

인공신경망의 은닉층 최적화를 통한 농산물 가격예측 모델

An Agricultural Estimate Price Model of Artificial Neural Network by Optimizing Hidden Layer

저자 (Authors)	배경태, 김창재 Kyung-Tae Bae, Chang-Jae Kim
출처 (Source)	한국정보기술학회논문지 14(12) , 2016.12, 161-169(9 pages) The Journal of Korean Institute of Information Technology 14(12) , 2016.12, 161-169(9 pages)
발행처 (Publisher)	한국정보기술학회 Korean Institute of Information Technology
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07079511
APA Style	배경태, 김창재 (2016). 인공신경망의 은닉층 최적화를 통한 농산물 가격예측 모델. 한국정보기술학회논문지, 14(12), 161-169
이용정보 (Accessed)	경희대학교 163.***.18.29 2020/02/21 16:42 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

인공신경망의 은닉층 최적화를 통한 농산물 가격예측 모델

배경태*, 김창재**

An Agricultural Estimate Price Model of Artificial Neural Network by Optimizing Hidden Layer

Kyung-Tae Bae*, Chang-Jae Kim**

요 약

인공신경망의 성능의 최적화에 영향을 미치는 요인은 은닉층의 수, 학습률, 모멘텀 등이 있다. 이 값들이 실제 농산물 가격예측에 어떤 영향을 미치는 가를 알아보는 것이 본 연구의 목적이다. 이를 검증하기 위해 인공신경망 알고리즘을 통한 당해년도 안동지역의 오이 가격예측을 한다. 독립변수는 기상청의 2013~2015년도 안동지역 21개의 농업기상관측 데이터셋 중에 선형회귀분석을 통해 가장 연관성이 있는 7가지 데이터셋을 설정한다. 종속변수로는 농업진흥청 공공개방데이터인 2013~2015년도의 안동지역 실제 오이의 판매 가격을 설정한다. 또한, 인공신경망의 은닉층 구성은 가장 보편적인 3-tier model을 기본으로 구성하고, 은닉층의 수, 모멘텀과 학습률 계수를 차례대로 늘려나가면서 실제 가격과 예측 가격의 차이를 비교한다.

본 연구는 선형회귀분석으로만 가격예측하기 보다는 선형회귀분석(독립변수 선정)과 인공신경망 기법을 혼합 사용함으로써 예측 정확도를 전자에 비해 14.77% 높였다는 점에 의의가 있다.

Abstract

Factors affecting the performance of the artificial neural network are a number of hidden layer, learning rate, and momentum, etc. These values are the aim of this research to find out whether it affects the actual farm price forecast. In the year through artificial neural network algorithm to verify this prediction and the price of cucumbers, An-dong. Independent variables are the seven sets of data that are most associated with the three linear regression analysis during 2013 ~ 2015 An-dong 21 agricultural meteorological datasets. The dependent variable is set to the actual selling price of cucumbers in Andong Rural Development Administration of public open data 2013-2015 years. In addition, the configuration of the hidden layer neural network consists of the most common 3-tier model as a base, the number of hidden layer, learning rate and momentum factor compares the difference between the actual price and the predicted price.

This research is significant to the point by using the electronic mixing the ANN and linear regression model increasing 14.77% prediction accuracy rather than only using linear regression analysis to predict the price.

Keywords

artificial neural network, hidden layer, price model, momentum, learning rate

* 숭실대학교 SW특성화대학원 석사과정 · Received: Sep. 14, 2016, Revised: Nov. 10, 2016, Accepted: Nov. 13, 2016
** 숭실대학교 SW특성화대학원 교수(교신저자) · Corresponding Author: Chang-Jae Kim
· 접수 일: 2016년 09월 14일 Graduate school of Software, Soongsil University, 369, Sangdo-Ro,
· 수정완료일: 2016년 11월 10일 Dongjak-gu Seoul, 156-743, Korea,
· 게재확정일: 2016년 11월 13일 Tel.: +82-2-828-7019, Email: winchang@ssu.ac.kr

I. 서 론

농산물 중 수급 및 가격 변동이 가장 큰 품목은 채소류인데, 채소류는 노지에서 재배되기 때문에 기상 변화에 따라 생산량의 변화가 심하고, 계절 변화에 따라 주산지가 전국으로 이동하며, 유사 작목 간 생산 대체 관계가 존재하기 때문에 수급 및 가격 예측이 어렵다.

또한, 농산물 가격예측은 농가의 의사결정으로 인한 재배면적의 변화, 기상 변화에 따른 생산량 증감 등 공급부문에 예측할 수 없는 변수가 많아 가격예측이 어려운 실정이다[1].

이러한 수급 예측의 어려움을 보완하기 위하여 농업 기상 변화에 따른 2013-2014년 오이의 도매가격 변화를 측정하여서 2015년 농업관측 기상정보에 따른 2015년 오이의 도매가격을 예측하는 예측 모델을 설계한다.

본 연구는 인공지능경망 알고리즘을 사용하여 오이 가격예측 모델을 제시하였으며, 은닉층(Hidden Layer)의 네가지 구성요소(은닉층의 개수, 은닉층의 노드수, 학습률, 모멘텀) 중에 은닉층의 개수와 모멘텀 값, 학습률이 가격 예측에 어떤 영향을 미치는지를 알아본다. 이때 예측 기법으로는 가장 보편적인 3-Tier인공신경망 기법을 사용하고 SPSS v.21와 WEKA 3.8을 이용하여 공공개방데이터를 분석한다.

본 논문은 다음과 같이 구성한다. 2장에서는 국내·외의 관련연구들을 나열하고, 3장에서는 독립변수를 선정하고 모델을 제시한다. 4장에서는 실험 및 분석을 수행하고 마지막 5장에서 본 논문의 결론을 맺는다.

II. 관련 연구

2.1 다중 선형회귀분석과 인공신경망 분석

다중선형회귀분석은 여러 개의 독립변수(X_1, X_2, \dots, X_n)와 종속변수(Y) 사이에는 선형 관계식(Linear Relationship)이 있음을 가정하고 주어진 학습 데이터를 바탕으로 각 독립변수의 영향력인 회귀 계수(β)를 다음과 같은 식을 이용하여 추정하는 모형이다.

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (1)$$

다중선형회귀분석에서 많은 입력 변수는 노이즈 데이터에 민감하게 반응하거나 불필요한 정보를 포함하여 회귀분석의 예측 성능을 저하시킬 위험이 있다[2].

인간은 뇌의 기본 구조 조직인 뉴런과 뉴런이 연결되어 일을 처리하는 것처럼 수학적 모델로서의 뉴런이 상호 연결되어 네트워크를 형성하는데 이를 신경망이라 한다. 뉴런의 기본적인 기능은 정보의 수용과 연산처리, 정보의 출력이며 다수의 뉴런들이 서로 결합되어 신경망을 이루고 있다. 인공적으로 이루어진 신경망의 이론을 보면 다수의 입력에 대해서 미리 결정된 비선형적인 함수에 의해서 출력이 이루어지는 형태가 된다.

신경망 모델의 모든 뉴런은 보통 기능에 따라 입력층(Input Layer), 은닉층, 출력층(Output Layer)으로 나뉘며 각 층은 기능적으로 연결되어 있다. 입력층은 외부 입력 모드를 연결하고 입력 단위에 따라 은닉층 단위로 보내진다. 여기서 은닉층은 신경망의 내부처리 단위 층으로 신경망의 모드 전환이 주 역할이고 출력층은 산출 모드를 생성하기 위해 사용된다.

신경망의 장점은 처리 노드가 많기 때문에 몇 개의 노드나 연결이 가진 결합이 비교적 시스템 전체의 결합을 초래하지는 않는 결합 내구성(Fault Tolerance)에 있다. 또한 불완전하거나 사전에 알 수 없었던 입력을 표현하는 경우에 신경망은 합리적인 반응을 생성할 수 있는 일반화(Generalization)가 있으며 마지막으로 신경망은 새로운 환경에서 학습한다. 새로운 경우는 즉각적으로 프로그램을 갱신하고 유지하는데 사용하는 적응성(Adaptability)에 있다. 신경망이 가지는 3가지 특징은 첫째 수학적인 모형이 필요 없으며 둘째 자료의 해석보다 예측에 더 유용하다는 것이고 세 번째는 자료의 크기가 커야 한다는 것이다[3].

인공신경망의 성능, 즉, 최적화에 영향을 미치는 요인으로서 은닉층의 수, 은닉층의 노드 수, 학습 반복횟수, 학습률, 모멘텀 등이 있는 것으로 알려져 있다[4]. 한편, 인공신경망을 학습하기 위한 방법으

로는 역전파 알고리즘(Back Propagation Algorithm)이 가장 널리 활용되는데, 이는 목표로 하는 결과값과 모델에 의해 계산된 결과값의 차이를 학습을 통해 계속 조정해 나감으로써 오차값을 최소화하도록 학습시키고 학습결과를 이용하여 예측 대상에 검증을 실시하는 방법이다[5].

2.2 인공신경망 모델에서 은닉층 수가 예측에 미치는 영향

대부분의 인공 신경망 모형은 단일의 입출력층(Input/output Layer)과 소수의 은닉층으로 구성된다. 입력층은 학습에 필요한 입력 요소 노드의 집합이다. 이들 노드들은 은닉층의 각 노드와 완전 연결(Fully Connected) 형태로 연결되나 동일 계층내의 노드들은 상호 연결되지 않는다. 각 층별로 최적의 노드수를 결정하는 명확한 규칙이 없기 때문에 노드의 수는 휴리스틱적인 방법이나 시행착오적인 방법으로 결정된다.

은닉층은 없거나 1개 이상의 계층으로 구성되며, 은닉층의 수도 노드수의 결정과 같이 명확한 결정 규칙이 없다. Tam & Kiang(1992)의 연구에 따르면 은닉층이 없는 경우와 1개가 있는 경우를 비교한 결과, 1개가 있는 인공 신경망이 우수한 결과는 나타났다. 그리고 Collins et al.(1988), Dutta & Shekhar (1988), 그리고 Salchenberger et al.(1992) 등의 연구에 의하면 인공 신경망의 은닉층 수를 2개 이상으로 증가시켜도 단일 은닉층을 갖는 인공 신경망보다 예측률이 나아지지 않는 것으로 나타났다. 따라서 대부분의 선행연구에서 인공 신경망의 성과를 높이고 학습시간을 줄이기 위해 단일 은닉층을 갖는 인공 신경망을 이용하고 있다[6].

은닉층을 포함한 망구조에서는 은닉층의 마디수를 증가시키면 분류능력이 향상되어 더 많은 패턴을 인식할 수 있지만 자칫 과잉 맞춤의 결과를 초래하기 쉽다. 따라서 적절한 은닉마디 수를 결정하는 것이 필요하다.

신경망에서 은닉 노드의 수를 몇 개로 하는 것이 좋더라는 규정은 없다. 다만, 적정 수준의 노드를 이용하여 모형을 만드는 것이므로 노드가 너무 많으면 overfitting의 문제 때문에 설명력이 있지만, 예

측에서는 의미 없는 값이 되는 경우가 많다. 또한 몇 개의 층으로 할 것인가는 한 개의 Layer로 노드의 수와 연결함수, 활성화함수를 잘 조절해야한다. 필요이상의 은닉층은 인공신경망의 수행 속도에 저해를 가져다준다[7].

2.3 인공신경망 모델에서 은닉층의 노드수가 예측에 미치는 영향

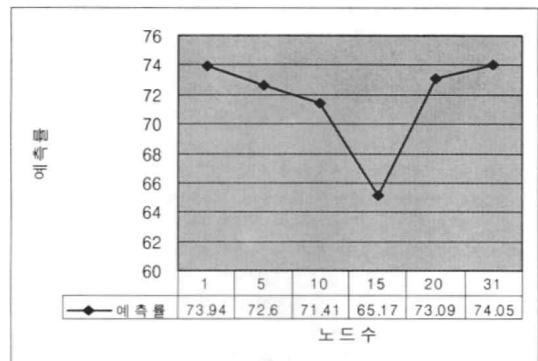


그림 1. 은닉층의 노드수에 따른 예측률 변화

Fig. 1. Variation of prediction by changing node numbers

은닉층의 노드수를 변화시킬 경우 예측률 변화는 그림 1과 같았다. 노드수를 증대시킴에 따라 예측률은 하락하다가 다시 일정지점(15개)이 지나면 상승하는 것으로 나타났다. 특히 은닉층의 노드수가 최대치인 31개일 때 최대 예측정확률인 74.05%를 나타냈으며, 은닉층의 노드수가 15개일 때 최소 예측정확률인 65.17%를 나타냈다. 은닉층의 노드수에 따른 예측률은 V자형 포물선을 따르는 형태가 일반적이다[6].

2.4 인공신경망 모델에서 학습률과 모멘텀이 예측에 미치는 영향

인공신경망 모델의 학습은 일반적으로 상당히 긴 학습시간이 요구되며, 빠르고 올바르게 학습시키기 위해서는 초기 연결강도, 학습률(Learning Rate), 은닉층의 노드수, 모멘텀(Momentum) 등 여러 요인을 고려해야한다.

모멘텀은 신경망의 연결 가중치 조절식에 관성을 줌으로써, 학습시간을 단축시키고 학습 성능을 향상

시킴을 위해 고안된 것이다. 그리고 통상적으로 학습률 값은 0.001 ~ 10사이의 값을 사용한다[8].

Nii O. Attah-Okin의 연구에 따르면 학습률은 예측하는 동안 가중치 값의 크기를 결정한다. 학습률 값을 작게 설정한다면 가중치의 변화가 작게 되고, 크게 설정하면 예측값의 변화가 크게 일어난다. 최적의 값은 확실하지 않으며, 만약 학습률 값을 0으로 설정하면 인공지능망은 학습을 하지 않게 된다[9].

2.5 인공지능망을 통한 기존의 가격예측 모델

기상관련 변수들 및 과거 몇 시점 이전까지의 전기가격, 즉 자기회귀변수들로는 그림 2와 같이 신경망을 통하여 실제 값을 예측한다. 신경망은 시계열 예측에 많이 사용되어 왔는데 이는 그 구조가 시계열 데이터를 훈련시키는데 유연하기 때문이다. 특히, 신경망의 일종인 시간 지연 네트워크에서는 시계열 데이터가 갖는 time-lag를 반영하기도 한다. 신경망 입력층의 노드 수는 주어진 입력변수의 수와 바이어스, 그리고 실험에 의해 결정될 자기회귀차수의 수를 합한 수로 결정한다. 마찬가지로 은닉 노드의 수도 실험에 의해 정해진다. 출력 층의 노드의 수는 한 개로, 신경망 학습이 끝나면 이로부터 전기 판매단가의 예측된 출력 값을 얻는다[10].

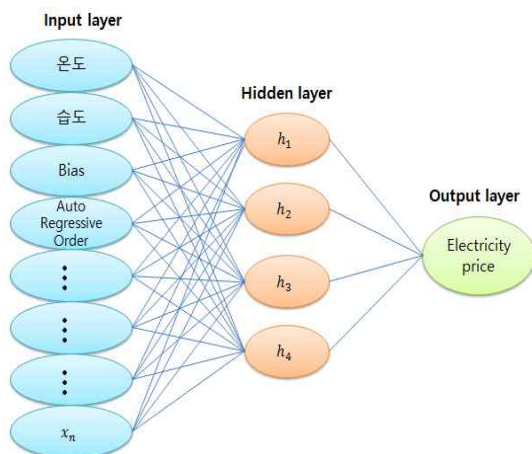


그림 2. 인공지능망을 이용한 날씨별 전기가격 예측 모델
Fig. 2. An example of the ANN network composed of weather related indexes

III. 인공지능망을 이용한 농산물가격 예측모델

3.1 훈련용 데이터셋과 테스트용 데이터셋의 선정

본 연구에서는 훈련용 데이터셋으로 2013-2014년도 안동지역의 21개 농업기상관측 데이터셋과 오이의 가격을 선정하였다. 그리고 테스트용 데이터 셋으로는 2015년도 안동지역의 21개 농업기상관측 데이터셋과 오이의 가격을 선정하였다.

3.2 인공지능망에 입력할 독립변수의 선정

보다 정확한 농산물 가격예측 인공지능망 모델 제시를 하기 위해서 R^2 (결정계수)이 높고 p-value(유의도) 값이 낮은 독립변수를 선정하여야 한다.

이를 위해, SPSS v21을 통한 선형회귀분석을 진행하여 전체 21개 변수 중에 R^2 값이 높고 p-value값이 낮은 독립변수를 선정하여 인공지능망 모델의 실제 인공지능망의 독립변수로 활용하였다.

그림 3은 가장 R^2 값이 높은 독립변수의 선형회귀 그래프이고, 전체 독립변수들의 R^2 값과 p-value는 표 1과 같다.

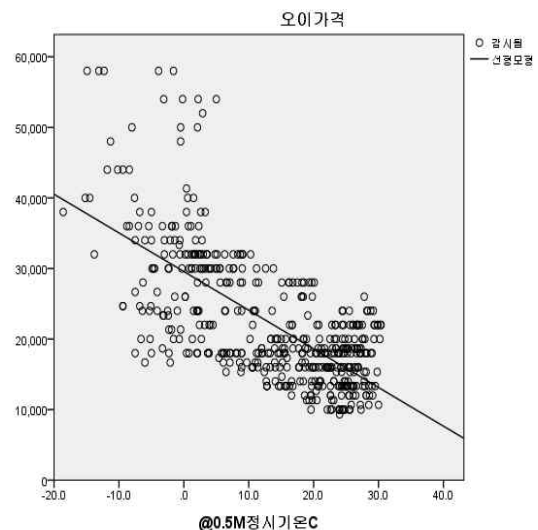


그림 3. 0.5M 정시기온과 오이가격간의 관계
Fig. 3. Relation of 0.5M temperature with cucumber price

표 1. 독립변수 선정을 위한 R^2 과 p-value 정리
Table 1. Independent variables of R^2 and p-value

	R^2	p-value
0.5M정시습도	0.001	0.594
1.5M정시습도	0.000	0.878
4.0M정시습도	0.002	0.344
10CM정시토양수분	0.256	0.000
20CM정시토양수분	0.107	0.000
30CM정시토양수분	0.143	0.000
50CM정시토양수분	0.215	0.000
0.5M정시기온(C)	0.467	0.000
1.5M정시기온(C)	0.459	0.000
4.0M정시기온(C)	0.463	0.000
1.5M정시풍속(ms)	0.007	0.076
4.0M정시풍속(ms)	0.006	0.088
정시지면온도(C)	0.429	0.000
5CM정시지중온도(C)	0.450	0.000
10CM정시지중온도(C)	0.465	0.000
20CM정시지중온도(C)	0.469	0.000
30CM정시지중온도(C)	0.476	0.000
정시누적순복사(MJm ²)	0.095	0.000
정시누적전복사(MJm ²)	0.020	0.002
정시누적반사복사(MJm ²)	0.073	0.000
정시조도(lux)	0.358	0.000

이중에서 0.5M 정시기온, 1.5M 정시기온, 4.0M 정시기온, 5CM정시지중온도, 10CM정시지중온도, 20CM정시지중온도(C), 30CM정시지중온도의 총 7개의 데이터 셋을 인공신경망의 독립변수로 선정하였다.

3.3 농산물 가격 예측 시스템에서 은닉층 노드수 수가 예측에 미치는 영향

은닉층의 노드 수는 독립변수와 종속변수 개수의 합을 n 이라 할 때 $n/2$, n , $3n/2$, $2n$ 의 총 4가지 경우로 구분하여 실험하는 널리 모델이 알려져 있다[10].

따라서 본 실험에서는 n 을 독립변수(7개)와 종속변수(1개)의 합인 8개로 정하였고, 따라서 은닉층의 노드 수는 각각 4개, 8개, 12개, 16개로 설정하여 실험을 진행하였다.

3.4 농산물 가격 예측 시스템에서 은닉층 수가 예측에 미치는 영향

인공신경망 구조에서 ‘은닉층의 단의 수는 얼마로 해야되는가’의 문제는 아직은 확실한 답은 없는 상태이다[4].

하지만 다층 퍼셉트론은 단층 퍼셉트론과 유사한 구조를 가지고 있지만 여러 개의 층으로 구성되어 있으며, 역전파 알고리즘을 통해 단층 퍼셉트론에서 하나의 조정층으로 구성되는 모델들의 한계점들과 선형 분리 문제를 해결하였다[11].

따라서 본 실험에서는 각각의 훈련용 데이터셋과 테스트용 데이터셋에 3-tier 다층 퍼셉트론 방식을 사용하여 가격 예측을 진행하였다.

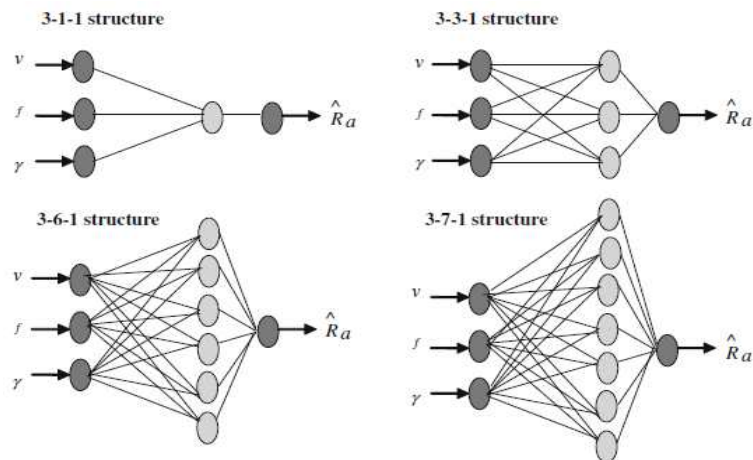


그림 4. 3-depth의 은닉층으로 이루어진 인공신경망 모델

Fig. 4. Artificial neural network of 3 depth model

3-tier 인공신경망에 추천되는 은닉층 수는 입력이 n 일 때 ‘ $n/2$ ’, ‘ $1n$ ’, ‘ $2n$ ’ 그리고 ‘ $2n+1$ ’개 이다. 예를 들어 $n=3$ 일 때 은닉층은 $(3)/2 = 1.5 \approx 1$, $1(3) = 3$, $2(3) = 6$, and $2(3) + 1 = 7$ 가 된다.

따라서 은닉층 depth가 3일 때는 위의 추천방식에 따라 그림 4와 같이 3-1-1, 3-3-1, 3-6-1, 3-7-1 같은 모델이 도출되었다. Chang Jou, Shih-Shien You, Long-Wen Chang의 연구(1994)에 의하면 인공신경망 모델에서 불필요한 은닉층 노드와 depth의 중복을 최소화 시키면, fault tolerance의 가능성도 줄어든다는 것을 증명하였다[12].

위의 연구들을 바탕으로 본 연구에서는 은닉층 depth가 3이고 $n=8$ 일 때 입력층, 은닉층, 출력층을 각각 8-4-1, 8-8-1, 8-16-1, 8-17-1와 같이 설정하였다 <표 2>.

표 2. 은닉층 depth에 따른 제1은닉층, 제2은닉층의 수
Table 2. Number of first and second hidden layers

	입력층	(제1) 은닉층	제2 은닉층	출력층
3-depth neural network model	8	4	-	1
	8	8	-	1
	8	16	-	1
	8	17	-	1

3.5 농산물 가격 예측 시스템에서 모멘텀 계수와 학습률이 예측에 미치는 영향

역전파 알고리즘에서 모멘텀 계수는 어느 정도로 기존방향을 그대로 유지할 것인가에 대한 변수이고 학습률은 새로운 정보에 대한 웨이트를 얼마나 줄 것인가에 대한 변수이다.

본 연구에서는 RMSE 값이 가장 낮을 때의 모멘텀과 학습률 계수를 찾기 위해서 다른 변수들을 동일하게 하고 반복하여 실험을 진행하였다.

IV. 실험 결과

4.1 은닉층의 수에 따른 가격 예측률의 변화

앞에서 설정한 은닉층 수 가설을 바탕으로 실험을 진행한 결과 연관성 계수(Correlation Coefficient)와 RMSE(Root Mean Square Error)는 다음과 같았다. 연관성 계수는 높을수록 독립변수와 종속변수간의 연관성이 높다는 의미이고, RMSE값은 낮을수록 실제값과 예측값이 비슷하다는 의미이다. 따라서 연관성 계수는 높고 RMSE값이 낮은 모델이 최적의 모델이라고 할 수 있다.

표 3. 은닉층 수에 따른 연관성 계수와 RMSE의 값
Table 3. Correlation coefficient and RMSE by hidden layer's number

은닉층 수	연관성 계수 (Correlation coefficient)	RMSE(Root Mean Square Error)
4	0.8422	0.27
8	0.8423	0.264
16	0.8490	0.26
17	0.8489	0.25

표 3에 따르면 은닉층 수가 17일 때 RMSE값이 0.25로 가장 낮게 나왔으므로 이때 최적의 모델이 됨을 알 수 있다.

4.2 모멘텀에 따른 가격 예측률의 변화

모멘텀은 신경망의 연결 가중치 조절식에 관성을 주는 계수이다. 표 4에 따르면 모멘텀 값이 0.5일 때 RMSE값이 0.256로 가장 낮게 나왔으므로 이때 최적의 모델이 됨을 알 수 있다.

표 4. 모멘텀에 따른 RMSE의 값
Table 4. Value of RMSE by momentum coefficients

모멘텀	RMSE(Root Mean Square Error)
0.1	0.2627
0.2	0.2621
0.3	0.2615
0.4	0.2603
0.5	0.2560
0.6	0.2568
0.7	0.2621
0.8	0.2703
0.9	0.2801

4.3 학습률에 따른 가격 예측률의 변화

학습률은 예측하는 동안 가중치 값의 크기를 결정하는 계수이다. 표 5에 따르면 학습률 값이 0.4일 때 RMSE값이 0.2511로 가장 낮게 나왔으므로 이때 최적의 모델이 됨을 알 수 있다.

표 5. 학습률에 따른 연관성 계수와 RMSE의 값
Table 5. Value of RMSE by learning rates

학습률	RMSE(Root Mean Square Error)
0.1	0.2541
0.2	0.2674
0.3	0.2560
0.4	0.2511
0.5	0.2858
0.6	0.2922
0.7	0.2934
0.8	0.2954
0.9	0.3028

4.4 본 논문의 인공지능망 모델과 선형회귀분석 모델과의 비교

표 6에 따르면 본 논문에서 제시한 오이 가격 예측 인공지능망 모델은 선형회귀분석 모델보다 연관성 계수가 0.1477만큼 높고 RMSE 값은 0.0496만큼 낮으므로 더욱 정교한 모델이라고 할 수 있다.

표 6. 인공지능망 모델과 선형회귀분석 모델의 연관성 계수와 RMSE의 값 비교

Table 6. Comparison two model's value of ANN and linear regression by RMSE and correlation coefficient

	본 논문의 인공지능망 모델	선형회귀분석 모델
연관성 계수	0.8466	0.6989
RMSE	0.2511	0.3007

4.5 본 논문에서 제시하고 있는 오이 가격예측 인공지능망 모델

본 연구에서 제시하고자 하는 안동지역의 오이 가격예측 인공지능망 모델은 그림 5와 같다.

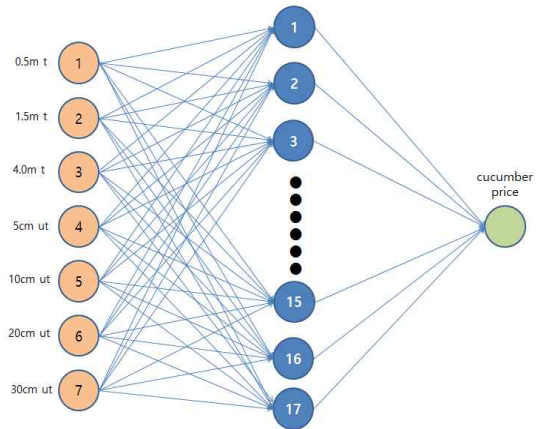


그림 5. 안동지역 오이 가격예측 인공지능망 모델
Fig. 5. Artificial neural network model of predicting cucumber's price

일단 입력층을 보면 21가지 실제 데이터셋 중에 7가지 독립변수(0.5M 정시기온, 1.5M 정시기온, 4.0M 정시기온, 5CM정시기온온도, 10CM정시기온온도, 20CM정시기온온도(C), 30CM정시기온온도)를 선정하여 구성하였음을 알 수 있다. 이는 실제 오이 가격과 연관성이 있는 독립변수를 추려내기 위해 21가지의 농업기상관측 데이터 중 선형회귀분석을 통하여 R^2 값(결정계수)이 높고 p-value 값이 낮은 순으로 나열하여 7가지 독립변수를 선정하였기 때문이다.

두 번째로 은닉층의 모델을 제시하기 위해 Chang Jou, Shih-Shien You, Long-Wen Chang의 연구(1994)를 참고하였다. 위 연구를 참고하여 fault tolerance 값을 줄이기 위해 은닉층의 depth를 3으로 하는 방향으로 모델을 정립하였다. 이후에, trial and error process를 통하여 은닉층의 개수가 17개이고 모멘텀 값이 0.5, 학습률 값이 0.4일 때 가장 실제 가격에 근접하게 예측됨을 확인하였다.

따라서 최종적으로 그림 5와 같은 농업기상관측 데이터를 통한 오이 가격예측 모델을 도출하였다.

V. 결론 및 향후 과제

본 연구에서 제시한 인공지능망을 통한 오이의 가격예측 모델에서는 은닉층 수가 17개, 모멘텀 값이 0.5, 학습률 값이 0.4일 때 실제 값과 근접한 예측 값이 나옴을 확인할 수 있었다.

또한, 연구에서 제시한 오이 가격 예측 모델의 RMSE값, 연관성계수의 비교를 통해 선형회귀분석 모델에 비해 인공지능망 모델이 농산물 가격 예측에 있어서 보다 탁월했음을 알 수 있다.

하지만 은닉층의 수가 너무 적거나 많을 경우에는 RMSE값이 높아짐에 따라 예측 값이 실제 값에 근접하지 않게 됨을 알 수 있었다. 모멘텀과 학습률 계수 또한 적절한 값을 찾지 못하고 너무 높거나 너무 낮은 경우에는 실제 값과의 오차가 많이 생겼음을 알 수 있다. 때문에 적절한 은닉층 수와 모멘텀 계수, 학습률 계수의 설정이 인공지능망 모델 설계에 중요한 부분이라 할 수 있다.

본 연구에서는 농업기상 데이터셋을 기반으로 오이의 가격을 예측하였으므로 타 농산물의 가격 예측으로도 확대할 수 있을 것으로 기대한다. 그러나 농산물 가격예측에 영향을 주는 독립변수가 농업기상관측 데이터 이외에 가격의 변화, 유통망 구조, 발의 면적 변화 등 다양하게 존재 하는데 이와 같은 부분을 고려하지 못한 점은 본 연구의 한계라고 할 수 있다. 공산품이나 타 사업 분야의 가격예측은 향후에 연구할 과제로 남겨둔다.

References

- [1] Kuk-Hyun Nam and Young-Chan Choe, "A Study on Onion Wholesale Price Forecasting Model", Journal of Agricultural Extension & Community Development, Vol. 22, No. 4, pp. 423-434, Nov. 2015.
- [2] Park Hee Jin, Jang Kyoung Ye, Lee Youn Ho, Kim Woo Je, and Kang Pil Sung, "Prediction of Correct Answer Rate and Identification of Significant Factors for CSAT English Test Based on Data Mining Techniques", KIPS Tr. Software and Data Eng, Vol. 4, No. 11, pp. 509-520, Aug. 2015.
- [3] Kwang-Kyu Seo, "Development of a Sales Prediction Model of Electronic Appliances using Artificial Neural Networks", Journal of Digital Convergence, Vol. 12, No. 11, pp. 209-214, Nov. 2014.
- [4] Yoonho Cho and Inhwan Kim, "Predicting the Performance of Recommender Systems through Social Network Analysis and Artificial Neural Network", Journal of Intelligent Information Systems, Vol. 16, No. 4, pp. 159-172, Dec. 2010.
- [5] Eun Chung Ryoo, Hyunchul Ahn, and Jae Kyeong Kim, "The Audience Behavior-based Emotion Prediction Model for Personalized Service", Korea Intelligent Information Systems Society, Vol. 19, No. 2, pp. 73-85, Jul. 2013.
- [6] Jin Baek Kim and Yoo Il Kim, "The Influence of Weight Adjusting Method and the Number of Hidden Layer's Node on Neural Network's Performance", Journal of Korean Association of Information Systems, Vol. 9, No. 1, pp. 27-44, Jul. 2000.
- [7] Il-Dong Jang and Seung-Min Wee, "The Analysis Telecommunication Service Market with Data Mining", Korea Information Science Society, Vol. 28, No. 2, pp. 1-3, Oct. 2001.
- [8] Han-Guk Hong, "A Study of Apartment Price Model Using Artificial Neural Networks", The International Industrial Information Systems Conference, Vol. 15, No. 32, pp. 220-226, May 2009.
- [9] Nii O. Atttoh-Okine, "Analysis of learning rate and momentum term in backpropagation neural network algorithm trained to predict pavement performance", Advances in Engineering Software, Vol. 30, No. 4, pp. 294-295, Jan. 1999.
- [10] Hang Seok Kim and Hyun Jung Shin, "Electricity Price Prediction Based on Semi-Supervised Learning and Neural Network Algorithms", Journal of the Korean Institute of Industrial Engineers, Vol. 39, No. 1, pp. 30-45, Feb. 2013.
- [11] Turker Ince, Serkan Kiranyaz, Jenni Pulkkinen, and Moncef Gabbouj, "Evaluation of global and local training techniques over feed-forward neural network architecture spaces for computer-aided medical diagnosis", Expert Systems with

Applications, Vol. 37, No. 12, pp. 8451-8452, Jun. 2010.

- [12] Azlan Mohd Zain, Habibollah Haron, and Safian Sharif, "Prediction of surface roughness in the end milling machining using Artificial Neural Network", Expert Systems with Applications, Vol. 37, No. 2, pp. 1755-1768, Mar. 2013.

저자소개

배 경 태 (Kyung-Tae Bae)



2010년 2월 : 숭실대학교
정보통신전자공학부(학사)
2015년 3월 ~ 현재 : 숭실대학교
SW특성화대학원(석사)
관심분야 : 인공지능망, 빅데이터,
정보보안 등

김 창 재 (Chang-Jae Kim)



2005년 2월 : 숭실대학교
정보과학대학원(석사)
2009년 8월 : 숭실대학교
컴퓨터학부(박사)
2009년 8월 ~ 현재 : 숭실대학교
SW특성화대학원 교수
관심분야 : SW공학, SW아키텍처,
데이터베이스, 빅데이터 등