the international economy during the Kim Jong-il period, but, ironically, coincided with criticisms of 'Western imperialism'. Kim Jong-un's period deals with a wide range of topics that differ from his predecessors, depicted in the research with words, such as 'power of knowledge', and 'power of science', 'modern management', 'economics', and 'statistics', as well as 'the international financial market'. More weight is being placed on practical economic reform ahead of the 'ideology of successful construction of a communist or socialist economy' than before.

4. Sustainability of the North Korean Economy

4.1. Overview of Kim Jong-un's Economic Strategy

Considering the above analysis, the policy directions, subjects, and environment in the Kim Jong-un period are clearly different from those of previous periods. Thus, it is necessary to examine the papers in *Economic Research* in this period in more detail. First, in terms of composition and form, the number of publications has increased remarkably since Kim Jong-un became leader, and the topics and researchers covered by the paper have diversified. Second, it can be inferred from this change that policymakers and economists in the country are learning about the open capitalist economy.

As Figure 15 shows, the frequency of different topics between the Kim Il-sung and Kim Jong-il periods and the Kim Jong-un period, such as production, knowledge, and technology, is increasing. Despite inheriting the 'self-reliance' ideology, Kim Jong-un is pursuing a different path than his predecessors did to rebuild the North Korean economy, as announced in the 2016 New Year's address, which focused on science and technology advancements and the achievement of a 'nuclear state, as shown in Table 2. The policies adopted by Kim Jong-un promote a less closed or alienated

economic system than those his predecessors favored. In other words, major keywords, such as 'foreign banking systems', 'currency and exchange rates', 'trade theory', 'competitiveness in the era of internationalization', and 'overseas banking system' show that Kim Jong-un aims to expand and develop engines of economic growth similar to those used in the Western world, and demands that knowledge and technology realize his plans.
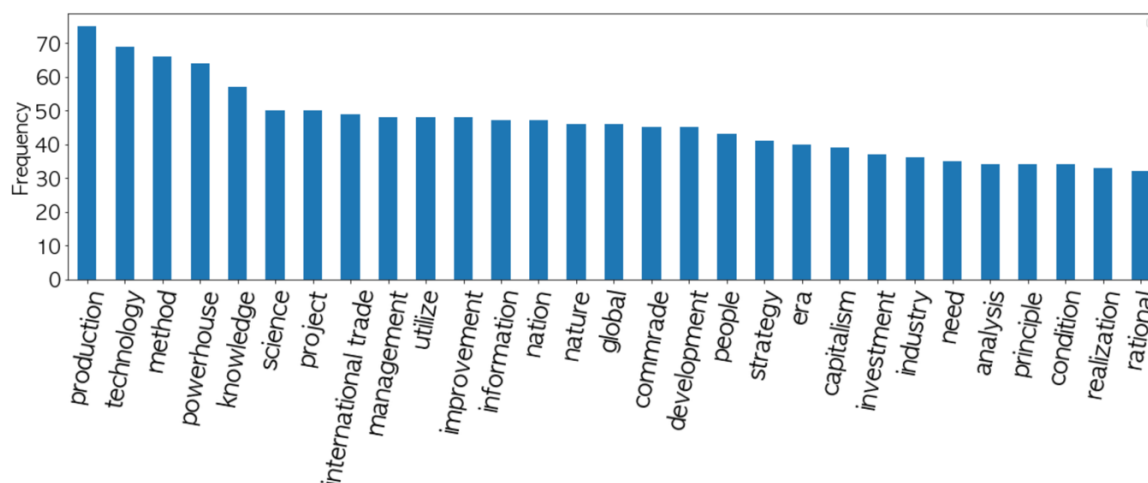

Figure 15. Frequency of keywords during Kim Jong-un period.

Additionally, from 2017, '*Silmari*', which corresponds to 'Keywords' in English, was added at the end of each paper. For example, if the title of the paper is, 'Our economic management method that embodies the construction of economic power and the self-reliance ideology', then 'economic power', 'self-reliance', and 'economic management' are selected as '*Silmari*'. This new addition means that managing keywords more systematically can make searching for related topics more efficient. The topics of the latest paper include 'knowledge economy', 'informationization', 'intellectual property rights', 'science and technology', 'tourism', 'economic and trade zones', 'special economic development zones', 'foreign exchange rates', 'accounting', 'finance', 'international banks', and 'Public-Private Partnerships (PPPs)', among others.

4.2. Economic Strategies and the Roles of the State

Specific changes in North Korea's economic strategy are represented, such as the Third Seven-Year Plan (1987–1994), the Revolutionary Economic Strategy (1994–1997), the Military First Economy (2003–2010), the Dual Strategy of a Nuclear-Economy (2013–2016), and the Five-Year Economic Development Plan, or New Strategy (2016–2020). Despite the various strategic changes made to overcome the economic crisis, the North Korean regime remains a centralized economy led by the state. The Kim Jong-un regime also emphasizes the role of the state as being an immutable principle for its economic policies.

There are two main reasons for this. First, North Korea's economy has not improved much since Kim Jong-un took power in 2012. The economic growth rate was 1.1% in 2013, and the economy retracted by 3.5% and 4.1% in 2017 and 2018, respectively, shortly after the ambitious New Strategy was launched. However, illegal and unofficial economic activities through the market have exposed weaknesses in the state management of economic resources. In addition, corruption, which has spread through illegal economic activities, is posing a threat to the centrally-planned economy by distorting the allocation of resources and threatening the socialist principle itself.

Second, the most imminent task for the regime is the restoration of authority and control of the state through 'economic management'. Since Kim Il-sung's death in 1994, North Korea has been suffering from an economic crisis, as well as a gradual weakening in the trust and authority of the supreme leaders, Kim Il-sung and Kim Jung-il. Under these circumstances, the prolonged economic

hardships have led to a vicious cycle of distrust in the state's functions and authority. Therefore, the first task of the Kim Jong-un regime was to restore the authority of the state (or of the Kim family), along with economic rehabilitation. Under this backdrop, the state's control of economic management, which the North Korean authorities emphasize, is quite clearly delineated.

An article entitled, 'The Status and Role of the State for Economic Management' published in the first issue of 2018, for instance, defines the status and the roles of the state as follows: "The nation's strategic management of economic projects means that the state functions in a commanding role to develop the economy, while actively taking various economic and organizational measures with a clear economic development vision". It also specifically describes three core roles of the state: First, 'achieving socioeconomic development based on quantitative growth of production'; second, 'defining the direction of the national economy with specific goals and development strategies'; and third, 'commanding and controlling the overall economic systems based on these development strategies'.

However, the following comments on the change in the role of the 'state', which distinguishes it from the previous Kim Il-sung and Kim Jong-il periods, are noteworthy: "The strategic management of the country's economic projects requires the state to map core sectors and core units of production and management activities, not all sector or units, but, to focus its [state] capabilities on its core areas to activate the entire economy" The strategy of the 'state' not controlling all sectors and all units, but focusing its core capabilities on key areas, can be interpreted as a deviation from the centralized state control system of an authentic socialist economy.

This subtle change accepted partial failures of the socialist economy, leading to the emergence of a new management style at state-run enterprises, the 'independent management system', in the early 2000s. Due to the breakdown of the national public distribution system, and the collapse of production facilities and capacity, North Korean authorities at that time ordered state-owned enterprises to be managed by individual companies to encourage 'the autonomy of the enterprise' and support senior managers and operations. This, in turn, meant the state could avoid the responsibility of running state-owned enterprises. There is a dilemma for the Kim Jong-un regime, which seeks to maintain the control function of the state, while responding to the damage that state authority caused through marketization from below. Internally, the Kim Jong-un regime could face several constraints, such as the loss of momentum for economic development, the weakening of state authority, and the deepening of economic dependence on China, due to economic sanctions.

Acknowledging this current policy environment and the economic situation in crisis, Kim Jong-un proposed holding the party congress and said, "(The) party will present a new five-year plan for national economic development, including the direction of the project in the next year, as well as the project of the party's central committee this year." The following sections elaborate specifically on strategies for the industrial sector and economic measures, to overcome these internal and external challenges.

4.3. Characteristics by Industrial Sector

By looking at the industrial strategies the Kim Jong-un regime initiated by sector, we can see the specific direction and goals of the economic strategy for which North Korea is aiming. Industry-related keywords mainly featured in the Kim Jong-un regime are 'the age of the knowledge economy' and 'construction of a strong science and technological state', along with those that represent maintaining the strategic status of the heavy and defense industries. By stressing that 'science and technology is the nation's most important strategic resource', it has set 'all of science and technology and production' as its policy goals.

To adapt to the age of the knowledge economy, in which knowledge is 'utilized as a basic source of production in the resource-intensive machine manufacturing industry', the regime stresses that the integration of science and technology and production and standardization should be made at production facilities to establish 'standardization in the production system, so the integration of human resources, research and development, production, sales, and even exports, can be used to build a production base for high-tech products. It also proposes specific measures, such as achieving 'modernization of the

people's economy' and 'informationization' through the development of science and technology, and in particular, the selection of talented workers for training professional, scientific personnel and educating at central institutions. Under the slogan of the 'establishment of a research-oriented university' and 'creative talent development', Kim Jong-un instructed that major universities, including Kim Il-sung University, be transformed into research-oriented universities to meet the needs of the 'knowledge economy era'.

The papers published in 2016 and 2017 need more attention. Before and after the announcement of the New Strategy in 2016, Kim Jong-un seemed to feel that North Korea had escaped the worst economic difficulties, and declared that fostering the tourism industry, attracting overseas investment, and developing special economic zones would be the main drivers of national development and restoring the country's authority. Specifically, the regime intended to boost the tourism industry by modernizing tourism facilities, industrializing hotel services, designating tourist development zones, and constructing ecological tourism bases. This strategy is seen to have been an attempt to resolve the foreign currency shortage, as the tourism industry is recognized as an exception from the economic sanctions.

The importance of attracting foreign capital and developing special economic zones during this period is being reflected in every issue of *Economic Research*. This is an external economic strategy that is very comparable to the Kim Il-sung and Kim Jong-il periods, and almost every volume now contains an introduction to international financial markets and foreign trade. It is worth noting that international financial market papers, for example, share very specific knowledge, such as hedging methods, accounting verification of foreign investment entities, credit at international banks, investment inducement plans, government and joint investment methods for foreign companies, tax credits, and tax favors for foreign companies.

Another notable feature is the emergence of papers related to credit transactions and exchange rate forecasts, which can be said to be a key to economic growth forecasts and the operation of international capital markets, as well as the introduction of foreign exchange account preparation methods and econometric models. For example, since the mid-2000s, there has been an increase in the number of papers that introduced basic aspects of finance and accounting, such as 'the general principles of statistical data collection' and 'International Accounting Standards', measuring derivatives trading and hedging, factors affecting the exchange rate, forecasting management statistics, methods for strategic preparation of national foreign exchange accounts, and methods for determining overdue collateral tax, thereby requiring more specific knowledge and techniques of the reader. Despite the level of quality, we notice that these introductions to capitalist market economic systems aim to educate policymakers and economists.

4.4. Prospect and Managerial Issues of Economic Reform in the North Korean Economy

As analyzed above, the North Korean economy of Kim Jong-un regime is aimed at expanding its openness to the outside world, realizing science and knowledge economy, and improving the standard of people's lives, while sticking to the *Juche* ideology and the planned economy centered on the state and the party. The question is whether the regime has the resources and national capacity to actually implement its economic strategies. Among the various studies related to national competence, recent studies have highlighted authority, acceptance, and ability as important factors. These three A's were initially used as analytical frameworks in various areas, including the study of the acceptance of information and technology [19], the communication between medical clinicians and patients [20], and the authority and acceptance of performance evaluation by employees in corporate management [21], and have recently been used as major variables in the execution of public policies [22,23].

An analysis of North Korea's economic policy implementation capabilities based on these three factors shows that it is not so optimistic. While the authority by the state and the party is alive in a situation where the public control and surveillance system by the state, and the party still operates

strongly, reform by 'authoritarianism without authority' is not sustainable. As seen in a study of Russia by Stoner-Weiss [24], 'authoritarianism without authority' can only worsen political corruption and social gaps. This weakening of authority will make the general public less receptive to policy, while increasing distrust in the state, and in the end, the authority will never be able to generate the policy effect intended by the state. Even under the Kim Jong-un regime, the nation's public economic system is not working properly, which further expands the informal economy and the illegal market economy. The growing informal economy, resulting from this weakening of the public sector, is a trigger to undermine the authority of the state and to damage the public's acceptance of state policies. Furthermore, as most of the nation's elite groups and resources are concentrated in the state political apparatus and defense sectors, the availability of human and material resources to be put into actual economic development is bound to be reduced.

5. Conclusions

This paper uses the periodical *Economic Research* (one of the representative pieces of primary literature published in North Korea) as the basic text to conduct an in-depth analysis of North Korea's economic policy and its changes over time. Based on the titles and keywords of the articles published from 1980 to 2018, a text mining and machine learning approach was employed to track changes in economic policy after a change of the national leader and in the economic environment.

We find that the title of the paper can be classified into the period of a leader with relatively high accuracy. In addition, the distribution of the thesis topic, along with the scope or direction, varies greatly from one period to another. Since the North Korean economy is inevitably influenced by a single leader, known as the Supreme Leader, the papers published in *Economic Research* can be regarded as reflecting the policy interests of the leader at the time. By extracting the keywords based on the characteristics of the titles, we gained an understanding of the trends in each leader's policy interests.

The implications of the analysis are as follows: Kim Jong-un is pursuing a different economic policy from those his predecessors pursued through reform and openness. As shown by the New Strategy, declared in 2016, the Kim Jong-un regime envisions a new economic strategy for economic development in the face of economic sanctions, including the knowledge economy, the industrial revolution in the new century, streamlining corporate management, and entering the international trade market. Specifically, the regime fosters the entry of international financial and commodity markets, and the exchange of advanced technologies and knowledge for competitive product development and commercialization. However, economic sanctions resulting from the nuclear weapons development act as fundamental constraints on the country. Nevertheless, improving U.S.–North Korea relations, and easing national security threats are expected to increase access to capital markets and overseas commodity markets; thus, the possibility of full-scale developmental cooperation cannot be excluded. Construction of infrastructure for the inflow of foreign capital, the standardization of financial markets, and general advancement will be needed in line with the reform and opening of North Korea. South Korea, which has geographical, cultural, and linguistic advantages, is in desperate need of cooperation with North Korea, and can take advantage of these positive aspects better than other neighboring countries can; in fact, South Korea may become a fund-raising hub, modeled after Hong Kong, to drive economic growth. It is time to develop realistic policy alternatives to lead a sound market economy with a deeper attention to the economic strategy and policy changes in North Korea.

In this study, we find that many articles appearing in *Economic Research* during the Kim Jong-un period refer to efficient resource input, corporate competitiveness enhancement, and expert utilization in management, which are emphasized in capitalist market economies. Nevertheless, North Korea still stresses the need to make itself more economically self-reliant to officially achieve its national goal of an independent socialist economy. In the absence of authority, acceptance, and ability, the gap between the state's political rhetoric and the actual economic reality is expected to remain wide.

Compared to the previous two Kim regimes, Kim Jong-un boldly admits to economic failures since his inauguration. It will seek to explore the possibility of economic development through the 'New Five-Year Plan for Economic Development' by 2021. Nevertheless, both 'reforms within the system' and 'reforms of the system', which is essential to North Korea's economic development, would not be an easy choice for Kim Jong-un. A wise alternative should be proposed so that the North Korean economy can go beyond 'muddling-through' and pursue sustainable development with international cooperation.

Appendix A

Appendix A.1. Logistic Regression

A logistic regression [25] is a probabilistic model in which linear combinations of independent variables account for or predict the likelihood of a particular event occurring. Unlike a general linear regression, a logistic regression corresponds to an independent variable whose value changes within the range of [0, 1] with a value of $[-\infty, \infty]$. Therefore, logistic regressions can be expressed by models of logit variable (l), independent variable vector (X), and regression coefficient vector (β) as shown below:

$$l(p) = log\frac{p(y = 1|x)}{1 - p(y = 1|x)} = \beta^T X \tag{A1}$$

The above expression can be expressed as: $p(y = 1|x) = 1/(1 + e^{-\beta^T X})$, which is called the logistic function. The logistic regression model estimates the probability that the dependent variable (y) belongs to zero or to one, and is the process of obtaining the decision boundary or hyperplane $\beta^T X$ that is the basis for the classification.

Appendix A.2. Unbalanced Data and Oversampling

Unbalanced data means that samples belonging to a particular category are significantly smaller or larger than those in other categories. If the classification of the data is severely unbalanced, the machine learning algorithm will classify all the smaller data points into categories to which the larger data points belong to improve the accuracy of the classification. In the case of the *Economic Research* titles used in this paper, the number of papers written during the leadership periods of Kim Il-sung, Kim Jong-il, and Kim Jong-un is 338, 1346, and 1073, respectively—the number of papers written during Kim Il-sung's rule is relatively small. In this case, if a large number of papers written during the Kim Il-sung period are classified as being from either the Kim Jong-il or Kim Jong-un period, the accuracy of the classification may be improved.

Oversampling can overcome these limitations when unbalanced data need to be classified. It is a technique that equalizes the size of samples in each category, and the most basic naïve random oversampling technique was used to make the sample size equal to 1346 for each period.

Appendix A.3. Cross-Validation and Evaluation

The performance of machine learning models can be optimized, by preventing either over- or under-fitting through cross-validation and the hyperparemeters settings during the learning process.

For logistic regression, the L2 regularization, which regulates the degree of overfitting, is used, and the model is selected by the corresponding hyperparameter. In other words, in minimizing the following cost functions (λ), the number and size of β_j estimated according to the size of hyperparameter c are estimated differently, and the cross-validation values of different models are compared to determine the degree of overfitting and underfitting.

$$\lambda = \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{n} x_{ij}\beta_j \right)^2 + \sum_{j=1}^{p} \frac{\beta_j^2}{c} \tag{A2}$$

The evaluation of the selected model mainly uses recall, accuracy, precision, and F1 score, which are most widely used as evaluation metrics. Table A1 shows the confusion matrix, which shows the combination of four coordinates according to the actual values in the classification and the prediction of the model. TP(TN) is the abbreviation for True Positive (Negative), representing when the model correctly predicts the true value of the feature. If the actual value is positive, or the model's prediction is the opposite, then FN(FP) is used. The following evaluation metrics are calculated by the values of TP, TN, FP, and FN in the confusion matrix.

Table A1. Confusion Matrix.

		Predicted Value	
		Positive	Negative
True Value	Positive	TP	FN
	Negative	FP	TN

First, accuracy shows the proportion of the actual value and the prediction of the model among the total values. It is used as a rating indicator even in multidimensional classifications, so it is most widely used among the model evaluation indices.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{A3}$$

Unlike accuracy, precision is the proportion of the actual value that the model predicts as positive.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{A4}$$

Recall is the ratio of what is predicted by the model as positive among those whose actual value is positive.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{A5}$$

The F1 score is calculated as the harmonic mean of the recall and precision indicators.

$$\text{F1score} = 2\frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{A6}$$

References

1. Haggard, S.; Noland, M. *Famine in North Korea: Markets, Aid, and Reform*; Columbia University Press: New York, NY, USA, 2007.
2. Frank, R. *North Korea's Economic Policy in 2018 and Beyond: Reforms Inevitable, Delays Possible. 38th North*; Stimson Center: Washington, DC, USA, 2018.
3. Kim, B.-Y. *Unveiling the North Korean Economy: Collapse and Transition*; Cambridge University Press: Cambridge, UK, 2017.

4. Yang, M.-S. An In-depth Reading of North Korean Documents: A Case Study of Gyeongjeyeongu (Economic Research). *Rev. North Korean Stud.* **2009**, *12*, 7–61.

5. Kim, B.-Y.; Kim, S.J.; Lee, K. Assessing the economic performance of North Korea, 1954–1989: Estimates and growth accounting analysis. *J. Comp. Econ.* **2007**, *35*, 564–582. [CrossRef]

6. Choi, C.; Jesse, D.L. A Semantic Network Analysis of Changes in North Korea's Economic Policy. *Governance* **2012**, *25*, 589–616. [CrossRef]

7. Park, J.H.; Eunjeong, P.; Jo, D.-J. Automated Text Analysis of North Korean New Year Addresses, 1946–2015. *Korean Political Sci. Rev.* **2015**, *49*, 27–61.

8. Lee, C.H. The analysis of North Korea's economy reform: Focused on the Journal of Political Economy (1986–1999). *North Korean Study* **2010**, *6*, 147–186. (In Korean)

9. Lim, E.C. Study on the Informationization Strategies as a Part of North Korea's Economic Modernization: Evidence from Economic Research. *Korean J. Mil. Art Sci.* **2019**, *75*, 253–278. (In Korean)

10. Na, J.W. Current Research Trends in Management Studies in North Korea during the Kim Jung Un's Era: Case study of Economic Research. *North Korean Stud. Rev.* **2018**, *22*, 57–85. (In Korean)

11. Choi, J.; Soo, H.L.; Seogki, L.; Yoojeong, C.; So, J.I. *Analysis on North Korea's Trade and Industrial Policy. KIEP Policy Analysis*; Korea Institute for International Economic Policy: Sejong, Korea, 2017.

12. Kim, S.; Youngjoon, L.; Jhinyoung, S.; Ki, Y.P. Text Mining for Macroeconomic Analysis. *J. Korean Econ. Anal.* **2020**, *25*, 2.

13. Rich, T.S. Deciphering North Korea's Nuclear Rhetoric: An Automated Content Analysis of KCNA News. *Asian Aff.* **2012**, *39*, 73–89. [CrossRef]

14. Rich, T.S. Propaganda with purpose: Uncovering patterns in North Korean Nuclear Coverage, 1997–2012. *Int. Relat. Asia-Pac.* **2014**, *14*, 427–453. [CrossRef]

15. Whang, T.; Lammbrau, M.; Joo, H.-M. Detecting patterns in North Korean military provocations: What machine-learning tells us. *Int. Relat. Asia-Pac.* **2016**, *18*, 193–220. [CrossRef]

16. Jung, K. Evaluation of Kim Dae Jung Government's North Korea Policy and the Future Tasks. *Unification Strategy* **2002**, *2*, 79–103.

17. Shin, J.; Kim, D.J. Government's Sunshine Policy and the Structural Limits. *Korean J. Int. Stud.* **2003**, *43*, 295–316.

18. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.

19. Dillon, A. User Acceptance of Information Technology. In *Encyclopedia of Human Factors and Ergonomics*; Karwowski, W., Ed.; Taylor and Francis: London, UK, 2001.

20. Stivers, T.; Timmermans, S. Medical Authority under Siege: How Clinicians Transform Patient Resistance into Acceptance. *J. Health Soc. Behav.* **2020**, *61*, 60–78. [CrossRef] [PubMed]

21. Nielsen, P.A.; Jacobsen, C.B. Zone of Acceptance under Performance Measurement: Does Performance Information Affect Employee Acceptance of Management Authority? *Public Adm. Rev.* **2018**, *78*, 684–693. [CrossRef]

22. Andrews, M. Authority, Acceptance, Ability and Performance-Based Budgeting Reforms. *Int. J. Public Sect. Manag.* **2004**, *17*, 332–344. [CrossRef]

23. Andrews, M.; Pritchett, L.; Woolcock, M. *Doing Problem Driven Work*; Harvard University: Cambridge, MA, USA, 2015.

24. Stoner-Weiss, K. *Resisting the State: Reform and Retrenchment in Post-Soviet Russia*; Cambridge University Press: Cambridge, UK, 2006.

25. Cox, D.R. The Regression Analysis of Binary Sequences. *J. R. Stat. Soc. Ser. B (Methodol.)* **1958**, *20*, 215–242. [CrossRef]