Algorithm Design to Judge Fake News based on Bigdata and Artificial Intelligence

Jangmook Kang, Sangwon Lee*

Department of Bigdata & Industry Security, Namseoul University
kangjm@nsu.ac.kr
*Department of Computer & Software Engineering, Wonkwang University
sangwonlee@wku.ac.kr

Abstract

The clear and specific objective of this study is to design a false news discriminator algorithm for news articles transmitted on a text-based basis and an architecture that builds it into a system (H/W configuration with Hadoop-based in-memory technology, Deep Learning S/W design for bigdata and SNS linkage). Based on learning data on actual news, the government will submit advanced "fake news" test data as a result and complete theoretical research based on it. The need for research proposed by this study is social cost paid by rumors (including malicious comments) and rumors (written false news) due to the flood of fake news, false reports, rumors and stabbings, among other social challenges. In addition, fake news can distort normal communication channels, undermine human mutual trust, and reduce social capital at the same time. The final purpose of the study is to upgrade the study to a topic that is difficult to distinguish between false and exaggerated, fake and hypocrisy, sincere and false, fraud and error, truth and false.

Keywords: Fake News, Bigdata, Artificial Intelligence, Algorithm.

1. Introduction

The contents of the research include the analysis of vocabulary, sentence analysis, and the extraction of a richer dictionary (verbal dictionary) to improve the quality of rule-based quality by using WordNet, such as 'when the title of fishing is attached to a news article and the content provides information that is not relevant to the main context of the news content'. The scope and results of the research are the process of determining the true and false facts of current news articles and securing the intellectual property rights of data flow diagrams, system schematics and algorithms based on countless discussions and countless discussions through the design of user-led or participatory news articles' Fact check system using SNS based on Bigdata [1-11]. The scope and results of the research are the process of determining the true and false facts of current news articles and securing the intellectual property rights through papers and patents based on numerous discussions

and numerous discussions through designing user-led or participatory news articles' Fact Check System.

2. Related Works

We look at the artificial intelligence [12-15] techniques and data (text) mining [16-23] techniques that are mainly used in this study.

2.1 Artificial Intelligence

Artificial intelligence is a branch of computer science and information technology that studies how computers can do thinking, learning and self-development with human intelligence, and says it is artificial intelligence that allows computers to imitate human intelligent behavior. Also, artificial intelligence does not exist in its own right, but has much to do with other areas of computer science, directly or indirectly. In particular, attempts to introduce artificial intelligence elements in various fields of information technology to utilize them in solving problems in the field are very active. (1) In the field of natural language processing, systems such as automatic translation are already put into practice, especially when further research is done, people will be able to communicate and exchange information with computers, which will lead to innovative changes in computer use. (2) In the field of expert systems, computers are allowed to replace many of the professional tasks that humans are currently doing (diagnosing doctors, evaluating the deposits of minerals, estimating the structure of compounds, and judging the damages insurance premiums, etc). It was the earliest development in many fields. (3) The task of analyzing images captured by a computer through a TV camera to find out what it is, or listening to a person's voice and converting it into sentences is very complicated, and impossible without the introduction of artificial intelligence theory. These imaging and speech recognition are key technologies for text recognition, robotics, and so on. (4) Theorem probing is an essential skill used in various fields of artificial intelligence, itself of great value, as a process of proving mathematical theorem logically from known facts. (5) Neural Network, which has appeared relatively recently, is not mathematical logic, but imitating the human brain to assume a neural network structure consisting of a network of numerous simple processors.

Artificial neural network (ANN) is a statistical learning algorithm inspired by the neural network of biology (especially the brain in the central nervous system of animals) in animals. Artificial neural network refers to the entire model in which an artificial neuron (node) formed a network through the coupling of synapses changes the binding strength of the synapses through learning, which has the ability to solve problems. In a narrow sense, it is sometimes referred to a multilayered perceptron using the error reverse propagation method, but this is a wrong use, and artificial neural networks are not limited to this. Artificial neural networks include teacher learning that is optimized for problems by input of teacher signals (correct answers) and comparator learning that does not require teacher signals. Teacher learning is used when there is a clear answer and comparator learning is used for data clustering. Artificial neural networks are generally used to guess and approximate veiled functions, depending on many inputs. Generally expressed by the interconnection of the neuron system, which calculates values from input, and with adaptability, machine learning such as pattern recognition can be performed. For example, a neural network for handwriting recognition is defined as a set of input neurons, which are activated by pixels in the input image. The deformation and weighting of a function (they are determined by the person who created the neural network) is applied and the activation of that neuron is transferred to another neuron. This processing is repeated until the last output neuron is activated, which depends on which character has been read. Neural networks, like other machine learning - learning from data

- are generally used to solve a wide range of problems, such as computer vision or voice recognition, which are difficult to solve with rule-based programming.

The survey of the human central nervous system inspired the notion of neural networks. Artificial neurons are connected to each other in the artificial neural network to shape a network that mimics a biological neural network. There is no official definition of what an artificial neural network is. However, if a set of statistical models has the following characteristics, that set is called Neural. It consists of a set of adjustable weights, i.e. parameters expressed in numbers that can be adjusted by the learning algorithm. Nonlinear function of input can be inferred. Adjustable weights mean the coupling strength between neurons, which operate during training or forecasting. Neural networks resemble biological neural networks in that various units perform functions in parallel or collectively by units rather than by assigned subtasks. The word 'neural network' usually refers to models used in statistics, cognitive psychology and artificial intelligence. Neural network models imitating central nerves are part of theoretical neuroscience and computational neuroscience. In modern software that implements artificial neural networks, biological approaches are not used primarily because of more realistic approaches based on signal processing and statistics. In some of these systems, parts of the neural network or neural network (artificial nerves) form large systems, which are combined into adjustable or dysfunctional components. While the general approach of these systems is suitable for solving many real-world problems, this is not the case with traditional AI-connected models. But they also have something in common: nonlinear principle dispersion, parallelism and regional processing and adaptation. Historically, the use of neural models is a paradigm shift from high-density artificial intelligence, which features an intelligent expert system expressed in if-then rules, to sub-symbolic mechanical learning, which features intelligence with parameters of dynamic systems.

There are many types of artificial neural network, such as complex multiple inputs and directional feedback loops, one-way or two-way, and different tiers. Overall, the algorithms of these systems will determine the control and connection of the functions, respectively. Most systems are used to modify the system's parameters by connecting "weights" with various nerves. Artificial neural networks can automatically learn from outside training or develop using data on their own. Feedforward Neural network: Artificial Neural network method of the simplest method. Neural network information is passed from the input node through the hidden node to the output node, forming a graph in which no circulation path exists. There are many different methods of structure, such as binary structure, perceptron, sigmoid, and so on. (1) Radial system function network: Radial artificial neural network has a very powerful capability for interpolation of multidimensional space. The radiation function may replace the shape in which the sigmoid function of the multistory layer is used in the hidden node. (2) Cohen self-organizing network: Self-organizing network algorithm is one of the representative neural network algorithms, using self-supervised and competitive learning methods, as opposed to most neural network algorithms using a map-based learning method. The neural network is divided into the input layer and the competition layer, and each neuron in the competing layer calculates how close the connection strength bag and input bag are. And each neuron wins the closest to competing against each other for the privilege of learning. This winner neuron is the only neuron that can send out an output signal. In addition, only this neuron and its adjacent neurons are allowed to learn about the proposed input backers. (3) Circulating Artificial Neural Network: Circular Artificial Neural Networks do exactly the opposite of the front neural network. Two-way data movement between nodes and data is linearly communicated. The operation may be performed by transferring data from the rear node to the front node.

2.2 Data Mining

Data mining refers to the process of discovering useful correlations hidden among many data, extracting actionable information in the future and using it for decision making. The process of discovering new data models from databases that were not previously known but were guided in the data, extracting actionable information in the future, and using them for decision making. In other words, they find patterns and relationships hidden in the data, and they discover information just as they do in the light vein. Where information discovery is the process of finding useful patterns and relationships by applying advanced statistical analysis and modeling techniques to the data. It is a key technology in database marketing. For example, a department store analyzes data from its sales database and finds out which products sell well Friday morning and what correlations are made between the products sold, and reflects them in marketing. Therefore, the essential element of data-making is sufficient reliable data. This is because sufficient reliable data enable accurate predictions. However, too much data may rather reduce the predictability of data-mapping, and it is necessary to secure meaningful data to produce the best results. The commercialization of data-mapping is under way as many data-ware houses are already built in Korea, which is the optimal system for data-mapping, and corporate requirements are moving toward database marketing, which focuses primarily on customer management. On the other hand, some say that the data warehouse will be reduced if software that can implement data-mapping well in terms of technology comes out. The process of finding knowledge of interest by any method (sequential pattern, similarity, etc.) within the database. Data mining is the process of finding useful information in large amounts of data, and it means the technology to find unexpected information as well as the information you expect. Data mining can help maximize profits by creating valuable information and applying it to decision making. It aims to discover hidden knowledge, unexpected trends, or new rules based on all available source data, including the daily transaction data held by the entity, customer data, product data, or customer response data from various marketing activities, etc., and to use them as information for making actual business decisions, etc.

Text mining is a data-mapping technique that utilizes natural language processing techniques based on linguistics, statistics, machine learning, etc. to stereotype semi-regular/non-regular text data, to extract features, and to discover meaningful information from extracted features. The following shows the process of natural language processing by ETRI's natural language-based query and response system. The main technologies of natural language processing-based text mining include natural language processing (fascinating, shapeshifting, adverbial tagging, relational extraction, semantic extraction), language modeling (language detection, rule-based object name and commercial recognition), machine learning algorithms (improving the ability to use information acquired through repeated training), and mining techniques (information classification and analysis techniques using each statistical technique). In order to implement natural language processing-based text mining technology, natural language text must be structured into elements that the computer can understand, extracting meaning from sentence-level text. The natural language processing-based text mining technology measures performance by response rate, accuracy of interpretation, reliability and consistency of extract results, processing speed, and scalability.

3. Research Design of Judge Fake News based on Artificial Intelligence

Based on the technical social construct theory, the whole research direction and the concept of research are developing algorithms that enable news article-specific pre-construction and understanding of news context, and through these results, self-learning on news identification through artificial intelligence-based software, and thus continuously improving data quality, screening ability, etc. This study can be divided into three major

parts as follows.

First, the corpus, etc., which is mainly used in the news, is written in the dictionary. At this stage, news-specific pre-production (theoretical exploration and hands-on experience through the biking pre-processing) is carried out. Request to the major academic societies of politics/law/media/social studies/philosophy/art/ fusion, present the results of the already developed S/W and test, identify the emerging issues, and analyze the definition of variables of humanities and sociologists and whether they can be coded.

Second, we discuss in-depth discussions on the issue: political, economic and social risks based on news context and value. Cleaning up learning data (by news section), some of the actual news data is controlled by difficulty level and learned after generating fake news. (Application of machine learning classification techniques) There are many cases of real-life damage among "real-life, fake-looking, fake-fake-fake, fake-real-looking" cases of real-life mistakes. (At the elementary level, the sensitivity of the fake news damage is identified and the weight is adjusted.)

Third, measure performance by weighting various variables, such as SNS and artificial intelligence, to determine the authenticity of the news in the previously developed news screening system. Simulated test analysis through actual news data, removal of appropriate stopwords for new words, and stepping to distinguish between multiple words. There is an elementary discussion of developed models and algorithms. (The humanities/socialists will be reminded of the algorithm and Feedback will be carried out.)

4. Algorithm Design of Judge Fake News based on Artificial Intelligence

The method of problem solving based on the 19th century scientific philosophy that the applicant has learned so far (the method of making assumptions, defining variables in a manipulative way, experimenting with them, and applying them to the world by laying theories) could not solve the value orientation and politically and economically complicated social phenomena and the results that have been achieved under numerous limitations and constraints, but only fitted to a hyper-connected society.

The contents of the research include the analysis of vocabulary, sentence analysis, and the extraction of a richer dictionary (verbal dictionary) to improve the quality of rule-based quality by using WordNet, such as 'when the title of fishing is attached to a news article and the content provides information that is not relevant to the main context of the news content'.

The scope and results of the research are the process of determining the true and false facts of current news articles and securing the intellectual property rights of data flow diagrams, system schematics and algorithms based on countless discussions and countless discussions through the design of user-led or participatory news articles' Fact check system using SNS.

To achieve the above-mentioned research objectives, the following specific strategies and methods are taken.

First, we analyze related algorithms at home and abroad through thorough literature research. This is done by reviewing and writing documents, including system schematics, data flow diagrams, and operational definitions of variables.

Second, it analyzes the opening of public data at home and abroad (e.g. information disclosure portal, open data in Seoul, data.gov in the U.S.), libraries to refer to thesaurus, and APIs of social networking services that can refer to true/false judging news. In particular, it analyzes news libraries (from ancient times to modern news, research on tone and speech unique to corpus and news) that have been shared by

Keynes(www.kinds.or.kr) and others. Through the old newspapers of Kinds, research should be carried out in a challenging way so that news language changes and changes in tone and tone of news can be discussed with humanities scholars and others and applied to algorithms of fake news screening systems to be developed in the future.

Third, the application technology (formal analyzer, natural language processing, sentence pattern processing, data classification machine, data compressor, query analysis by natural language processing, summary technology of news content, etc.) is reviewed and AI model and model evaluation indicators are analyzed.

Total population	Predicted condition			
Total population	Predicted Condition positive	Predicted Condition negative	**Prevalence** $= \dfrac{\Sigma \text{ Condition positive}}{\Sigma \text{ Total population}}$	
True condition — condition positive	**True positive**	**False Negative** (Type II error)	True positive rate (TPR), Sensitivity, Recall $= \dfrac{\Sigma \text{ True positive}}{\Sigma \text{ Condition positive}}$	False negative rate (FNR), Miss rate $= \dfrac{\Sigma \text{ False negative}}{\Sigma \text{ Condition positive}}$
True condition — condition negative	**False Positive** (Type I error)	**True negative**	False positive rate (FPR), Fall-out $= \dfrac{\Sigma \text{ False positive}}{\Sigma \text{ Condition negative}}$	True negative rate (TNR), Specificity (SPC) $= \dfrac{\Sigma \text{ True negative}}{\Sigma \text{ Condition negative}}$
Accuracy (ACC) = $\dfrac{\Sigma \text{ True positive} + \Sigma \text{ True negative}}{\Sigma \text{ Total population}}$	Positive predictive value (PPV), Precision $= \dfrac{\Sigma \text{ True positive}}{\Sigma \text{ Test outcome positive}}$	False omission rate (FOR) $= \dfrac{\Sigma \text{ False negative}}{\Sigma \text{ Test outcome negative}}$	Positive likelihood ratio $(LR+) = \dfrac{TPR}{FPR}$	Diagnostic odds ratio $(DOR) = \dfrac{LR+}{LR-}$
	False discovery rate (FDR) $= \dfrac{\Sigma \text{ False positive}}{\Sigma \text{ Test outcome positive}}$	Negative predictive value (NPV) $= \dfrac{\Sigma \text{ True negative}}{\Sigma \text{ Test outcome negative}}$	Negative likelihood ratio $(LR-) = \dfrac{FNR}{TNR}$	

Figure 1. Model evaluation measure

Fourth, this research method studies Accuracy, Precision, Recall, F1-measure (Precision and recall) through Naive Bayes TF, Naive Bayes TF-IDF, etc. as a learning model that distinguishes between true and false news. But can the above research methods really determine the true and false news? Of course, as shown in the study below, false or untrue or misleading news or angling news can be distinguished by discrepancies in the title and content of the news, and inserting sentences that are not related to the entire context of the news. However, it may be difficult to recognize the complex semantic perception between truth and falsehood of the value orientation that news has, the political/social/economic context that it is inherent in.

Fifth, subdivisions of words and synonyms are required according to the sections that news has (comprehensive, economic, political, social, cultural, entertainment, etc.). Writing dictionaries for synonyms and significance or internal tags goes beyond the previously open ETRI's pre-based learned data sets to carry out further studies of the significance and synonym dictionary and tag configurations by news section that are uniquely optimized for news.

Sixth, is a model with a high accuracy just a good price to be high? Is increased precision a good model? Will solving technical problems solve social challenges? If you only increase the evaluation index of precision, reproducibility and precision, do you differentiate between true and false? What kind of mission does artificial intelligence unlock the lies and truths that are everywhere? All sorts of techniques do not improve your performance. In order to use context-sensitive features, studies should be conducted to identify the

chapters/units/opportunities/threats of the model. In order to do this, the slogan of solving social problems as if the engineer would solve the math formula alone should go into the world, and the proposed model should be made through a process of bowing, in which the proposed model is broken and adjusted to a new media major to a political scientist.

Seventh, semantic recognition and context understanding are essential to distinguish between true and false news. Today, it is simply functional learning of 1.3 million news articles, creating rules within them, and comparing them with a dictionary of words to increase precision. However, we need to define quality data that is optimized for news, thereby adjusting the weights to select good features, and finally we plan to design better algorithms based on them. In this process, it is important to explore which features should be used before machine learning (EDA) under the assumption that "news articles have their own way of speaking" unlike the existing data. To this end, there is a need to reconstruct unit technologies and plans should be made to carry them out in future seven-step studies.

Eighth, based on existing prior learning, Feature {Content, Title}, {Life, Region], {Products}Vector Length is calculated using the ratio of composition and word count, and the process of determining various fake news such as Logistic Regress, Random Forest, and Naive Bayes TF parameter tuning are newly organized and proposed through coordination with humanities/social scientists.

Ninth, more than one study should be carried out to complete the true/false determination algorithms of news articles that are newly proposed in the following table as hardware/software architectures, and the task of designing and verifying the optimized algorithms to actual news data should be focused on the last three years. It should plan to design and propose artificial intelligence learning optimized for distributed and news environments based on Hadoop in the future. In order to help identify fake news, but to achieve more than a certain level of convergence, precision, and reliability, the Korean language analyst's wisdom in the dictionary, in-house search and methodologies is requested through this task.

Tenth, as we enter a hyperconnected society, it is limited to distinguish between true/false news by the power of engineers alone. Of course, good research can be achieved in a limited environment of functional improvement or statistical significance. Through this research, we want to develop fake news identification algorithms/architects by adding rational coordination processes and engineering logic to news supply sites where the values of social scientists collide, such as understanding the context of fake news screening processes, reflecting media characteristics of news, and meaningful use of existing data such as data quality data collection(Kinds.or.kr in addition to the meaningful use of a dictionary of significance and synonyms as a public database. In addition, various studies should be reviewed to make efficient use of the Folksonomy tag, to review the dynamic navigation link and the architectural model of the tag cloud, and to incorporate the tag cloud and dynamic data catalysts logs using Ontology into the design of future algorithms.

5. Conclusions

The results of this study will mean that engineers will be central and try to solve social challenges, such as MIT's MediaLab, by accepting the imaginations of the humanities and applying the rationality of the social scientist. Instead of simply solving problems presented by a technologist or a social scientist, the research team uses its logic to ask, "Why is the social issue?" and design algorithms based on the rationality of the humanities and social scientists. It will contribute to the development of the next generation of studies by fostering the ability of engineering students or researchers who wrote code only after graduating from college or university to look at social phenomena and experience the protocol of humanities/social scientists being implemented in

code rather than simply staying in social protocol. Based on the technical social construct theory, the whole research direction and the concept of research are developing algorithms that enable news article-specific pre-construction and understanding of news context, and through these results, self-learning on news identification through artificial intelligence-based software, continuously improving data quality, screening ability, and so on the basis.

Acknowledgement

References

[1] S. Park, J.S. Hwang, and S. Lee, "A Study on the Link Server Development Using B-Tree Structure in the Big Data Environment", Journal of Internet Computing and Services, Vol. 16. No. 1. pp. 75-82, 2015.
DOI: https://doi.org/10.7472/jksii.2015.16.1.75.

[2] S.B. Park, S. Lee, S.W. Chae, and H. Zo, "An Empirical Study of the Factors Influencing the Task Performances of SaaS Users", Asia Pacific Journal of Information Systems, Vol. 25. No. 2. pp. 265-288, 2015.
DOI: https://doi.org/10.14329/apjis.2015.25.2.265.

[3] S. Park, and S. Lee, "Big Data-oriented Analysis on Issues of the Hyper-connected Society", The E-Business Studies, Vol. 16. No. 5. pp. 3-18, 2015.
DOI: https://doi.org/10.15719/geba.16.5.201510.3.

[4] Jumin Lee, S.B. Park, and S. Lee, "Are Negative Online Consumer Reviews Always Bad? A Two-Sided Message Perspective", Asia Pacific Journal of Information Systems, Vol. 25. No. 4. pp. 784-804, 2015.
DOI: https://doi.org/10.14329/apjis.2015.25.4.784.

[5] J.K. Kim, S.W. Lee, and D.O. Choi, "Relevance Analysis Online Advertisement and e-Commerce Sales", Journal of the Korea Entertainment Industry Association, Vol. 10. No. 2. pp. 27-35, 2016.
DOI: https://doi.org/10.21184/jkeia.2016.04.10.2.27.

[6] S.W. Lee, and S.H. Kim, "Finding Industries for Big Data Usage on the Basis of AHP", Journal of Digital Convergence, Vol. 14. No. 7. pp. 21-27, 2016.
DOI: https://doi.org/10.14400/JDC.2016.14.7.21.

[7] S. Lee, and S.Y. Shin, "Design of Health Warning Model on the Basis of CRM by use of Health Big Data", Journal of the Korea Institute of Information and Communication Engineering, Vol. 20. No. 4. pp. 1460-1465, 2016.
DOI: https://doi.org/10.6109/jkiice.2016.20.8.1460.

[8] M. Nam, and S. Lee, "Big Data as a Solution to Shrinking the Shadow Economy", The E-Business Studies, Vol. 17. No. 5. pp. 107-116, 2016.
DOI: https://doi.org/10.20462/TeBS.2016.10.17.5.107.

[9] S.H. Kim, S. Chang, and S.W. Lee, "Consumer Trend Platform Development for Combination Analysis of Structured and Unstructured Big Data", Journal of Digital Convergence, Vol. 15. No. 6. pp. 133-143, 2017.
DOI: https://doi.org/10.14400/JDC.2017.15.6.133.

[10] Y. Kang, S. Kim, J. Kim, and S. Lee, "Examining the Impact of Weather Factors on Yield Industry Vitalization on Big Data Foundation Technique", Journal of the Korea Entertainment Industry Association, Vol. 11. No. 4. pp. 329-340, 2017.
DOI: https://doi.org/10.21184/jkeia.2017.06.11.4.329.

[11] S. Kim, H. Hwang, J. Lee, J. Choi, J. Kang, and S. Lee, "Design of Prevention Method Against Infectious Diseases based on Mobile Big Data and Rule to Select Subjects Using Artificial Intelligence Concept", International Journal

of Engineering and Technology, Vol. 7. No. 3. pp. 174-178, 2018.
DOI: https://doi.org/10.14419/ijet.v7i3.33.18603.

[12] I. Jung, H. Sun, J. Kang, C.H. Lee, and S. Lee, "Big Data Analysis Model for MRO Business Using Artificial Intelligence System Concept", International Journal of Engineering and Technology, Vol. 7. No. 3. pp. 134-138, 2018.
DOI: https://doi.org/10.14419/ijet.v7i3.33.18593.

[13] S. Kim, S. Park, J. Kang, and S. Lee, "The Model of Big Data Analysis for MICE Using IoT (Beacon) and Artificial Intelligence Service (Recommendation, Interest, and Movement)", International Journal of Engineering and Technology, Vol. 7. No. 3. pp. 314-318, 2018.
DOI: https://doi.org/10.14419/ijet.v7i3.33.21192.

[14] S.H. Kim, J.K. Choi, J.S. Kim, A.R. Jang, J.H. Lee, K.J. Cha, and S.W. Lee, "Animal Infectious Diseases Prevention through Big Data and Deep Learning", Journal of Intelligence and Information Systems, Vol. 24. No. 4. pp. 137-154, 2018.
DOI: https://doi.org/10.13088/jiis.2018.24.4.137.

[15] S. Lee, and I. Jung, "Development of a Platform Using Big Data-Based Artificial Intelligence to Predict New Demand of Shipbuilding", The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 19. No. 1. pp. 171-178, 2019.
DOI: https://doi.org/10.7236/JIIBC.2019.19.1.171.

[16] H. Hwang, S. Lee, S. Kim, and S. Lee, "Building an Analytical Platform of Big Data for Quality Inspection in the Dairy Industry: A Machine Learning Approach", Journal of Intelligence and Information Systems, Vol. 24. No. 1. pp. 125-140, 2018.
DOI: https://doi.org/10.13088/jiis.2018.24.1.125.

[17] Y. Shon, J. Park, J. Kang, and S. Lee, "Design of Link Evaluation Method to Improve Reliability based on Linked Open Big Data and Natural Language Processing", International Journal of Engineering and Technology, Vol. 7. No. 3. pp. 168-173, 2018.
DOI: https://doi.org/10.14419/ijet.v7i3.33.18601.

[18] T. Minami and K. Baba, "A Study on Finding Potential Group of Patrons from Library's Loan Records", International Journal of Advanced Smart Convergence, Vol. 2, No. 2, pp. 23-26, 2013.
DOI: https://doi.org/10.7236/IJASC2013.2.2.6

[19] S.H. Kim, M.S. Kang, and Y.G. Jung, "Big Data Analysis using Python in Agriculture Forestry and Fisheries", International Journal of Advanced Smart Convergence, Vol. 5. No. 1, pp. 47-50, 2016.
DOI: https://doi.org/10.7236/IJASC.2016.5.1.47

[20] W.Y. Kim, "A Practical Study on Data Analysis Framework for Teaching 3D Printing in Elementary School", International Journal of Internet, Broadcasting and Communication, Vol. 8, No. 1, pp. 73-82, 2016.
DOI: https://www.earticle.net/Article/A263475

[21] H.C. Kang, K.B. Kang, H.K. Ahn, S.H. Lee, T.H. Ahn, and J.W. Jwa, "The Smart EV Charging System based on the big data analysis of the Power Consumption Patterns", Vol. 9, No. 2, pp. 1-10, 2017.
DOI: https://www.earticle.net/Journal/Issues/821/22509

[22] Y.I. Kim, S.S. Yang, S.S. Lee, S.C. Park, "Design and Implementation of Mobile CRM Utilizing Big Data Analysis Techniques", The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 14, No. 6, pp. 289-294, 2014.
DOI: https://doi.org/10.7236/JIIBC.2014.14.6.289

[23] S.J. Oh, "Design of a Smart Application using Big Data", The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 15, No. 6, pp. 17-24, 2015.
DOI: https://www.earticle.net/Article/A259710