논문 2018-13-21

규칙기반 초미세먼지 상태 추론

(Particulate Matter (PM2.5) State Inference by Rule Induction)

최 락 현, 강 원 석, 손 창 식*
(Rock-Hyun Choi, Won-Seok Kang, Chang-Sik Son)

Abstract : Particulate Matter (PM2.5) has various adverse effects on health. Climate and industry activity and traffic volume are the main causes, especially in urban area. In order to construct an effective forecasting system, many measurement systems are required, but it is impossible in reality. Therefore, in this study, we propose a method to infer PM2.5 condition by using rule induction technique. The experimental results showed a classification accuracy of 71%.

Keywords : Air quality, Weather, Data mining, Rule induction, PM2.5

Ⅰ. 서 론

높은 농도의 초미세먼지 (PM, Particulate matter 2.5)는 건강에 다양한 영향을 끼친다. 2016년 WHO 보고에 의하면 매년 공기오염 관련 질병 사망이 600만 명을 초과하는 것으로 예측됐다. 더군다나 우리나라는 OECD 34국 중 터키에 이어 두 번째로 심한 것으로 보고되었다. 2015년도부터 시행된 연간 대기환경기준을 보면 대부분의 지역에서 기준치를 초과하고 있다 [1]. 이런 미세먼지는 천식과 같은 호흡기, 폐질환 및 심혈관계 질환 등을 유발하여 건강에 나쁜 영향을 끼친다 [2]. 불행히도 건강에 영향을 끼치는 미세먼지 (PM10) 예보는 2013년에 시작되었고 초미세먼지 (PM2.5)는 2015년에 예보를 시작했다. 최근 에어코리아 사이트를 통해 데이터가 공유되고 있다. 예보의 정확도는 농도를 측정값과 비교하고 적중률이나 감지 확률 등을 계산하여 조사되고 있다 [3].

미세먼지와 초미세먼지의 원인은 자연적, 인위적 요인에 의해 발생하는데 특히 산업 활동, 화석연료 등의 인위적인 요인에 많이 발생된다고 보고 있다. 최근 연구 조사에 따르면 서울시에서 시행중인 미세먼지 저감 대책 사업 이전 2002년 2007년에 교통량 증가와 함께 미세먼지 농도 증가 확인 되었다 [4]. 기후 변화에 따른 공기질 영향 관련 연구에 따르면 가장 큰 영향을 주는 것으로 지형, 그리고 습도로 보고 있다 [5]. 이런 초미세먼지 문제와 같은 대책을 수립하기 위해서는 측정이 중요한데 아직 인구밀도가 높은 도심에 주로 설치되어 있고 서울에도 25개소 밖에 없어 효과적인 측정망 구축이 되지 못했다 [6].

현실적인 문제로 효과적인 예보시스템을 구축할 수 없으면 예측하는 연구 역시 병행되어져야 한다. 따라서 본 논문에서는 기후정보를 근거로 초미세먼지 상태를 예측할 수 있다고 가정 후 실험하였다. 초미세먼지 상태를 추론하기 위해서 인공지능 분야의 지식획득 및 처리방법 중에 하나인 LEM1 규칙 생성 알고리즘을 활용하였다. 본 연구에서 초미세먼지 상태를 추론하기 위한 흐름도는 그림 1과 같다.

Ⅱ. 연구 방법

1. 실험환경

데이터 분석을 위한 소프트웨어 구성은 다음과 같다. 운영체제는 64비트 환경의 Windows10, 개발언어로 파이썬 3.5.2를 사용하고 외부 라이브러리로 pandas 0.18.1이 사용되었다. 하드웨어 구성으로 메인프로세서는 i7-6700 3.4GHz, 메모리는 32.0GB, 하드디스크는 SSD 500GB를 사용하였다.

*Corresponding Author (changsikson@dgist.ac.kr)
Received: Jan. 31 2018, Revised: Mar. 31 2018, Accepted: Apr. 3 2018.
R. Choi, W. Kang, C. Son: DGIST
※ 본 연구는 산업통상자원부에서 지원하는 산업핵심기술개발사업 (10063553)과 미래창조과학부의 대구경북과학기술원 기관고유사업 (18-IT-02)에 의해 수행되었습니다.

ⓒ IEMEK J. Embed. Sys. Appl. 2018 Aug. 13(4) 179-185
ISSN : 1975-5066
http://dx.doi.org/10.14372/IEMEK.2018.13.4.179

표 1. 수집된 데이터 특성
Table 1. Collected data characteristics

No	Variable	Mean±SD	No	Variable	Mean±SD
1	Temp (°C)	14.04±4.51	13	Snowfall_3H (cm)	NaN±NaN
2	Rain (mm)	1.01±1.52	14	Amount of cloud (10)	5.09±3.88
3	Wind speed (m/s)	2.43±1.31	15	Middle low amount of cloud (10)	2.82±3.46
4	Wind direction (16)	216.01±105.43	16	Lowest height of clouds (100m)	13.29±9.60
5	Humidity (%)	54.79±23.08	17	Visibility (10m)	1506.38±615.21
6	Stemp pressure (hPa)	8.29±2.69	18	Ground condition code	0.55±0.50
7	Dew point temp (°C)	3.59±4.71	19	Phenom number (domestic)	2410.78±27303.47
8	Local surface pressure (hPa)	1003.14±5.21	20	Ground temp (°C)	16.83±9.34
9	Sea surface pressure (hPa)	1013.34±5.29	21	5cm ground temp (°C)	15.45±5.16
10	Sunshine (hr)	0.51±0.46	22	10cm ground temp (°C)	14.86±3.06
11	Irradiation (MJ/m2)	1.07±0.92	23	20cm ground temp (°C)	14.49±1.93
12	Snowfall (cm)	NaN±NaN	24	30cm ground temp (°C)	13.52±1.54

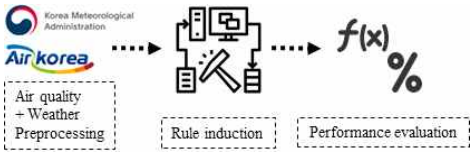


그림 1. 데이터분석 처리 과정
Fig. 1 Data analysis process

2. 실험데이터와 실험방법

사용된 데이터는 크게 두 가지로 분류되는데 기상청 기후 데이터와 에어코리아의 공기질 데이터다. 기상청 데이터는 기상자료개방포털에서 공개된 서울특별시의 2016년 4월 자료와 2017년 4월 자료를 각각 수집하였다. 2016년 4월 자료의 기상 데이터 특성은 표 1과 같고 [7], 그림 2는 이들 데이터 특성에 대한 히스토그램 분포를 보여준다. 또한 공기질 데이터는 에어코리아를 통해 수집되었고 날씨 데이터와 같이 2016년 4월과 2017년 4월 데이터를 수집했다. 공기질 데이터를 통해 수집할 수 있는 데이터는 각각 PM10, PM2.5, 오존, 일산화질소, 일산화탄소, 아황산가스이다. 그 중 본 연구에서 사용된 데이터는 PM2.5 초미세먼지의 등급 (“좋음”, “보통”, “나쁨”, “매우나쁨”)을 사용하였다. 공기질이 측정된 위치는 한국환경정책평가연구원이 위치한 서울 은평구 진흥로 215번지이며, 2016년과 2017년 4월 한 달 동안 매시간 측정된 데이터 1440건의 데이터를 수집했다 [8]. 그 중 등급분류가 결측인 부분을 제외한 데이터를 본 실험에 활용하였다. 규칙생성을 위해 사용된 데이터의 분포는 그림3과 같다. 정규분포에 가까운 특성을 가진 속성 값으로 5cm 지중온도, 10cm 지중온도, 20cm 지중온도, 30cm 지중온도, 온도, 습도, 풍속, 이슬점온도, 지면온도, 현지기압, 해수면기압, 증기압 정도가 있다. 2016년 4월에는 눈이 오지 않았기 때문에 적설량 데이터는 존재 하지 않았기 때문에 차트에서 데이터 분포를 확인할 수 없다.

표 2. 전반적인 흐름 의사코드
Table 2. Overall flow pseudo-code

```
1  BEGIN
2    x_tr, y_tr = dataRead()
3    x_categorized = discretize(x_tr)
4    rules = rule_induction(x_categorized)
5    classify(rules, x_test)
6    evaluator(y_test)
7  END
```

전반적인 실험은 표 2 의사코드 (Pseudo-code)와 같이 수행된다.

그림 3은 의사코드를 클래스 다이어그램으로 표현한 그림으로 실행함수는 LEM1_algorithm_start이고 5가지 모듈 (즉 File_Reader, MAD, LEM1,

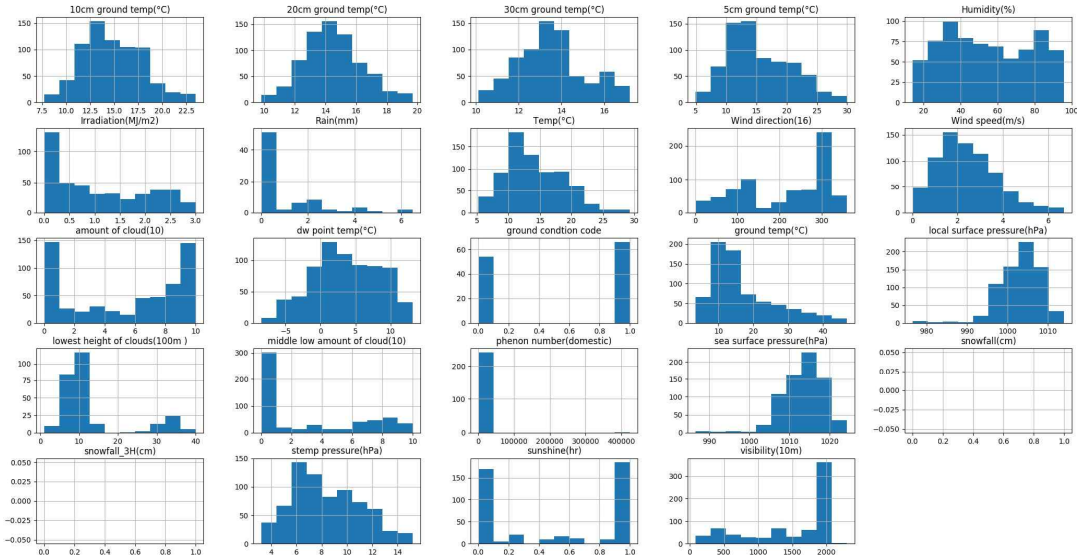


그림 2. 날씨 데이터 히스토그램
Fig. 2 Weather data histogram

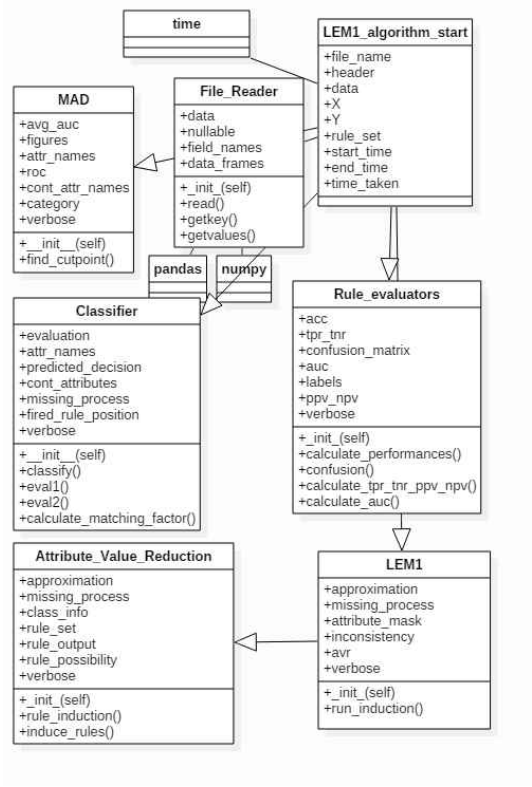


그림 3. LEM1 클래스 다이어그램
Fig. 3 LEM1 class diagram

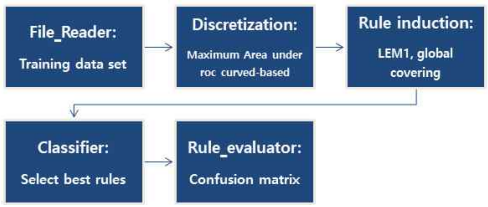


그림 4. 함수 흐름 블록 다이어그램
Fig. 4 Functional flow block diagram

Rule_evaluators, Classifier)과 상속관계에 있으며, LEM1 모듈은 Attribute_Value_Reduction 모듈과 상속관계로 구성된다.

그림 4는 각 모듈의 흐름을 나타내며 5가지의 동작 순서를 블록 다이어그램으로 표현하였다.

전체적인 흐름을 요약하면 File_Reader 모듈을 통해 훈련 데이터 셋을 로드하고, 이산화알고리즘을 통해 데이터의 구간 의미화를 진행한 후, Rule induction 모듈을 통해 규칙을 생성한다. 그 후 Classifier 모듈을 통해 테스트 데이터를 분류하고 Rule_evaluator 모듈의 confusion matrix로 산출된 규칙의 성능을 평가하였다. 알고리즘 선정이유와 각 모듈에 대한 설명은 다음과 같다.

많은 기계학습 알고리즘들이 그렇듯 모든 수치 값으로 표현된 데이터 범주화해야 하며 이를 데이터 이산화 (Data discretization)라 한다. 실험에서

규칙 생성을 위한 이산화 과정에서는 2016년과 2017년 4월 기상 데이터를 모두 이용하였고, 초미세먼지 상태 추론을 위한 규칙은 훈련 데이터 셋으로 2016년 4월의 기상 데이터를, 실험 데이터 셋으로 2017년 4월의 기상 데이터를 이용하였다. 본 연구에서는 다중-클래스 ROC (Receiver operating characteristic) 분석방법 [9]을 이용하여 이산화 과정을 수행하였다. MAD 알고리즘의 선정 이유는 2가지 대표적인 이산화 알고리즘 (즉 MDLP [10]와 Chimerge [11]) 보다 보편적으로 우수한 성능을 보여주며, 통계적 평가척도로서 ROC (Receiver Operating Characteristic)와 AUC (Area Under the ROC) 값을 사용하기 때문이다.

또한 규칙생성 알고리즘으로는 훈련 데이터 셋에서 비일관적인 특성을 가진 패턴, 즉 동일한 입력 패턴에 대해서 서로 다른 출력을 포함하는 샘플의 데이터로부터 규칙을 유도할 수 있는 장점을 가진 LERS (Learning from Examples based on Rough Sets)의 global covering 기법인 LEM1 알고리즘 [12]을 이용하였다. LEM1은 소모적인 탐색 (Greedy search) 전략을 사용하고 검색 공간이 모든 속성값 집합이다. 지식을 생성하기 위해 규칙을 유도하는 방법에는 커버링 탐색전략 (AQ family, LERS systems, RULES family 등) 방법과 분할정복 (ID3, C4,5 등) 방법으로 구분된다. 의사결정트리 (Decision tree)로 불리는 유명한 알고리즘인 ID3 는 탑다운 방식의 탐색전략을 사용하며 한 번에 모든 클래스를 다룬다. 이에 ID3와 같은 방식은 가지치기 (Pruning) 깊이에 따라 결과가 매우 복잡하고 이해하기 힘들 때가 있다. 반면에 LEM1은 한 번에 하나의 클래스에 집중하여 산출된 규칙을 검토하는데 효과적이고 전역적 탐색을 통해 추가적인 특징선택 (Feature selection)과정을 필요로 하지 않는 장점을 가진다 [13].

표 3에서, 속성 혹은 속성들의 하위집합 B의 동치관계 (Indiscernibility relation), 즉 IND(B)는 샘플 x와 y의 모든 값이 동일할 때를 의미하며, 전체 집합 U에 대해 $(x, y) \in$ IND(B)는 $x, y \in$ U와 같이 정의된다 [12].

[정의 1] U: 샘플들에 대한 전체집합 (universe)

[정의 2] B: 입력속성들의 부분집합 (subset)

그러므로 LEM1 알고리즘은 위의 개념을 활용하여 동치관계를 나타내는 모든 동치류 (Equivalence classes) 즉 일관성 있는 속성 값들의 쌍들을 찾는 방법을 나타낸다. 기호 '*'는 위에서 논의된 것처럼, 동치관계를 만족하는 각 데이터 샘플들에 대한 동치류를 나타낸다. 그러므로 첫 번째 조건문인 if $A^* \subseteq \{d\}^*$ 의미는 임의의 속성에 대한 하위집합 A의 동치류 (A^*)가 출력 속성 (즉 의사결정 속성)의 동치류 ($\{d\}^*$)에 포함될 때, 일관성 척도에 위배되지 않음을 의미한다. 따라서 begin부터 시작되는 반복문은 전체 입력 속성들 가운데에 불필요한 속성을 하나씩 제거하면서 출력 속성의 동치류를 만족하는 입력 속성의 하위집합을 탐색하는 과정을 보여준다 [12].

표 3. LEM1 의사코드

Table 3. LEM1 pseudo-code

```
input: the set A of all attributes, partition {d }* on U
output: a single global covering R
1  begin
2  compute partition A*
3  P : = A
4  R := ∅
5      if A* ⊆ {d}*
6          then
7                  begin
8                  for each attribute a in A do
9                      begin
10                     Q := P − {a };
11                      compute partition Q*
12                     if Q* ⊆ {d }* then P := Q
13                  end {for }
14                  R := P
15             end {then }
16 end {algorithm }
```

또한 본 연구에서는 규칙기반 분류기를 설계하기 위해서, LEM1 알고리즘을 통해 생성된 규칙집합과 실험 데이터의 속성들 간에 유사도 (Similarity degree)를 평가하여 추론 결과를 추정하였다. 이를 위해 Minkowski의 거리척도(Distance metric) 개념을 활용하였고, k개의 후보규칙들 가운데에 가장 높은 유사도를 나타낸 후보규칙의 출력을 주어진 샘플의 출력으로 결정하였다.

Ⅲ. 실험결과

훈련 및 실험 데이터 셋에서 초미세먼지 상태 (혹은 등급)가 존재하지 않는 모든 샘플은 제거하였고, 이산화 결과에 따라 각 변수는 최소 3개에서 최대 6개의 데이터 범주로 구분하였다. 그로인해 2016년 4월의 기상 데이터 셋은 688개의 샘플 데이터로, 2017년 4월의 기상 데이터 셋은 665개의 샘플 데이터로 재구성되었다. 또한 추론 과정에서

각 속성 값에서 발견된 손실값 (Missing value)들은 해당 특징 (혹은 속성)값만을 고려하지 않는 “LOST” 방식으로 처리하였다 [13].

688개의 훈련 데이터로부터 생성된 규칙은 199개이었고, 초미세먼지 상태 별 규칙 수는 다음과 같았다: “나쁨” 7개, “좋음” 43개, “보통” 149개. 그 중 빈도가 가장 높은 상위 5개의 규칙의 조건을 확인한 결과 다음과 같았다.

1. 미세먼지 상태 “나쁨”

Rule 3) IF 67.5<습도<83.5 AND 10.0<현상번호<1901.5 AND 15.05<5cm지중온도<19.75 AND 7.7<10cm지중온도<18.55 AND 14.15<20cm지중온도<15.55 AND 10.1<30cm지중온도<13.55

Rule 6) IF 67.5<습도<83.5 AND 986.5<해면기압<1015.95 AND 10.0<현상번호<1901.5 AND 7.7<10cm지중온도<18.55 AND 7.7<20cm지중온도<15.55 AND 10.1<30cm지중온도<13.55

Rule 1) IF 4.5<중하층운량<63.5 AND 10.0<현상번호<1901.5 AND 7.7<10cm지중온도<18.55 AND 7.7<20cm지중온도<15.55 AND 10.1<30cm지중온도<13.55

Rule 2) IF 0.5<중하층운량<4.5 AND 10.0<현상번호<1901.5 AND 15.05<5cm지중온도<19.75 AND 14.15<20cm지중온도<15.55

Rule 5) IF 10.0<현상번호<1901.5 AND 14.15<20cm지중온도<15.55 AND 10.1<30cm지중온도<13.55

2. 미세세먼지 상태 “좋음”

Rule 23) IF 0.0<풍속<1.85 AND 0.0<풍향<330.0 AND 19.75<5cm지중온도<30.0 AND 9.8<20cm지중온도<14.15 AND 10.1<30cm지중온도<13.55

Rule 2) IF 986.5<해면기압<1015.95 AND 0.685<일사<3.0 AND 0.0<전운량<0.5 AND 0.5<지면상태<1.0 AND 4.9<5cm지중온도<15.05

Rule 16) IF 986.5<해면기압<1015.95 AND 0.005<일사<0.685 AND 4.5<중하층운량<6.5 AND 2.0<최저운고<11.0 AND 9.8<20cm지중온도<14.15

Rule 18) IF 2.0<최저운고<11.0 AND 0.5<지면상태<1.0 AND 7.7<10cm지중온도<18.55 AND 14.15<20cm지중온도<15.55 AND 10.1<30cm지중온도<13.55

Rule 27) IF 0.0<풍속<1.85 AND 1015.95<해면기압<1017.75 AND 18.15<10cm지중온도<23.5 AND 14.15<20cm지중온도<15.55 AND 10.1<30cm지중온도<13.55

3. 미세먼지 상태 “보통”

Rule 74) IF 8.5<이슬점온도<9.35 AND 976.5<현지기압<1003.95 AND 4.9<5cm지중온도<15.05 AND 14.15<20cm지중온도<15.55 AND 13.55<30cm지중온도<15.85

Rule 2) IF 0.0풍속<1.85 AND 14.0<습도<67.5 AND 0.0<전운량<0.5 AND 4.9<5cm지중온도<15.05

Rule 3) IF 14.0<습도<67.5 AND 976.5<현지기압<1003.95 AND 0.0<중하층운량<0.5 AND 10.1<30cm지중온도<13.55

Rule 8) IF 0.5<전운량<7.5 AND 0.0<중하층운량<0.5 AND 19.75<5cm지중온도<30.0 AND 10.1<30cm지중온도<13.55

Rule 10) IF 0.0<중하층운량<0.5 AND 7.7cm지중온도<18.15 AND 14.15<20cm지중온도<15.55 AND 10.1<30cm지중온도<13.55

제안된 방법의 예측 정확도는 혼동행렬 (Confusion matrix)을 이용하여 평가하였다.

$$ACC=(TP+TN)/(TP+FN+TN+FP)*100 \quad (1)$$

식 (1)에서 *TP*, *TN*, *FP*, 그리고 *FN*은 각각 True Positive, True Negative, False Positive, ‘False Negative’를 나타내고, *TP*와 *TN*의 합의 비율은 예측출력(Predicted output)이 목표출력 (Target output)을 정확하게 분류한 경우를 의미하고, FP와 FN은 예측출력이 목표출력을 서로 다르게 예측한 경우를 의미한다.

2017년 4월 초미세먼지의 전체 추론 정확도는 71%이었다. 선행연구 [3]에서 보고된 초미세먼지 예보모델 (WRF v3.3과 CMAQ v4.7.1 활용한 경우 68%)과 간접적인 성능을 비교해 볼 때, 본 연구에서 제안된 방법이 약 3% 개선된 결과를 보였다.

표 4. 2017년 4월 초미세먼지 추론 결과
Table 4. PM2.5 inference results in Apr. 2017

Confusion matrix		Actual class		
		Bad	Good	Normal
predicted class	Bad	0	9	2
	Good	0	11	53
	Normal	7	119	464

Ⅳ. 결 론

초미세먼지와 미세먼지가 건강에 다양한 영향을 미치고 있다. 원인은 다양한데 특히 인구 밀도가 높은 도심에서는 기후 및 산업 활동과 교통량을 주요 원인으로 볼 수 있다. 효과적인 예보 시스템을 구축하기 위해서는 다수의 측정망이 필요한데 현실적인 문제가 있어 초미세먼지상태 예측 연구의 필요성이 커지고 있다. 따라서 본 연구에서는 인공지능 분야의 지식획득 및 처리방법 중에 하나인 LEM1 규칙 생성 알고리즘을 활용하여 2016년 4월의 기상 데이터를 활용해 2017년 4월의 초미세먼지 상태를 추론 하는 접근방법을 제안하였다. 기존의 선행된 예보모델과의 간접적인 비교에서 3% 개선된 71%의 추론 정확도를 보여주었다.

향후 연구에서는 다양한 지형 및 교통량 정보를 활용하여 각 지형의 기상정보를 통해 추론된 초미세먼지 상태를 지식기반 시스템으로 구축하고자 하며, 유사한 지형에 초미세먼지 측정 장치가 없는 장소에서도 근사적으로 추론할 수 있는 방안을 모색할 계획이다.

References

[1] J. Lee, Y. Kim, Y. Kim, “Spatial Panel Analysis for PM2.5 Concentrations in Korea,” Journal of the Korean Data and Information Science Society, Vol. 28, No. 1, pp. 473 - 481, 2017 (in Korean).

[2] G. Choo, G. Lee, M. Jung, “Analysis of Empirical Multiple Linear Regression Models for the Production of PM2.5 Concentrations,” Journal of the Korean Earth Science Society, Vol. 38, No. 4, pp. 283 - 292, Aug. 2017 (in Korean).

[3] G. Ghim, Y. Choi, S. Kim, C. Bae, J. Park H. Shin, “Model Performance Evaluation and Bias Correction Effect Analysis for Forecasting PM2.5 Concentrations,” Journal of Korean Society for Atmospheric Environment, Vol 33, No. 1, pp. 11 - 18, 2017 (in Korean).

[4] J. Park, Y. Choi, W. Jung, “Understanding on Regional Characteristics of Particular Matter in Seoul - Distribution of Concentration in Borough Spatial Area and Relation With the Number of Registered Vehicles,” Journal of Environmental Science International, Vol. 26, No. 1, pp. 55 - 65, 2017 (in Korean).

[5] D.J. Jacob, D.A. Winner, “Effect of Climate Change on air Quality,” Journal of Atmospheric Environment, Vol. 43, No. 1, pp. 51 - 63, 2009.

[6] B. No, G. Choi, “Development of IoT-based PM2.5 Measuring Device,” Journal of the Korean Society of Safety, Vol. 32, No. 1, pp. 21 - 26, 2017 (in Korean).

[7] “Korea Meteorological Administration.” [Online]. Available on http://web.kma.go.kr/eng/index.jsp. [Accessed: 31-Jan-2018].

[8] “Airkorea” [Online]. Available on http://www.airkorea.or.kr/index. [Accessed: 13-Oct-2017].

[9] M. Kurtcephe, H.A. Guvenir, "A Discretization Method Based on Maximizing the Area Under ROC Curve," Journal of Pattern Recognition and Artificial Intelligence, Vol. 27, No. 1, 2013.

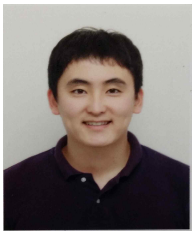
[10] U. Fayyad, K. Irani, “Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning,” 1993.

[11] R. Kerber, “ChiMerge: Discretization of Numeric Attributes,” Proceedings of the Tenth National Conference on Artificial Intelligence, pp. 123 - 128, 1992.

[12] J.W. Grzymala-Busse, “Rule Induction,” in Data Mining and Knowledge Discovery Handbook, Springer, pp. 277 - 294. 2005.

[13] J.W. Grzymala-Busse, W.J. Grzymala-Busse, “Handling Missing Attribute Values,” in Data Mining and Knowledge Discovery Handbook, Springer, 2009, pp. 33 - 51.

Rock-Hyun Choi (최 락 현)



He is received his B.S and M.S degrees in information and communication system engineering from Daegu University, Korea in 2010 and 2012, respectively. He is currently a researcher in the Convergence Research Center for Wellness at DGIST. His research interests include data mining, machine learning, and Wireless Network Control System.

Email: choimosi@dgist.ac.kr

Won-Seok Kang (강 원 석)

He is received his B.S. and M.S. degrees in computer engineering from Yeungnam University, Korea in 1998 and 2000, respectively. He is currently a senior researcher in the Convergence Research Center for Wellness at DGIST. His research interests include distributed simulation algorithms, mobile robot systems, and intelligent embedded systems.

Email: wskang@dgist.ac.kr

Chang-Sik Son (손 창 식)

He is received his M.S. and Ph.D. degrees in intelligent systems from Catholic University of Daegu, Daegu, Korea in 2002 and 2006, respectively. He is currently a senior researcher in the Convergence Research Center for Wellness at DGIST. His research interests include artificial intelligence areas in biomedical informatics and biomedical engineering.

Email: changsikson@dgist.ac.kr