

딥러닝을 이용한 다변량, 비선형, 과분산 모델링의 개선 : 자동차 연료소모량 예측

Improvement of Multivariable, Nonlinear, and Overdispersion Modeling with Deep Learning: A Case Study on Prediction of Vehicle Fuel Consumption Rate

한대석	HAN, Daeseok	정회원 · 한국건설기술연구원 수석연구원 · 교신저자 (E-mail : handaeseok@kict.re.kr)
유인균	YOO, Inkyoon	정회원 · 한국건설기술연구원 연구위원 (E-mail :ikyoo@kict.re.kr)
이수형	LEE, Suhung	정회원 · 한국건설기술연구원 수석연구원 (E-mail : shlee1@kict.re.kr)

ABSTRACT

PURPOSES : This study aims to improve complex modeling of multivariable, nonlinear, and overdispersion data with an artificial neural network that has been a problem in the civil and transport sectors.

METHODS : Deep learning, which is a technique employing artificial neural networks, was applied for developing a large bus fuel consumption model as a case study. Estimation characteristics and accuracy were compared with the results of conventional multiple regression modeling.

RESULTS : The deep learning model remarkably improved estimation accuracy of regression modeling, from R-sq. 18.76% to 72.22%. In addition, it was very flexible in reflecting large variance and complex relationships between dependent and independent variables.

CONCLUSIONS : Deep learning could be a new alternative that solves general problems inherent in conventional statistical methods and it is highly promising in planning and optimizing issues in the civil and transport sectors. Extended applications to other fields, such as pavement management, structure safety, operation of intelligent transport systems, and traffic noise estimation are highly recommended.

Keywords

Data modeling, Artificial neural network, Deep learning, Multiple regression, Fuel consumption

Corresponding Author : Daeseok, HAN, Senior Researcher
Highway&Transportation Research Institute, Korea Institute of
Civil Engineering and Building Technology(KICT), 283,
Goyangdae-ro, Ilsanseo-gu, Goyang-si, Gyeonggi-do, 10223, Korea
Tel : +82.31.910.0066 Fax : +82.31.910.0161
E-mail : handaeseok@kict.re.kr

International Journal of Highway Engineering

<http://www.ksre.or.kr/>

ISSN 1738-7159 (Print)

ISSN 2287-3678 (Online)

Received Feb. 07, 2017 Revised Jul. 20, 2017 Accepted Jul. 25, 2017

1. 서론

본 연구에서는 최근 4차 산업혁명의 핵심기술로 주목 받고 있는 인공지능(Artificial Intelligence; AI)의 토목/교통분야에 대한 활용방안을 모색하고자 하였다. 최근 도로분야에서는 센서와 AI 기술을 접목한 첨단 모니

터링 시스템, 빅데이터 정보분석 시스템, 자율주행 등 기존에 한계에 머무르고 있던 분야에 4차 산업혁명 기술을 도입하기 위한 노력이 진행 중이다. 그러나, 토목·교통분야의 근간이자 대부분의 공학자가 빈번하게 직면하는 문제 중 하나가 수치모델링이다. 변수가 많고 추

세가 분명하지 않은 다변량(multivariables)/비선형(nonlinear)모형, 특히 분산이 심한 데이터에 대한 모델링은 종종 공학자들을 곤란하게 만든다.

다변량/비선형 분석에 가장 보편적인 기법은 회귀분석이다. 회귀분석은 선형성을 가정하기 때문에 추정이 편하고 계산이 쉽다는 장점이 있다. 반면 확보한 데이터가 선형성을 만족하지 못하면 해당 모형은 일반적으로 낮은 적합도(R-square)를 갖게 되며, 모집단이나 샘플을 조정하지 않는 한 개선의 여지는 거의 없다. 데이터가 비선형일 경우 다항식(polynomial)형태로 변환하여 대응이 가능하지만 다변량 분석이 불가능하게 되며, 이 외에도 홀수/짝수 차수에 따른 함수의 방향성 문제, 실측 데이터가 가지고 있는 주요 영역(x 축 domain)을 벗어나게 되면 과소/과대 추정되는 경향이 빈번하게 나타나기도 한다. 특히 데이터 고유의 분산이 큰 경우 하나의 추세선으로 설명하는 회귀분석의 특성상 모형의 설명력은 급격하게 낮아지며, 이 역시도 개선의 여지는 거의 없다. 분석 과정에서 이러한 문제들에 봉착하게 되면 이를 데이터의 특성으로 간주하고 모형의 추정력을 높이기 위한 노력을 쉽게 단념하곤 한다. 그러나 여기서 중요한 것은 모형의 낮은 추정력은 데이터가 아닌 회귀분석 고유의 통계적 특성에 기인한다는 것이다. 즉, 복잡한 특성을 가진 데이터의 패턴분석과 예측 모형의 개발에는 새로운 접근 방식이 필요하다.

이에 본 연구에서는 다층퍼셉트론(multilayer perceptron)을 기반으로 하는 딥러닝(deep learning) 기법을 다변량/비선형/과분산 모델링의 대안으로써 제시·검토하고자 한다. 실증분석에서는 동일한 데이터를 활용하여 인공신경망 모형과 다중회귀모형을 구축하고, 모형의 추정력과 추정 특성을 비교하는 과정을 통해 인공신경망의 활용 가능성을 검토하였다. 분석에 활용된 자료는 본 연구의 선행연구(Han et al. 2016)에서 구축되었던 대형버스 실시간 순간연료소모량 정보이며, 설명변수로는 순간속도, 순간경사도, 포장의 중단평탄성(IRI; International Roughness Index)을 포함하고 있다.

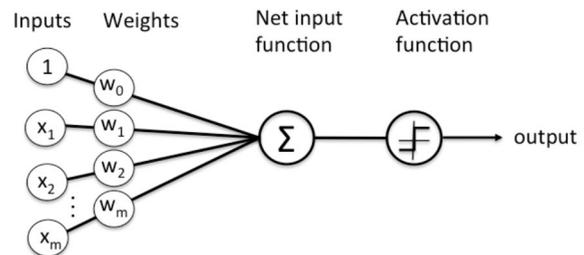
본 연구는 기존연구에서 제안되었던 연료소모량 예측 모형의 추정력을 개선했다는 점에도 의의가 있으나, 인공신경망이 토목/교통분야의 복잡한 예측모델링에 보다 쉽고 광범위하게 응용될 수 있음을 실증사례를 통해 검증한다는 것에 더 큰 의미를 두고 있다.

2. 인공신경망의 개요

인간은 사람의 걸음걸이나 목소리만으로도 누구인지 유추할 수 있다. 한편 그 패턴의 종류가 많아지고 정보의 분산이 커지게 되면 불완전한 근사치를 추정하게 되며, 주어지는 정보량이 많아지고 반복(학습) 횟수가 증가할수록 추정의 정확도가 높아지는 특성을 갖는다. 인공신경망은 이러한 인간의 패턴인식 특징을 수학적으로 모델링 한 일종의 알고리즘이다.

인공신경망의 구조는 입력층(input layer), 출력층(output layer), 은닉층(hidden layer)으로 구성되며, 이 중 은닉층이 2개 이상으로 많아질 경우 심층신경망(deep neural network)으로 구분하는데, 이런 다층 인공신경망을 학습하는 알고리즘을 포괄적으로 딥러닝이라고 한다(Wikipedia, 2017).

인공신경망이 인간의 신경망을 모사하였다고는 하나, 원리 자체는 간단해서 회귀분석을 이해하는 사람이면 누구나 쉽게 이해할 수 있다. 먼저 인공신경망의 핵심인 퍼셉트론(Rosenblatt, 1958)의 이해가 선행되어야 할 필요가 있다(Fig. 1 참조).



Source: Raschka (2015)

Fig. 1 Schematic of Rosenblatt's Perceptron Rule

Fig. 1을 살펴보면 우리가 사용하는 데이터는 입력 값 x_m , 그리고 입력 값을 얼마나 반영할지를 정하는 가중치(weight)는 w_m 으로 표현하고 있다. 두 값을 곱한 값의 총합을 가중치 합(weighted sum, Σ)이라고 하며, 이 값이 역치(threshold)를 넘으면 1, 그렇지 않으면 0이 되는 활성화 과정(Activation function)을 거친다. 즉, 최종적으로 단층구조에서의 퍼셉트론이 하는 일은 1과 0으로 구분하는 것으로 정의될 수 있다(Raschka, 2015).

그러나, 단층 퍼셉트론은 더미(dummy)수준(예: 합격, 불합격)의 간단한 문제는 쉽게 풀 수 있지만 비선형/다변량에 기반하는 주거나 아파트의 가격예측, 강우확률 등 복잡한 문제들은 풀 수 없다는 한계를 갖는다. 특히 과거 'XOR' 과 같은 가장 간단한 비선형 문제도 풀

수 없다는 연구결과(Rumelhart, et al. 1986)가 나오면서 주목받던 신경망 연구는 한 때 암흑기를 맞기도 하였다(Fig. 2 참조). 그러나 2000년대 이후 다층퍼셉트론을 기반으로 하는 딥러닝 기법이 출현하면서 다시 4차 산업혁명의 핵심기술로 주목받고 있다.

Structure	Description of decision regions	Exclusive or Problem	Classes with meshed regions	General region shapes
Single layer	Half plane bounded by hyperplane			
Two layer	Arbitrary (complexity limited by number of hidden units)			
Three layer	Arbitrary (complexity limited by number of hidden units)			

Source: Lilien and Rangaswamy (1999)

Fig. 2 Role of Multi-hidden Layer Structure in a High Dimensional Input Space

다층퍼셉트론의 개념을 다중회귀분석에 대입하여 살펴보기로 한다. 다중회귀분석의 기본 모형식은 $\hat{y} = a + b_1x_1 + b_2x_2, \dots, b_mx_m$ 의 형태로 표현된다. 여기서 x 는 입력치, b 는 기울기 파라미터, a 는 절편, \hat{y} 는 예측치가 되고, x_m 은 입력치, b_m 은 각 노드가 가지고 있는 가중치 w_m 에 해당된다. 회귀분석에서 b 는 설명변수를 대표하는 m 개의 값을 갖지만, 사실 인공신경망에서는 각 층을 구성하는 모든 노드(node)마다 별도의 다중회귀분석이 이루어진다. 그리고 인공신경망에서 종종 이야기 되는 ‘학습’은 각 층의 입력 데이터 중 어떤 값을 얼마나 반영할 것인지를 결정하는 과정이라 할 수 있다(DL4J, 2016).

인공신경망의 학습은 한 계층을 학습하는 퍼셉트론(perceptron) 알고리즘과 계층 간 오류정보를 전달하는 역전파법(backpropagation)으로 구분된다. 퍼셉트론에서는 종속변수의 추정값 \hat{y} 를 계산하고, 역전파법에서는 오차를 줄이기 위한 가중치를 업데이트하는 훈련과정을 거치게 된다(Hastie et al., 2009). 이렇게 결정된 \hat{y} 는 출력되기 전에 비선형 함수를 통과하게 되는데, 이는 함수의 비선형성 조건을 충족함은 물론 노드의 출력값이 너무 과다하지 않도록 제어하는 역할을 한다. 각 노드에서 적용되는 비선형 함수는 S자 형태를 갖는 경우가 대부분인데, 시그모이드 함수(Sigmoid function)가 여기에 해당된다(DL4J, 2016).

즉, 인공신경망과 일반적인 회귀분석의 분명한 차이

는 인공신경망에서 활용하는 가중치의 개수가 회귀분석보다 훨씬 많다는 것이다(RIMF, 2016). 이는 인공신경망이 종속변수와 설명변수들 간의 비선형관계를 한층 더 유연하게 모델링할 수 있는 핵심적인 특징이라고 할 수 있다.

3. 실증분석

본 연구에서는 동일한 데이터를 활용하여 다중회귀모형과 인공신경망 모형을 구축하고, 모형의 추정력과 추정특성을 비교하는 과정을 통해 인공신경망의 활용가능성을 검토하기로 한다. 모형구축에 활용된 데이터는 선행연구(Han et al., 2016)에서 구축한 대형버스의 순간연료소모량 자료로 주행속도, 경사도, 종단평탄성(IRI)을 설명변수로 포함하고 있다. 본 절에서는 수집 데이터와 구축된 모형의 통계적 특성을 중심으로 기술하기로 하며, 데이터 구축에 관한 구체적인 정보는 선행연구(Han et al., 2016)를 참조하길 바란다.

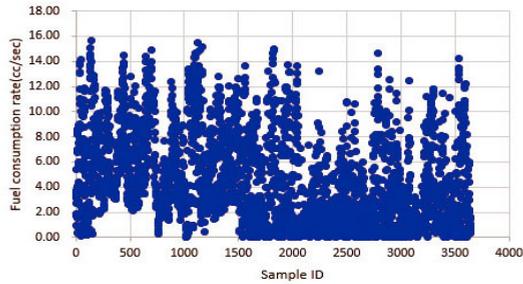
3.1. 데이터의 통계적 특성과 분포

연료소모량은 도로 기하구조 설계, 포장의 유지관리 수준, 교통운영 등 효율적·환경친화적 도로계획을 위한 경제성 분석에 중요한 요소로 활용된다. 순간연료소모량(Instant fuel consumption rate)은 특정조건하에서의 연료소모량을 초 단위로 가공한 지표로, 거리를 기반으로 하는 평균연료소모량보다 설명변수의 영향력을 민감하고 정확하게 반영할 수 있다는 장점이 있다.

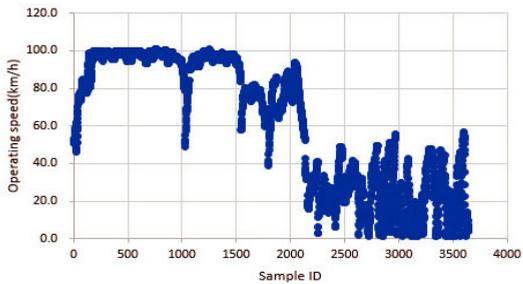
본 연구에서는 ‘순간연료소모량=f(순간주행속도, 순간경사도, 종단평탄성)’ 형식으로 퓨전(fusion)된 데이터를 모형구축에 활용하였다. 순간연료소모량은 차량의 OBD(On-Board Diagnosis)에서 초 단위로 직접 수집되며, 순간주행속도와 순간경사도는 GPS좌표정보를 유클리디안거리법(Euclidean distance)을 통해 가공되었다. 포장의 종단평탄성 IRI는 ASTM(2015)의 측정규격을 준용한 자동노면상태조사장비(PES; Pavement Evaluation Surveyor)를 활용하여 조사되었다. 조사구간은 서울-인천공항 간 62.5km 버스노선으로 단속류와 연속류 통행특성을 모두 포함하고 있으며, 사전분석 결과 주행속도, 경사도, IRI의 범위가 고르게 분포되어 모형 구축에 적합한 조건을 갖추었다. 먼저 수집된 정보의 기초통계치와 분산도를 살펴보기로 한다(Table 1, Fig. 3 참조).

Table 1. Basic Statistics of the Sample Data Set

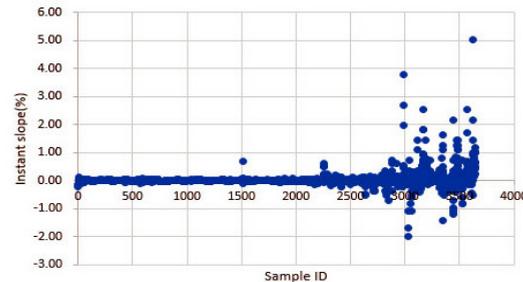
Contents	Num. of sample set	Min	Max	Avg.	Std. dev.	Correlation coefficient with dependent variable
Fuel consumption rate (cc/sec)	3,816	0.0002	15.68	4.38	3.68	-
Operating speed (km/h)		1.01	101.24	61.76	34.21	0.431
Slope (%)		-1.93	5.04	0.02	0.21	-0.055
IRI (m/km)		0.81	6.43	2.61	1.57	-0.313



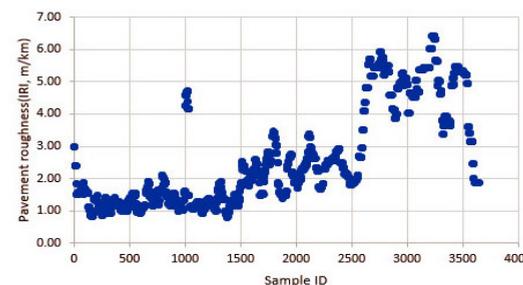
(a) Fuel Consumption Rate



(b) Operating Speed



(c) Instant Slope



(d) Pavement Roughness

Fig. 3 Scatter plots of Variables

Table 1에서 유심히 살펴보아야 할 부분은 연료소모량의 평균과 표준편차이다. 종속변수의 평균은 4.38cc/sec, 표준편차는 3.68cc/sec 수준으로 회귀모형 구축 시 출력 값의 범위가 상당히 제한적이 될 것을 쉽게 예상할 수 있다. 다음으로 Fig. 3의 각 산점도를 살펴보면, 각 설명변수의 산점도 Fig. 3(b)~(d)의 분포특성을 시각적으로만 살펴보아도 종속변수인 Fig. 3(a)와 유사한 추세를 갖는 변수는 찾아보기 어렵다. 종속변수와의 상관계수 R값을 도출해 보면 각각 주행속도 0.431, 경사도 -0.055, 종단평탄성 -0.313 수준에 그치고 있다 (Table 1 참조). 이러한 간단한 통계적 수치만을 검토해보아도 회귀분석모형은 데이터의 특성을 충분히 반영하기 어렵다는 것을 쉽게 알 수 있다.

3.2. 모형의 구축

다중회귀모형은 기존연구(Han et al., 2016)와 동일한 자료를 활용하였기 때문에 원문을 그대로 인용하기로 한다(Table 2 참조).

Table 2. A Large Bus Fuel Consumption Model by Multiple Regression

Regression Eq.: $Y = 0.62 + 0.05 \text{ Vehicle Speed (km/h)} + 0.16 \text{ Slope (\%)} + 0.18 \text{ IRI (m/km)}$, $R^2 = 0.1876$					
Parameter estimates					
Predictor	Coef. Est	Std. Error	t-value	P-value	VIFs
Constant	0.6195	0.3046	2.0341	0.0420	
Vehicle speed (km/h)	0.0531	0.0027	19.8846	0.0000	2.7370
Slope (%)	0.1585	0.2604	0.6087	0.5428	1.0254
IRI (m/km)	0.1818	0.0578	3.1469	0.0017	2.6953

Source: Han et al. (2016)

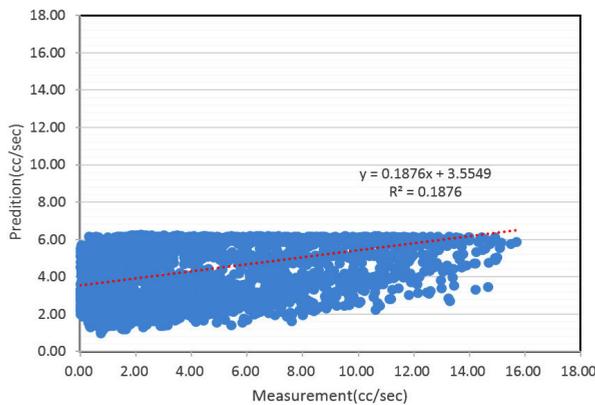
Table 2에서 각 설명변수의 통계치 t-value와 P-value를 살펴보면 주행속도가 연료소모량에 가장 강한 영향을 미치고 있고, IRI도 t-value 3.15(P-value 0.0017)수준으로 유의한 변수로 분류되었다. 다만, 운전자의 주행습관에 많은 영향을 받을 수 있는 경사도는 유의하지 못한 변수로 나타났다. 최종적으로 모형의 적합도는 R-sq.=18.76%로 앞서 예상한 바와 같이 상당히 낮은 수준으로 분석되었다.

본 연구에서 활용한 인공신경망 틀은 Palisade사의 NeuralTools™ 7.0로, 다층퍼셉트론을 기반으로 하는 심층신경망의 일종이다. 특히 심층신경망 중 가장 폭 넓게 활용되고 있는 다층앞먹임네트워크(Multi-Layer Feedforward Networks; MLFN) 알고리즘을 기본으

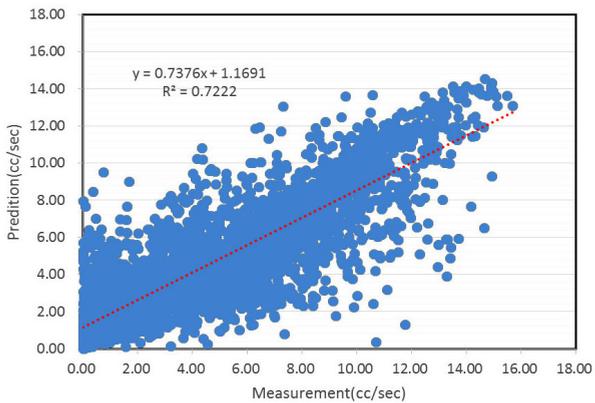
로 한다(Palisade, 2015). MLFN의 신경망구성(network topology)은 데이터의 특성에 따라 1~2개의 은닉층이 활용되며, 활성화함수로는 쌍곡탄젠트함수(hyperbolic tangent function)를 활용한다(Palisade, 2015). 해당 이론에 대한 구체적인 이론적 배경은 Bishop et al. (1995), Master(1995), Reed and Robert(1999)을 참조한다. 여기서 유의할 점은 신경망의 학습과정과 결과는 인간의 두뇌와 유사한 개념으로 통계기법과는 달리 수식으로 표현될 수 없다는 것이다. 즉, 모형의 공유는 학습된 신경망 파일(두뇌)의 공유를 통해서만 가능하다.

3.3. 모형별 예측추세와 추정력 비교

본 절에서는 본 연구의 핵심결과 중 하나인 회귀모형과 인공신경망 모형의 추정력과 추정특성을 비교해 보기로 한다.



(a) Multiple Regression



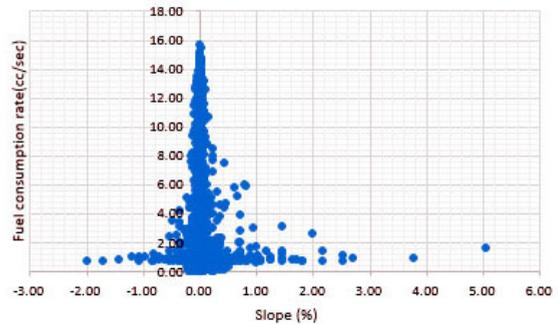
(b) Neural Network

Fig. 4 Scatter Plots of Measurement VS Prediction

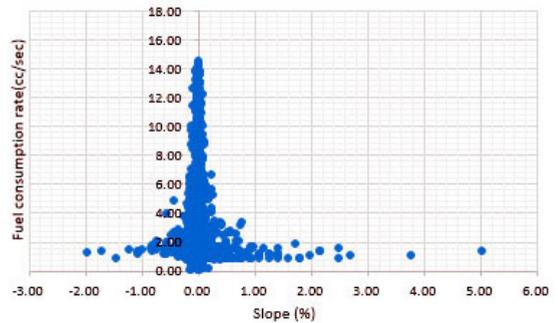
Fig. 4에서 두 그래프의 차이는 명확하다. 회귀모형 적용 시 $R-sq.=18.76\%$ 였던 것이 인공신경망 모형을 통해 $R-sq.=72.22\%$ 까지 개선되었다. 그러나 단순한 통계적 적합도보다 모형의 예측추세에 주목할 필요가 있다. Fig.

4(a)를 살펴보면 회귀모형은 약 $6.2cc/sec \sim 1.0cc/sec$ 의 범위에서만 출력되고 있으며, 특히 전 영역에서 $6.2cc/sec$ 이상의 고출력 영역을 전혀 반영하지 못하는 특성을 가지고 있다. 반면 인공신경망 모형은 출력영역이 실측값과 거의 동일한 수준으로 매우 탄력적이며, 실측값과의 상관계수도 85% 수준으로 분석되었다. 회귀분석의 이러한 특성에 대한 구체적인 이해를 위해 종속변수와 각 설명변수 간의 관계를 그래프를 통해 살펴보기로 한다.

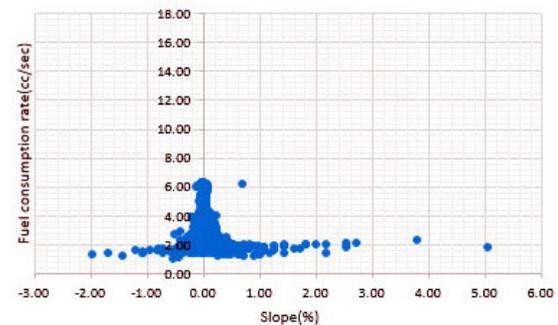
Fig. 5~7은 각 설명변수와 종속변수의 관계를 측정값(a), 인공신경망(b), 회귀모형(c)으로 구분하여 비교하고 있다. 먼저 인공신경망(b)은 모든 설명변수와의 종속변수와의 관계가 실측값(a)과 유사함을 알 수 있다. 반면 회귀모형은 상당한 차이가 있다. 경사도(Fig. 5(c))



(a) Measured

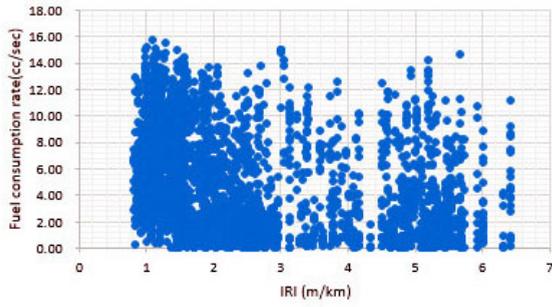


(b) Neural Network

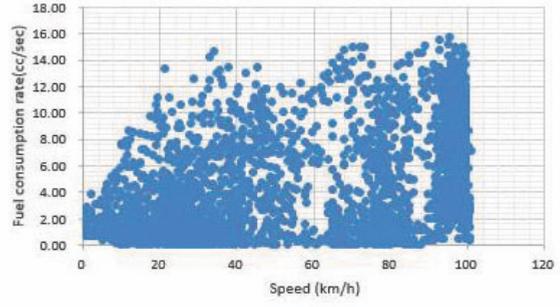


(c) Regression

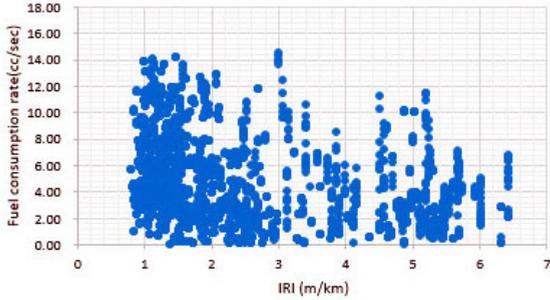
Fig. 5 Relation between Slope and Fuel Consumption Rate



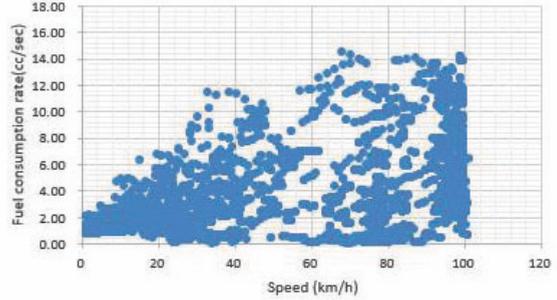
(a) Measured



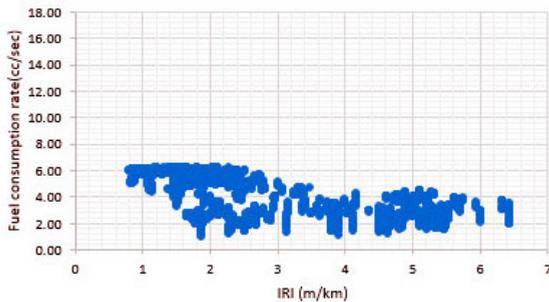
(a) Measured



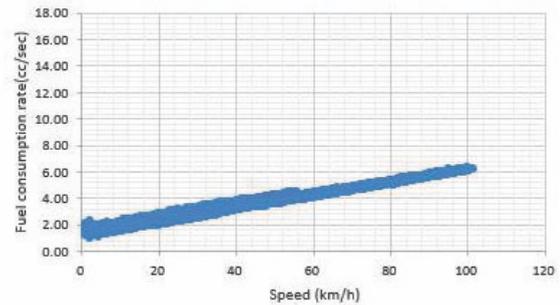
(b) Neural Network



(b) Neural Network



(c) Regression



(c) Regression

Fig. 6 Relation between IRI and Fuel Consumption Rate

Fig. 7 Relation between Speed and Fuel Consumption Rate

와 IRI(Fig. 6(c))의 경우 추세는 다소 반영하고 있다고 판단되나 6.2cc/sec 이상의 고출력 범위를 반영하지 못하고 있다. 특히 Fig. 7(c)를 살펴보면 측정값의 고유 특성과는 상관없이 주행속도와 선형관계를 갖도록 모델링되어 있는 것을 알 수 있다. 즉, 회귀모형은 통계적으로 선형적 상관성이 높은 변수에 가중치를 두고 추세를 결정하는 특성을 반영하고 있다. Fig. 8의 반응과적도(trace-plot)를 통해 모형별 예측추세를 비교해 보면 그 특성을 보다 쉽게 이해할 수 있다.

Fig. 8은 실측값과 각 모형의 예측치를 주행속도를 기준으로 내림차순으로 정리한 반응과적도이다. 그래프를 살펴보면 인공지능경망은 실측치의 분산과 추세를 충실히 반영하고 있으나, 회귀모형은 주행속도를 중심으

로 모형이 구성되면서 추정의 범위가 상당히 제약적임을 알 수 있다.

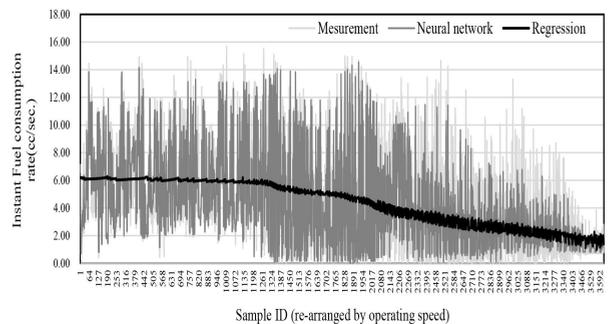


Fig. 8 Comparison of Trace-Plots among the Measurement, Neural Network, and Multiple Regression (Re-arranged by Vehicle Operating Speed)

4. 결론

본 연구에서는 기존 다변량/비선형/과분산 샘플에 대한 모델링의 개선방향으로 인공신경망, 특히 다층 퍼셉트론을 기반으로 하는 딥러닝의 활용가능성에 대해 검토해 보았다.

먼저, 선행연구에서 R-sq.가 18.76%에 불과했던 회귀모형의 설명력은 인공신경망 적용 시 72.22% 수준까지 대폭 개선되었다. 두 모형의 추정특성을 살펴본 결과, 회귀모형은 추정범위가 표준편차 수준으로 상당히 제한적이고 상대적으로 설명력이 강한 변수에 의존하는 특성으로 인해 실측데이터의 분산과 추세를 반영하지 못하는 특성을 보인 반면, 인공신경망은 유연한 학습구조를 통해 실측데이터의 특성을 충실히 반영하고 있음을 확인하였다.

참고로 선행연구에서는 추정력 개선을 위한 방안으로 운전자의 운전행태, 차량하중 등 연료소모량에 영향을 주는 추가적인 변수의 발굴을 제안한 바 있다. 그러나 본 연구를 통해 추가적인 변수 도입 없이 인공신경망 기법의 도입만으로 모형의 추정력이 대폭 개선될 수 있음이 확인되었다. 인공신경망은 종속변수의 변화에 많은 요소가 복합적으로 기인하고 있어 상관관계가 뚜렷한 변수가 존재하지 않는 경우, 그리고 추정력이 절대적으로 낮고 개선의 한계에 도달한 경우에 유용한 도구로 활용될 수 있을 것으로 판단된다.

토목/교통분야에서의 모델링 수요는 광범위하다. 특히 포장파손과정/요인의 해석, 교량/터널/사면 등 대형 구조물의 안전등급 저하, 혼잡한 도심부 구간에서의 통행시간 예측, 도로네트워크 OD통행량 예측, 교통소음 예측 등 알려진 혹은 알려지지 않은 다양한 요소들의 복잡한 영향을 받아 불확실한 결과를 도출하는 요소들이 대부분이다. 본 연구에서는 인공신경망이 기존의 전통적인 통계기법에 의존해왔던 토목/교통분야의 모델링의 한계를 보완할 수 있는 혁신적인 도구로 활용될 것으로 기대하고 있다. 향후 연구에서는 본 연구의 특성상 다루지 않았던 학습데이터의 선별기법, 설명변수의 조합과 선택 등 신경망 학습의 고도화를 위한 연구가 수행되어야 할 필요가 있다.

REFERENCES

ASTM (2015), ASTM E1926-08; Standard Practice for Computing International Roughness Index of Roads from Longitudinal

Profile Measurements, ASTM International, West Conshohocken, PA.

Bishop, C. M., (1995), Neural Networks for Pattern Recognition, Clarendon Press - Oxford.

DL4J(2016), Introduction to Deep Neural Networks, DEEPLARNING4J, A Web Page, (Available in : <https://deeplearning4j.org/neuralnet-overview>).

Han, D., Yoo, I., and Lee, S. (2016), "Development of Heavy Bus Fuel Consumption Model considering Road Pavement Roughness", J. of Korean Society of Hazard Mitigation, Vol.16, No.5, pp.41-46.

Hastie, T., Tibshirani, R. and Friedman, J. (2009), The Elements of Statistical Learning; Data Mining, Inference, and Prediction (2nd edition), Springer Series in Statistics, Springer-Verlag New York.

Lilien, G. L. and Rangaswamy, A. (1999), Marketing Engineering, 1999-20 Predicting Individual Responses using Multinomial Logit Analysis uModeling an Individual's Response to Marketing Effort, Marketing Engineering; Computer-Assisted Marketing analysis and Planning (Available in: <http://slideplayer.com/slide/5270066/>).

Master, T., (1995), Advanced Algorithms for Neural Networks, Wiley.

Palisade (2015), NeuralTools; Neural Network Add-in for Microsoft Excel (ver.7), Palisade Corporation, NY.

Raschka Sebastian (2015), Single-layer Neural Networks and Gradient Descent, A Web Page, (Available in: http://sebastianraschka.com/Articles/2015_singlelayer_neurons.html).

Reed, R. D. and Robert, J. M. (1999), Neural Smithing; Supervised Learning in Feedforward Artificial Neural Networks, MIT Press.

RIMF (R Is My Friend) (2016), Variable importance in neural networks, A web page, (Available in: <https://beckmw.wordpress.com/2013/08/12/variable-importance-in-neural-networks/>).

Rosenblatt, F. (1958), "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", Psychological Review, Vol.65, No.6, pp.386-408.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986), "Learning Internal Representations by Error Propagation", Parallel Distributed Processing, Vol. 1, Cambridge MA: MIT Press, pp.318-362.

Wikipedia (2017), Searching key word: Deep learning, Wikipedia (date: Jan, 2017).