JKIICE Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지 Vol. 24, No. 5: 600~608, May. 2020

머신러닝을 위한 온톨로지 기반의 Raw Data 전처리 기법

황치곤[1] · 윤창표[2*]

Pre-processing Method of Raw Data Based on Ontology for Machine Learning

Chi-Gon Hwang[1] · Chang-Pyo Yoon[2*]

[1]Invited Professor, Department of Computer Engineering, Institute of Information Technology, Kwangwoon University, Seoul, 01897, Korea
[2*]Assistant Professor, Department Of Computer & Mobile Convergence, GyeongGi University of Science and Technology, Siheung-si, 15073, Korea

요 약

머신러닝은 학습 데이터로부터 목적함수를 구성하고, 테스트 데이터를 통해 목적함수의 확인함으로써 발생하는 데이터에 대한 예측을 수행한다. 머신러닝에서 입력데이터는 전처리 과정을 통해 정규화 과정을 거친다. 이런 정규화는 입력데이터의 평균과 표준편차를 이용하여 표준화하거나, 수치 데이터가 아닌 nominal value는 one-hot 코드 형태로 변환하는 방식을 이용한다. 그러나 이 전처리 과정만으로 문제를 해결할 수 없다. 이러한 이유로 본 논문에서 입력데이터의 정규화를 위해 온톨로지를 이용하는 방법을 제안한다. 이를 위한 테스트 데이터는 모바일 기기로부터 수집된 와이파이 장치의 RSSI값을 이용하고, 수집된 데이터의 노이즈와 이질적 문제는 온톨로지를 이용하여 정제하는 방법을 제시한다.

ABSTRACT

Machine learning constructs an objective function from learning data, and predicts the result of the data generated by checking the objective function through test data. In machine learning, input data is subjected to a normalisation process through a preprocessing. In the case of numerical data, normalization is standardized by using the average and standard deviation of the input data. In the case of nominal data, which is non-numerical data, it is converted into a one-hot code form. However, this preprocessing alone cannot solve the problem. For this reason, we propose a method that uses ontology to normalize input data in this paper. The test data for this uses the received signal strength indicator (RSSI) value of the Wi-Fi device collected from the mobile device. These data are solved through ontology because they includes noise and heterogeneous problems.

키워드 : 온톨로지, 기계학습, 데이터전처리, 입력데이터 정규화, RSSI

Keywords : Ontology, Machine Learning, Data Pre-processing, Input Normalization, RSSI

Received 16 April 2020, **Revised** 29 April 2020, **Accepted** 9 May 2020

*** Corresponding Author** Chang-Pyo Yoon(E-mail:cpyoon@gtec.ac.kr, Tel:+82-31-496-6410)
Assistant Professor, Department Of Computer & Mobile Convergence, GyeongGi University of Science and Technology, Siheung-si, 15073, Korea

Open Access http://doi.org/10.6109/jkiice.2020.24.5.600 print ISSN: 2234-4772 online ISSN: 2288-4165

Ⅰ. Introduction

Machine learning based on real-world data may show a variety of learned results depending on the accuracy of the input data. As such, input data for machine learning is an important factor. Machine learning is based on the amount of learned data for learning, the learning model, and the number of features of the data[1]. Input data is the unprocessed data with real-life information. The accuracy and reliability of the learned results can be improved if the information is used to make sense of these data and then used as learned data for machine learning. In this paper, we propose a method for processing the collected data using ontology and normalizing it by preprocessing for machine learning input. Data preprocessing is a method of classifying, adjusting, and eliminating data that are necessary for learning and unnecessary prior to learning. Preprocessing such information can improve the performance, such as processing speed and accuracy, of the generated objective function[2]. Unnecessary data is called noise, which is an important goal of data preprocessing because it interferes with various types of data analysis[3].

Existing data processing methods cause errors in processing results due to incomplete data collection. Although this focuses on eliminating noise, which is an error in the data results, irrelevant data may have a significant impact on the results of the analysis.

Therefore, if the goal is to improve the accuracy of the results of the data analysis, the noise should be removed through the refinement of the input data rather than correction of the errors in the results. As a result, there is a need for a data refinement technique that removes such noise[4].

This is because the set of collected data may contain a lot of noise. The types of noise can be classified according to the difference in the format of the data, the difference in the structure, and the difference in the meaning implied. The difference in type is the same as type, value, unit, and occurrence time. The difference in structure is caused by a combination of two or more data, or separable data. The difference in meaning is due to words, such as similar words and compound words, caused by semantic heterogeneity. If this is used as machine learning data such as the above noise, the result will have many errors and the reliability of the accuracy of the result will be poor[4].

There is an ontology as an emerging concept of solving the heterogeneous and semantic problems of data. The ontology is a shared concept for the domain, which defines the relationship between data and it is used as the applied knowledge to the intelligent system[5].

This can be used as a way to remove the noise from the collected data. In this paper, we collect Wi-Fi signals received through a mobile device indoors and use this data as input data for machine learning in order to identify the corresponding areas. We use an ontology to solve the noise and heterogeneity issues in the collected data. Furthermore, we describe the process of data change and system configuration[4, 6, 7].

The structure of this paper is as follows. Section 2 describes the ontology and related research on machine learning. Section 3 describes the method and reason for constructing an ontology. Section 4 describes the raw data preprocessing method based on the ontology and system configuration for it. Finally, Section 5 describes the conclusion and future work of this study.

Ⅱ. Relative Work

2.1. Ontology

The definition of an ontology used as an essential element in an intelligent system is described in various fields. Gruber defined ontology as "a formal and explicit specification of the shared conceptualization of the domain". A closer look at this definition reveals that the following four terms are compounded: (a) Conceptualization is an abstraction of the real world in which people think about things, and it is generally discussed only in specific areas; (b) Explicit means that the type of concept or usage constraints are clearly explained; (c)

Formal refers to the method of expressing artificial intelligence, such as predicate logic, that a machine can read and process; and (d) Shared refers to agreed knowledge by all members of the group, and it is not limited to any particular individual. The domain of interest is the concept of a specific domain. What is indicated in the four definitions above is the term ‘ontology’[5].

These ontologies are necessary for semantic reasoning and data heterogeneity, and they are used for interoperability in distributed processing environments of different models[8]. Therefore, ontology improves data accessibility through the standardization of common knowledge, and enables web-based knowledge processing, sharing, and reuse. In addition, ontology provides a common language communication environment for communicating between people and applications, and applies associations for websites that are semantically comparable but structurally distinctive[9].

2.2. Machine learning

Many scholars are currently studying machine learning. Yoshua defined machine learning as "one of the techniques by which a machine learns its own learning data and provides appropriate services for newly collected data"[10]. Meanwhile, Mohri defined machine learning as "a computational method that uses experience to improve performance or predict accurately"[11]. Based on this definition, machine learning refers to a method of solving problems by learning the data that occurs in a specific area by itself. If the current method is a passive method based on knowledge, then machine learning is an active methodology based on prior learning experience. Machine learning is divided into Supervised Learning and Unsupervised Learning by providing label (target value) information. Supervised Learning is the problem of learning to map from x to y when training sets of pairs (xi, yi) are provided. Unsupervised Learning is a problem of estimating the structure or density that is likely to have generated X under the training set X = (x1, x2, ..., xn) and without the target value. Semi-Supervised Learning is a combination of Supervised Learning and Unsupervised Learning. When adding unlabelled data, the algorithm may provide some map information, but not all examples. In this case, semi-supervised learning is used[12].

This machine learning method has a preprocessing technique of data to be conducted before learning is performed with input data. This process includes normalization, nominal value, and one-hot method. However, this technique is insufficient to solve the structural and semantic heterogeneity issues in the data.

Ⅲ. Ontology for Machine Learning

Data generated by various devices has a heterogeneous nature depending on the structure and the constructed metadata information. When using different device types other than the same device, it is difficult to maintain the same data at all times. In this paper, we use the signal input from the Wi-Fi device as information to determine the current location. At this time, Wi-Fi devices may have different signal strengths from the installed device types. An ontology is used to refine the data collected from these devices, and to solve heterogeneous issues based on information, such as time, device, measured location, signal strength, etc., collected from a Wi-Fi device.

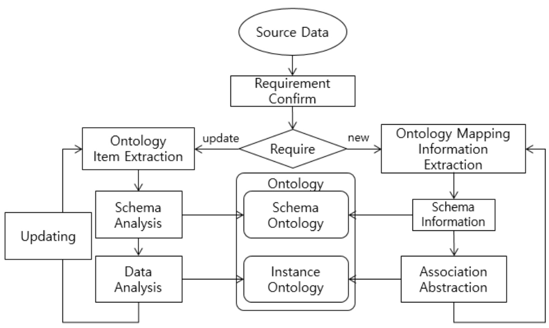

Fig. 1 Ontology Generation Process from Raw Data

Fig. 1 shows the process of creating an ontology from raw data, adding a new ontology for the recognition of new devices, and expressing an additional update process for existing devices. The recognition of a new device is a process of extracting the concept of metadata through schema information, extracting heterogeneity, analysing association, and adding the data. Updating the existing device is a process of updating the ontology by analysing the schema and information of the added data after predicting the existing ontology item. In this case, there is a difference among the data size, unit, qualitative data, and quantitative data, thus resulting in a slow convergence speed of the objective function in machine learning.

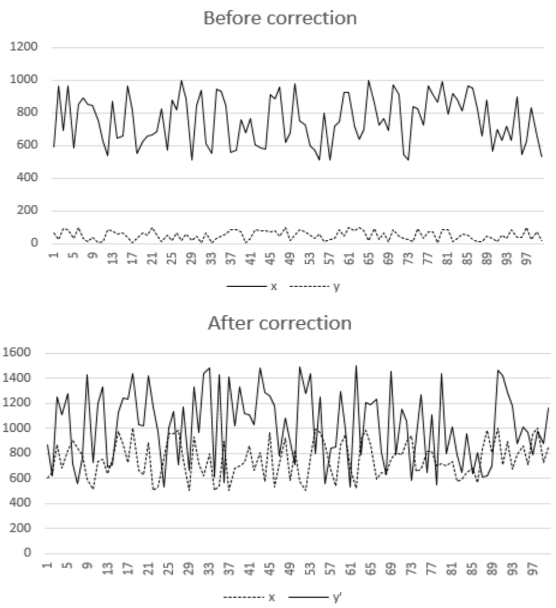

Fig. 2 Before and after correction of the signal for the unit difference

3.1. Problems due to data unit differences

In general, the unit difference between data can be solved by adjusting the units for machine learning. For example, units for length are m for meters and cm for centimeters, units for weights are kg for kilograms and g for grams, units for frequency strength are kHz for kilohertz and Hz for Hertz, etc. There are many other unit differences. Because of the difference between these units, when used in combination, data noise is affected. Fig. 2 shows the calibration in order to solve the problems caused by the above difference in advance. The former is before the correction, whereas the latter is the data change after the correction.

3.2. Problems due to qualitative / quantitative differences

This is a problem that arises from the differences between numerical and numeric data. Data that can be represented numerically, such as weight, height, frequency, etc., can be analyzed through numerical analysis, but data, such as sex and grade, cannot be expressed numerically. Analysis is difficult when these data are combined. This is one of the preprocessing methods used in conventional machine learning.

3.3. Problems with different representations of the same data

This part is the same device, which is a problem that arises when it is recognised as a different device depending on the machine being measured or other problems. There are cases where it is necessary to recognize that data with a similar meaning is the same, and there are cases where it is necessary to recognize them as similar, which is different. The solution for this case can be solved through the definition of affinity in the form of knowledge about similarity. Wi-Fi devices may be labelled differently depending on the equipment used for the measurement. For example, the measured location labels were marked as '326', 'C326', and '326R' when measuring the Wi-Fi signal values in the area used for testing. However, it was measured differently depending on the Wi-Fi equipment and measurement equipment installed. (The Wi-Fi device in the test space uses the Wi-Fi device used in the existing space.) These signals should be recognised as '326' in the representative label to indicate that they are the same area.

Fig. 3 shows the relationship for expressing the ontology to solve these heterogeneous issues. Expressed relations are relationships between schemas, relationships between data, structural heterogeneity, unit heterogeneity, etc., and these are solved through the process shown in Fig. 1.

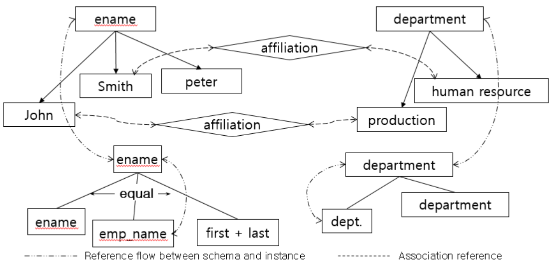

Fig. 3 Expression of relationship between schema information and instance information for ontology

Ⅳ. Raw Data Pre-processing

We install multiple Wi-Fi devices in a building as an environment for data collection. In order to collect the data received from the Wi-Fi device, we generate a label to identify the recognised area using the Wi-Fi signal, and define the meaning of the collected data through the association between the generated labels. The ontology used in this paper is constituted by the method presented in Section 3 by establishing an association based on heterogeneous issues with the system construction environment. This ontology refines the collected data and adjusts it to machine learning input data. Fig. 4 shows the preprocessing process for constructing the data collected from the Wi-Fi device to the mobile terminal as machine learning input data. Data preprocessing is divided into three areas, namely, Source Data Recogniser (SDR), Data Controller (DC), and Data Selector (DS).

These three areas play the following roles:

- SDR accumulates the original data collected from the mobile device, analyzes the accumulated data, and filters out duplicate or abnormal data through an analysis process.
- DC is a data regulator that extracts and removes noise-causing data from SDR, and determines and extracts data necessary for the analysis. The extracted data is merged and normalized.
- DS is merged in the DC, accumulating the regular data, interpreting the relationship between the data, and verifying the data.

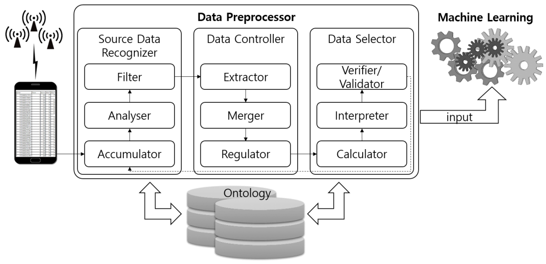

Fig. 4 Ontology-based data preprocessing for machine learning

In this process, the SDR analyser, filter, DC regulator, DS interpreter, Verifier/Validator is refined and relatively modified by using ontology. This process describes the process based on the collected data.

Table. 1 The raw data for the Wi-Fi signal strength collected at each place in the building

floor	store	f8:e7:1e:cc:83:4c	f8:e7:1e:8c:75:08	f8:e7:1e:4c:6b:d8	12:07:89:14:6b:30	f8:e7:1e:cc:36:d8	f8:e7:1e:8c:6b:d8	00:12:17:74:a:c9	f8:e7:1e:cc:7e:0c	f8:e7:1e:4c:75:08	f8:e7:1e:8c:71:48	f8:e7:1e:8c:36:dc
3	326	-86	-72		-62	-62		-79		-83		
3	326	-89	-70		-61	-61		-81				
3	c326	-89	-74	-79	-62	-60	-81	-80		-82		
3	326	-87	-70		-61	-61	-82	-80				
3	c324	-89	-68	-79	-61	-60	-81					-85
3	324	-88		-78	-62	-61	-83					
1	116	-79						-79				-82
1	116	-84			-78	-80		-74				
1	116	-84			-75	-80		-73				-82
1	116	-85			-80							-85
1	117	-83			-83	-82		-73	-89			-86
1	117	-80			-81			-74				-84
1	117	-81						-75				-85
2	225	-79						-69				-87

Table 1 is the data collected from the Wi-Fi device and received by the mobile device that corresponds to the RSSI value. In this paper, SSID, MAC, and RSSI are used to identify the location of the signal measured by the Wi-Fi device. In addition, floor refers to the floor receiving the Wi-Fi signal, while store represents the room number of the measured building. This signal is acquired by an Accumulator, and then measured and received by the mobile device. The Wi-Fi device used for testing was the data received from 494 devices, and the number of measurements was 7,000 times. Table 1 shows the sample groups.

The Wi-Fi signal consists of SSID, BSSID, MAC, Supplicant state, RSSI, Link SPEED, and Net ID. SSID is the identifier of the Wi-Fi device being accessed. MAC refers to a unique identifier of the Wi-Fi device being accessed. RSSI refers to the strength of the received signal. This appears as a negative value due to the nature of the signal measurement itself, and the higher the intensity, the higher the intensity is perceived. This signal usually determines whether there is a valid signal based on -80 dBm. A value higher than -80 dBm is recognised as a valid signal.

Table. 2 Filtering for Wi-Fi devices that do not affect positioning

floor	store	f8:e7:1e:cc:83:4c	f8:e7:1e:8c:75:08	f8:e7:1e:4c:6b:d8	12:07:89:14:6b:30	f8:e7:1e:cc:36:d8	f8:e7:1e:8c:6b:d8	00:12:17:74:a:c9	f8:e7:1e:cc:7e:0c	f8:e7:1e:4c:75:08	f8:e7:1e:8c:71:48	f8:e7:1e:8c:36:dc
3	326	-86	-72		-62	-62		-79		-83		
3	326	-89	-70		-61	-61		-81				
3	c326	-89	-74	-79	-62	-60	-81	-80		-82		
3	326	-87	-70		-61	-61	-82	-80				
3	c324	-89	-68	-79	-61	-60	-81					-85
3	324	-88		-78	-62	-61	-83					
1	116	-79						-79				-82
1	116	-84			-78	-80		-74				
1	116	-84			-75	-80		-73				-82
1	116	-85			-80							-85
1	117	-83			-83	-82		-73	-89			-86
1	117	-80			-81			-74				-84
1	117	-81						-75				-85
2	225	-79						-69				-87

Table 2 shows the data extracted from the raw data accumulated in the data of Table 1 through the SDR process. The raw data shown in Table 1 are all Wi-Fi signals generated in the building. It is difficult to select the location based on this data alone. In this case, there are signals that render positioning difficult, such as a Wi-Fi MAC signal received at all positions, a MAC signal where no signal is received, and a MAC signal its reception is not constant. These signals are determined by the Analyzer. The MAC signals received at all locations (f8:e7:1e:cc:83:4c, 12:07:89:14:b5:30, f8:e7:1e:cc:36:d8, 00:12:17:74:fa:c9, f8:e7:1e:cc:3a:e8, f8:e7:1e:cc:78:fc, and f8:e7:1e:cc:36:fc) must be extracted and removed since it is not used to measure the location caused by the strong signal strength outside the building or for a specific purpose. Signals for which no signal is received (f8:e7:1e:8c:71:48) must be removed as they are not valid for the region classification because no signal is received within that region. However, Wi-Fi signals, when reception is inconsistent, determine the presence or absence of a signal reception according to the location. Therefore, Wi-Fi signal is used for location selection via filtering. SDR was constructed in order to perform these tasks.

Table. 3 Assigning a null value to a Wi-Fi signal for which no signal is received

(a) Filtered data

floor	store	f8:e7:1e:8c:75:08	f8:e7:1e:4c:6b:d8	f8:e7:1e:8c:6b:d8	f8:e7:1e:cc:7e:0c	f8:e7:1e:4c:75:08	f8:e7:1e:8c:36:dc
3	326	-72				-83	
3	326	-70					
3	c326	-74	-79	-81		-82	
3	326	-70		-82			
3	c324	-68	-79	-81			-85
3	324		-78	-83			
1	116						-82
1	116						
1	116						-82
1	116						-85
1	117				-89		-86
1	117						-84
1	117						-85
2	225						-87

(b) Data with adjusted values

floor	store	f8:e7:1e:8c:75:08	f8:e7:1e:4c:6b:d8	f8:e7:1e:8c:6b:d8	f8:e7:1e:cc:7e:0c	f8:e7:1e:4c:75:08	f8:e7:1e:8c:36:dc
3	326	-72	1	1	1	-83	1
3	326	-70	1	1	1	1	1
3	c326	-74	-79	-81	1	-82	1
3	326	-70		-82	1	1	1
3	c324	-68	-79	-81	1	1	-85
3	324	1	-78	-83	1	1	1
1	116	1	1	1	1	1	-82
1	116	1	1	1	1	1	1
1	116	1	1	1	1	1	-82
1	116	1	1	1	1	1	-85
1	117	1	1	1	-89	1	-86
1	117	1	1	1	1	1	-84
1	117	1	1	1	1	1	-85
2	225	1	1	1	1	1	-87

Table 3 shows the method of changing the data in the process of performing the DC process by using the data of Table 2, where (a) is the data changed after filtering through the SDR process and (b) is the data extracted and normalized through the DC process. The process from (a) to (b) is performed through the DC Extractor, Merger, and Regulator. The Regulator recognizes the data (Null), for which the RSSI value is not received, and converts it to a value (1) that does not affect the data processing of this paper. The reason for the conversion is to achieve a degree to which the measured value is evenly distributed. An even distribution of values indicates a Wi-Fi signal at a given location, which maintains a constant signal strength. Receiving a null value means correcting the signal by increasing the standard deviation value, assuming that the signal is unreliable.

For this reason, a signal with a low standard deviation of the Wi-Fi signal value measured at the same location is obtained, and learning data is obtained in such a way that a position with a low standard deviation value can be recognized as the corresponding position. Table 3 (a) shows the measurement location label as the stored items. According to the setting of the existing Wi-Fi device, 326 and c326 are the same values from the point of view, but they can be recognized as different values in machine learning. These parts are recognized as the same data using the association data of the ontology. We emphasized that it is necessary to make adjustments, such as different expressions of the same value and representation by unit difference, as described in Section 3. The application of this is shown in Table 3, which is performed by a DC regulator.

Table. 4 Labelling with normalization process using ontology and select representative signals for each label

floor	store	f8:e7:1e:8c:75:08	f8:e7:1e:4c:6b:d8	f8:e7:1e:8c:6b:d8	f8:e7:1e:cc:7e:0c	f8:e7:1e:4c:75:08	f8:e7:1e:8c:36:dc
3	326	-72	1	1	1	-83	1
3	326	-70	1	1	1	1	1
3	326	-74	-79	-81	1	-82	1
3	326	-70	1	-82	1	1	1
stdev		1.91	40	47.63	0	48.21	0
3	324	-68	-79	-81	1	1	-85
3	324	1	-78	-83	1	1	1
stdev		40.69	59.34	64.82	0.5	23.61	42.84
1	116	1	1	1	1	1	-82
1	116	1	1	1	1	1	1
1	116	1	1	1	1	1	-82
1	116	1	1	1	1	1	-85
stdev		0	0	0	0	0	42.0238
1	117	1	1	1	-89	1	-86
1	117	1	1	1	1	1	-84
1	117	1	1	1	1	1	-85
stdev		0	0	0	51.96	0	1

Table 4 shows the standard deviations of the RSSI values of each Wi-Fi device by grouping them based on the stores through the ontology. It is based on data generated by refining labelled stores using ontology and transforming null values. The Calculator obtains the value of the standard deviation of the RSSI signals, which is the signal strength of the Wi-Fi devices, based on each store. The standard deviation represents the degree of dispersion of data based on the used data. The smaller the standard deviation, the less the variation in the mean. This indicates that the Wi-Fi signals are constantly generated. In other words, this signal is an indication that the store can be confirmed. In the preprocessor of Fig. 4, the Calculator of the DS obtains the standard deviation of the signal value among Wi-Fi devices based on each store. The Interpreter selects the representative signal by comparing the standard deviation between the Wi-Fi devices that are interactively measured. The Verifier/Validator identifies the problem that occurs in the selection of a selected device and plays the role of feedback. Therefore, the data constructed is used as input data for machine learning.

Ⅴ. Conclusion

The proposed method starts by recognizing the importance of training data and test data used in machine learning. As part of our ongoing project, we needed to identify the location of the store by using the strength of the Wi-Fi signal recognised by the mobile device in a specific area. For this purpose, Wi-Fi signals measured in mobile devices are used to specify the positions through machine learning. However, there was insufficient collected data to use the learned objective function due to the large amount of noise, and it took a long time to learn. In order to solve this problem, we need a method to filter the difference between semantic relations and measured values using ontology. For this reason, we applied the proposed method. The adaptation rate for the objective function after application has increased as opposed to not using this method, the time applied has improved, and the effect of reducing the size of the input data of machine learning has been shown. However, this is a technique applied only to a specific purpose, so it needs to be generalized, and the performance evaluation should be performed by presenting quantitative test results. In the future, we will complement these disadvantages so that we will be able to identify the need to refine the input data.

REFERENCES

[1] L.C.Navarro, A.K.W. Navarro, A.Gregio, A.Rocha, and R.Dahab, "Leveraging ontologies and machine-learning techniques for malware analysis into Android permissions ecosystems," *Computers & Security*, vol. 78, pp. 429-453, Sep. 2018.

[2] X.Wang, and H.C.Kim, "Text Categorization with Improved Deep Learning Methods," *Journal of information and communication convergence engineering*, vol. 16, no.2, pp.106-113, 2018.

[3] H. Xiong, G. Pandey, and M. Steinbach, V. Kumar, "Enhancing data analysis with noise removal," *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 3, pp.304- 319, 2006.

[4] A. H. Doan, J. Madhavan, P. Domingos, and A. Halevy, "Ontology matching: A machine learning approach," In Handbook on Ontologies. Springer, Berlin, Heidelberg. pp. 385-403, 2004.

[5] T.R.Gruber, "Towards principles for the design of ontologies used for knowledge sharing," *International Journal of Human-Computer Studies*, vol. 43, no. 5, pp.1-2, 1995.

[6] C. P. Yoon, and C. G. Hwang, "Efficient indoor positioning systems for indoor location-based service provider," *Journal of the Korea Institute of Information and Communication Engineering*, vol.19, no.6, pp.1368-1373, 2015.

[7] M. S. Devi, and H. Mittal, "Machine learning techniques with ontology for subjective answer evaluation," *International Journal on Natural Language Computing (IJNLC)*, vol. 5, no. 2, 2016.

[8] B. Amann, C. Beeri, I. Fundulaki, and M. Scholl, "Ontology- Based Integration of XML Web Resources," In International Semantic Web Conference (ISWC), pp.117-131, 2002.

[9] Architecture Committee, "Semantic Web Services Architecture requirements," Working draft, Version 1.0, [Internet]. Available: http://www.daml.org/services/swsa/swsa-requirements.html, 2004.

[10] Y.Bengio, "Learning deep architectures for AI," *Foundations and trends in Machine Learning*, vol.2, no.1, pp.1-127, 2009.

[11] A.Singh, "Foundations of Machine Learning," Available at SSRN 3399990, 2019.

[12] T. Miyato, S. Maeda, M. Koyama, and S. Ishii, "Virtual adversarial training: a regularization method for supervised and semi-supervised learning," *IEEE transactions on pattern*

analysis and machine intelligence, vol. 41, no. 8, pp.1979-1993, 2018.

황치곤(Chi-Gon Hwang)

2012년 광운대학교 컴퓨터과학과 (공학박사)
2006년~2015년:(주)인찬 연구원
2016년~2018년: 경민대학교 인터넷정보과 교수
2019년~현 재: 광운대학교 정보과학교육원 컴퓨터공학과 교수
※관심분야 : 모바일 클라우드, 멀티미디어 온톨로지, 기계학습, NLP

윤창표(Chang-Pyo Yoon)

2012년 : 광운대학교 컴퓨터과학과 (공학박사)
2012년~현 재: 경기과학기술대학교 컴퓨터모바일융합과 교수
※관심분야 : 기계학습, 모바일 시스템, 네트워크 보안, 무선 네트워크, 온톨로지