

LSTM (Long-short Term Memory)과 GRU (Gated Recurrent Units) 모델을 활용한 양식산 넙치 도매가격 예측 연구

이가현 · 김도훈*

부경대학교 해양수산경영학과

Forecasting the Wholesale Price of Farmed Olive Flounder *Paralichthys olivaceus* Using LSTM and GRU Models

Ga-hyun Lee and Do-Hoon Kim*

Department of Marine Fisheries Business and Economics, Pukyong National University, Busan 48513, Republic of Korea

Fluctuations in the price of aquaculture products have recently intensified. In particular, wholesale price fluctuations are adversely affecting consumers. Therefore, there is an emerging need for a study on forecasting the wholesale price of aquaculture products. The present study forecasted the wholesale price of olive flounder *Paralichthys olivaceus*, a representative farmed fish species in Korea, by constructing multivariate long-short term memory (LSTM) and gated recurrent unit (GRU) models. These deep learning models have recently been proven to be effective for forecasting in various fields. A total of 191 monthly data obtained for 17 variables were used to train and test the models. The results showed that the mean average percent error of LSTM and GRU models were 2.19% and 2.68%, respectively.

Keywords: Deep learning, Price forecasting, Olive flounder, *Paralichthys olivaceus*, Wholesale price

서 론

우리나라 수산업에서 양식어업이 차지하는 산업적 중요도는 점점 증가하고 있다. Fig. 1에서 보는 바와 같이, 해면양식업 생산량은 2006년을 기점으로 연근해어업 생산량을 뛰어넘어 지속적인 상승세를 보이다 2021년에는 연근해어업 생산량의 약 2.6배를 생산했다. 당해 해면양식업 생산량은 약 240만톤으로, 전체 수산업 생산량의 약 63% 수준이다(KOSIS, 2022). 2021년 기준 양식어업 전체 생산량 중 가장 큰 비중을 차지하는 것은 해조류이며(약 77%), 어류는 약 3.7%에 불과하다. 하지만 생산금액을 기준으로 어류가 약 37%에 해당하는 1조 2천억원으로 가장 큰 비중을 차지하고 있어 양식어업에서 어류가 차지하는 산업적 중요도는 상당히 높다고 할 수 있다. 특히 넙치는 양식어류 생산에서 가장 큰 비중을 차지하는 어종으로, 2021년 기준 전체 양식어류 생산량의 약 47% 그리고 생산금액으로는 약 55%를 차지하고 있다(KOSIS, 2022). 최근 넙치의 가격 변동 문제가 대두되고 있다. 인천 지역 도매가격을 기준으로, 2010년

부터 2019년까지의 연간 월별 가격 변동율의 절댓값의 평균은 약 13%에 수준이었으나, 2020년부터 2022년까지는 약 21%로 수준으로 상승하였다. 특히, 2019년 5월부터 2021년 4월까지의 가격 변동율은 무려 약 38% 수준으로, 이전에 비해 매우 높은 현상이 발생하고 있다(KMI, 2022). 가격은 소비 결정에 중요한 요소로 작용하기 때문에 물가의 변동은 소비자가 민감하게 받아들일 수 있는 현상이다(Lee and Kim, 2010). 특히 수산물과 같은 주요 소비 품목의 가격 폭등 및 폭락은 소비자의 부담을 가중시킨다. 수산물 가격 변동은 소비자뿐만 아니라 수산업계 종사자들에게 불확실성으로 인한 위험성을 증대시켜 소득을 불안정하게 만들기도 한다(Ko, 2009). 이에 따라 수산물 가격을 보다 정확히 예측 및 분석하는 것은 수산업 경영의 안정성과 이를 통한 수산업의 산업적 발전을 도모하는 데 필수적이라 할 수 있다. 본 연구에서는 현재 다양한 분야에서 가격 예측 기법으로 주목받고 있는 딥러닝 모델을 수산물 가격 예측에 적용하고자 한다. 예측을 위해 분석 대상 변수 자료를 수집하고, 가격 예측에 영향을 줄 것으로 예상되는 독립변수들을 선정하여 다변

*Corresponding author: Tel: +82. 51. 629. 5954 Fax: +82. 51. 629. 5953

E-mail address: delaware310@pknu.ac.kr



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

<https://doi.org/10.5657/KFAS.2023.0243>

Korean J Fish Aquat Sci 56(2), 243-252, April 2023

Received 10 January 2023; Revised 19 February 2023; Accepted 21 February 2023

저자 직위: 이가현(대학원생), 김도훈(교수)

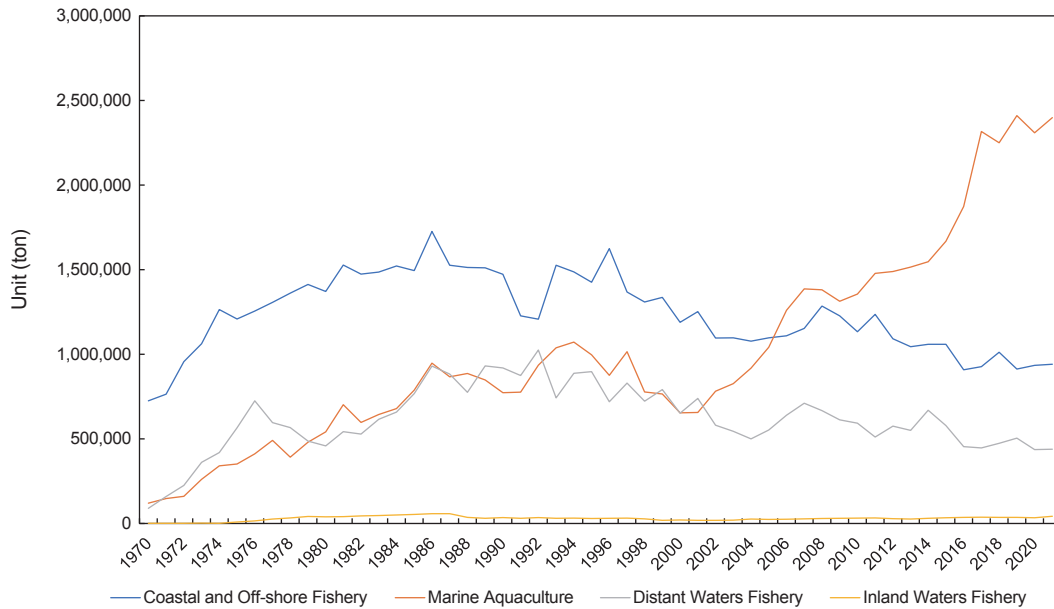


Fig 1. Annual statistics of production be fishery type.

량 모델을 구축하였다. 서론에서는 관련 선행연구를 검토하고, 재료 및 방법에는 분석에 활용한 변수의 선정 근거와 특성, 그리고 모델 구축 과정을 기술하였다. 예측 결과에 대한 내용은 결과에 기술하고, 고찰에서는 연구의 요약 및 결론으로 본 연구의 시사점 및 한계점 등을 제시하였다.

딥러닝 모델은 비지도 학습, 우수한 일반화 능력 및 빅데이터 훈련 등의 특성 때문에 최근 다양한 분야에서 주목받고 있다(Wang et al., 2019). 특히 딥러닝 모델을 활용하여 주가를 예측하는 국내외의 연구가 다수 존재하며, 최근 국내에서는 농산물 가격 예측을 위해 딥러닝 모델을 활용하고 있다. Shin et al. (2018)의 연구에서는 딥러닝 알고리즘을 이용하여 자연재해가 농산물 가격에 미치는 영향을 고려한 학습 모델을 만들고 농산물 가격을 예측하였다. Bae and Kim (2016)은 인공지능망의 은닉층 최적화를 통해 농산물 가격에 최적화된 모델 구조를 도출하는 연구를 수행하였다. Yun et al. (2016)는 국제곡물가격 예측을 위해 인공지능망을 이용하였으며, Im et al. (2018)는 딥러닝 기술을 활용하여 과일 가격을 예측하였다.

딥러닝 모델을 활용하여 수산물 가격을 예측한 선행연구로 Hwang et al. (2012)은 인공지능망 기법을 활용하여 고등어 위판가격을 예측하고, 위판량 및 위판가격이 위판단가에 미치는 영향을 분석하였다. Song (2021)은 MLFN (multi-layer feedforward network)과 GRNN (generalized regression neural network) 모델 그리고 다중회귀모형을 사용하여 2019년과 2020년의 양식산 넙치의 산지가격을 예측하였다. 모델 간 비교 결과, GRNN 모델의 평균절대백분율오차(mean absolute percentage error, MAPE)가 2019년에는 5.26%, 2020년에는 6.04%로, 가장 우수한 것으로 평가되었다. Woo and Shin

(2022)은 LSTM (long short term memory)과 MLP (multi-layer perceptron) 모델, 그리고 ARMA (autoregressive moving average) 모형을 활용하여 양식산 조피볼락의 도매가격을 예측하였다. 모델 간 비교 결과, MLP 모델의 MAPE가 17%로 가장 낮게 나타났다. Lee and Kim (2022)는 다변량으로 LSTM과 GRU (gated recurrent unit)를 구축하여 양식산 조피볼락의 산지가격을 예측하였고, 단변량 모델과의 예측력을 비교하였다. 그 결과, LSTM 모델의 예측력이 더 우수한 것으로 나타났으며, 단변량 LSTM의 MAPE는 13.05%, 다변량 LSTM의 MAPE는 4.66%로 나타나 수산물 가격 예측 시 다변량 딥러닝 모델 활용의 필요성이 제시되었다.

이 외의 수산물 가격 예측에 관한 연구는 대부분 전통적 시계열 모형을 활용하였다. Guttormsen (1999)은 CAD (classical additive decomposition), HW (holt winters) 지수평활법, ARMA, VAR (vector auto regression) 모형 및 두 가지 Naive 모델로 연어의 주간 산지가격을 예측하였다. MPE (mean percentage error), MAPE, RAF (ratio of accurate forecasts)를 이용하여 예측 정확도를 평가한 결과, MPE와 RAF를 이용해 평가한 예측력은 CAD 모델이, MAPE를 이용해 평가한 예측력은 VAR 모형이 가장 높은 것으로 나타났다. Nam et al. (2014)는 ARIMA (autoregressive integrated moving average) 모형과 VECM (vector error correction model) 모형, 그리고 다중회귀분석으로 김 위판가격의 단기 예측력을 검증하였다. 세 모형의 예측력을 비교한 결과, RMSE (root mean squared error), MAE (mean absolute error), MAPE, Theil IC (theil inequality coefficient) 모두 다중회귀모형이 가장 낮은 것으로 평가되었다. Gordon (2017)은 ARMAX (autoregressive moving aver-

age with exogenous variable) 모델과 EC (error-correction) 모델로 캐나다 수산물의 산지가격에 영향을 미치는 요소를 분석하고, 시뮬레이션을 통해 산지가격을 예측하였다. ARMAX 모델의 외생 변수로 미국과 캐나다의 환율을 설정하여 분석한 결과, 미국과 캐나다의 환율 하락은 수산물 산지가격 상승으로, 미국과 캐나다의 환율 상승은 산지가격 하락으로 이어지는 것으로 나타났다. EC 모델을 사용하여 생산자 가격과 공정 가격 (process price) 간의 연관성을 분석한 결과, 공정 가격과 산지가격은 장기간에 걸친 균형 관계가 존재하며, 공정 가격이 산지가격 결정을 선도하는 것으로 나타났다. 이처럼 수산물 가격에 관련한 기존 연구들은 대부분 전통 시계열 모형을 활용한 것이며, 수산 분야에서 딥러닝 모델을 활용하여 가격을 예측한 연구는 아직 미흡한 실정이다.

재료 및 방법

변수 선정

본 연구에서는 방대한 양의 데이터 활용이 이점인 딥러닝 모

델의 특성에 따라 투입되는 모든 변수를 고려하여 수집 가능한 자료의 최대 기간인 2006년 10월부터 2022년 8월까지의 월별 자료를 사용하였다(Najafabadi et al., 2015).

분석에서 사용된 변수는 양식산 넙치의 지역별 도매가격(인천, 하남, 부산), 지역별 산지가격(제주, 완도), 출하량 및 양성 물량, 수출량 및 수입량, 경쟁어종인 양식산 조피볼락의 도매가격, 산지가격, 출하량, 양성 물량, 연어 및 방어 수입량, 양식 환경 변수로 제주지역 수온, 그리고 특수한 시장적 변수인 COVID-19 더미 변수 등 총 17가지이며, 이들 변수들에 대한 기초 통계량 및 자료 출처는 Table 1에서 정리된 바와 같다.

종속변수는 인천 지역의 양식산 넙치 900 g-1.0 kg 품목 도매 가격으로 사용하였다. 양식산 넙치의 크기별 품목 중 생산 비중이 가장 높은 것은 1 kg 품목이다(KMI, 2022). 그리고 우리나라에서 양식산 넙치의 주요 산지는 제주와 완도로, 2021년 기준 제주는 넙치 양식생산량의 50% 이상을 차지하고 있다(KOSIS, 2022). 인천, 하남, 그리고 부산의 도매시장 중 제주산 넙치의 유통 비중이 가장 큰 곳은 인천이며(Lee and Ma, 2020; MOF, 2021), 특히 최근 인천 지역 양식산 넙치 도매가격의 변동성이

Table 1. Basic statistics of analysis variables

	Incheon wholesale price (won) ¹	Hanam wholesale price (won) ¹	Busan wholesale price (won) ¹	Jeju producer price (won) ¹
Average	13,453	13,875	14,265	11,355
Maximum	18,575	19,625	19,625	16,358
Minimum	9,750	9,750	9,250	7,526
Median	13,375	13,625	14,000	11,326
	Wando producer price (won) ¹	Shipment volume (ton) ¹	Growing volume (ton) ¹	Export volume (ton) ²
Average	12,614	1,128	1,759	260
Maximum	17,887	2,094	3,082	562
Minimum	8,274	499	746	82
Median	12,452	1,120	1,618	240
	Import volume (ton) ²	Wholesale price of Korean rockfish (won) ¹	Producer price of Korean rockfish (won) ¹	Shipment volume of Korean rockfish (ton) ¹
Average	18	11,619	8,971	1,932
Maximum	187	20,906	17,205	4,055
Minimum	0	7,229	5,638	663
Median	0	11,264	8,550	1,842
	Growing volume of Korean rockfish (ton) ¹	Import volume of salmon (ton) ²	Import volume of yellow tail (ton) ²	Water temperature of Jeju (°C) ³
Average	21,790	856	74	19
Maximum	44,985	2,896	1,164	27
Minimum	10,371	0	0	14
Median	18,227	483	7	19
COVID-19 Dummy Variable				
0	COVID-19 (O)			
1	COVID-19 (X)			

¹KMI (2022), Fisheries outlook statistics. ²KSC (2022), Trade statistics. ³KMA (2022).

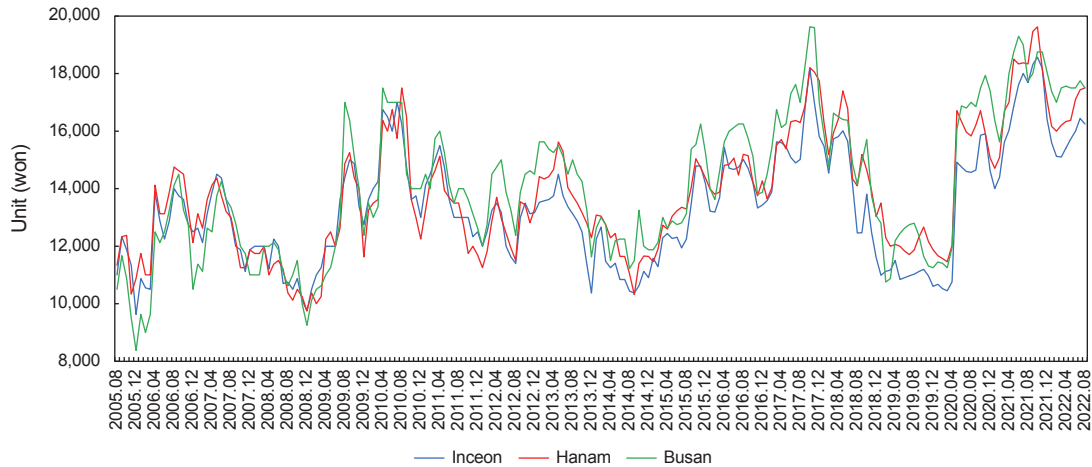


Fig 3. Comparison of wholesale price changes of olive flounder *Paralichthys olivaceus* by region (900 g–1 kg).

수산개발원 수산업관측센터의 관측통계 월별 자료를 활용하였으며, 전체 크기 및 지역의 평균치로 이용하였다.

지형이나 해류의 특성으로 인한 수온의 변화는 양식 넙치의 식욕부진, 성장불량, 대량폐사 등을 유발한다(Lee et al., 2002). 이처럼 수온의 변화는 넙치 생산에 직접적으로 연관되며, 가격에도 역시 영향을 미칠 것으로 판단된다. 이에 양식산 넙치 생산량 비중이 가장 높고 인천에서 가장 큰 비중으로 유통되는 제주 지역의 수온을 독립변수로 활용하였다. 기상청 기상자료개발포털 기상관측 일별 해양자료 중 제주 지역 해양기상부이 2곳, 등표기상관측 1곳, 파고부이 12곳의 일별 수온을 월별 평균 자료로 변환하여 사용하였다.

수입량과 수출량은 식품의 수급 중 공급 및 수요 요소에 해당한다(KREI, 2022). 수급과 가격은 상호 연관성을 갖기 때문에 (Wohlgenant, 1985), 넙치의 수입량 및 수출량은 도매가격 예측에 필수적인 변수라 할 수 있다. 이에 본 연구에서는 관세청 수출입무역통계의 넙치 수입 및 수출 통계량 자료를 활용하였다. 우리나라에서 넙치는 활어회 선호 1위 어종이다(Kim and Kang, 2019). 한편, 연어와 방어는 최근 횡감 대체 어종으로 수입량이 크게 늘고 있으며, 특히 연어는 선어회 선호 1위 어종이다(Kim and Kang, 2019; Song, 2021). 이에 조피볼락과 더불어 연어와 방어가 넙치의 경쟁 어종으로서 가격에 영향을 미칠 것으로 예상된다. 이에 본 연구에서는 관세청 수출입무역통계

의 신선 냉장 연어 및 활어 방어 수입 통계량 자료를 활용하였다. Fig. 2를 살펴보면, COVID-19 확산 시점인 2020년 이후로 넙치의 도매가격이 높게 형성되어 팬데믹이 수산물 가격에도 영향을 미치는 것으로 판단된다. Song (2021)에서도 이러한 상황을 고려하여 COVID-19 발생 시점 전후로 시계열을 구분하여 예측을 수행하였고, Lee and Kim (2022)에서는 조피볼락 산지가격 예측을 위해 COVID-19 관련 더미 변수를 활용하였다. 이에 본 연구에서도 COVID-19를 더미 변수화하여 분석에 활용하였다.

분석 방법

분석 방법에 대한 설명은 Lee and Kim (2022)을 참고하여 작성하였다.

딥러닝 모델 (Deep Learning Model)

인공신경망은 인간의 뇌를 모방한 머신러닝 기법이다. 정보를 처리하는 노드가 계층적으로 구성되어 네트워크를 형성한다. 인공신경망의 기본 모델은 독립변수에 해당하는 입력층과 종속변수에 해당하는 출력층으로 구성되어 있다. 모델의 작동 원리는 입력층에 가중치를 부여하고 이를 합산하여 최종 출력으로 변환하는 것이다(Cooper, 1999).

인공신경망은 계층 구조에 따라 다양한 모델로 세분화된다. 초기 모델인 단층퍼셉트론(single-layer perceptron, SLP)부터 딥러닝 기술이 적용된 다층퍼셉트론(multi-layer perceptron, MLP), 합성곱신경망(convolutional neural network, CNN), 순환신경망(recurrent neural network, RNN), 그리고 생성적 적대 신경망(generative adversarial network, GAN) 등이 있다. 본 연구에서는 시계열 예측에 특화된 RNN의 일종인 LSTM 과 GRU로 예측 모델을 사용하였다.

RNN은 입력층-은닉층-출력층으로 이루어진 각 시퀀스의 출력값의 일부가 입력값에 수평적으로 포함되어 뒤로 갈수록 가

Table 3. Correlation of IWP, HWP and BWP

	IWP	HWP	BWP
IWP	1		
HWP	0.940	1	
BWP	0.922	0.926	1

IWP, Incheon wholesale price; HWP, Hanam wholesale price; BWP, Busan wholesale price.

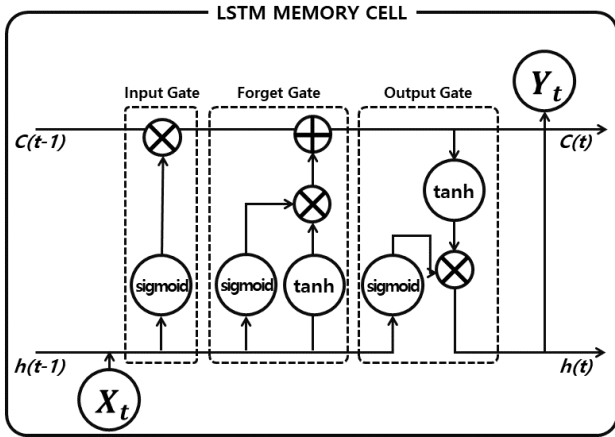


Fig 4. Structure of LSTM. LSTM, Long-short term memory.

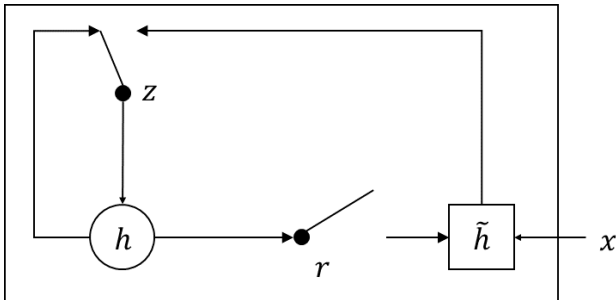


Fig 5. Structure of GRU. GRU, Gated recurrent units.

중치가 업데이트되는 모델이다(Mun et al., 2016). 하지만 시퀀스가 길어질수록 이전 데이터와의 연관성이 떨어져 예측력이 저하되는 장기 의존성 문제가 나타나기도 하는데, 이러한 문제점을 해결한 모델이 LSTM이다. LSTM은 1997년 Hochreiter와 Schmidhuber에 의해 제안된 모델로, 기존의 RNN 은닉층을 memory cell이라 불리는 은닉층 유닛으로 대체하여 RNN의 한계를 보완하였다(Hochreiter and Schmidhuber, 1997).

LSTM의 memory cell은 Fig. 4에서 보는 바와 같이, input gate, output gate, 그리고 forget gate로 불리는 세 게이트와 sigmoid layer와 tanh layer로 구성되어 있다(Hochreiter and Schmidhuber, 1997). Input gate로 입력된 정보를 forget gate를 거치며 출력 정보 망각 여부 및 시기가 결정된다. 활성화 결과가 높은 정보가 셀 상태에 저장되어 가중치 이외의 추가적인 정보를 반영하도록 하는 것이다(Fu et al., 2016; Shahid et al., 2020). Sigmoid layer는 0-1의 값을 생성하여 정보가 어느 정도로 저장될지 결정한다. Tanh layer는 새로운 벡터를 생성하고, 이는 memory cell에 추가된다. 이러한 과정을 거친 정보는 output gate로 출력되고, 다음 memory cell에 전달되는 과정이 반복된다(Selvin et al., 2017). 각 셀은 식 (1)-(5)과 같은 수학적 방정식을 사용해 나타낼 수 있다.

식 (1)-(5)에서 W와 b는 각 게이트와 memory cell에 적용되는 가중치와 편향을, X_t 는 투입 시계열 변수를, h_{t-1} 는 각 게이트마다 가중치가 부여되는 이전 단계의 은닉층 값을 의미한다. Input gate 및 forget gate에서는 식 (1)-(2)과 같이 X_t 와 h_{t-1} 에 가중치 및 편향을 적용하고 sigmoid 활성화 함수를 통해 비선형성을 부여한다. Input gate와 forget gate의 연산 결과는 식 (3)의 과정을 거쳐 memory cell인 C_t 로 변환되고, tanh 활성화 함수가 적용되어 식 (4)에 해당하는 output gate 연산 결과와 곱해져 현시점의 은닉층 값 h_t 로 반환된다. 이와 같은 과정이 반복되어 최종적으로 Y_t 로 산출된다(Fu et al., 2016; Shahid et al., 2020).

$$i_{\text{input gate}} = \text{sigmoid}(W_{ig} X_t + W_{hig} h_{t-1} + b_{ig}) \dots \dots \dots (1)$$

$$f_{\text{forget gate}} = \text{sigmoid}(W_{fg} X_t + W_{hfg} h_{t-1} + b_{fg}) \dots \dots \dots (2)$$

$$C_t = C_{t-1} \times (f_{\text{forget gate}})_t + (i_{\text{input gate}})_t \times (\tanh(W_c X_t + W_{hc} h_{t-1} + b_c)) \dots \dots (3)$$

$$o_{\text{output gate}} = \text{sigmoid}(W_{og} X_t + W_{hog} h_{t-1} + b_{og}) \dots \dots \dots (4)$$

$$h_t = o_{\text{output gate}} \times \tanh(C_t) \dots \dots \dots (5)$$

GRU 모델은 LSTM의 구조를 보다 간소화한 모델로, 구조는 Fig. 5에서 보는 바와 같다. r은 Reset Gate로 이전 노드의 정보를 보존하거나 버리는 선택을 하며, 이는 LSTM의 forget gate와 유사한 기능을 한다. z는 update gate로 다음 노드로 전송할 출력값을 결정한다. \tilde{h} 는 이전 노드의 상태이다(Cho et al., 2014; Fu et al., 2016).

LSTM 및 GRU 모델은 계층 구조, 노드 수, 최적화 알고리즘, epochs 등의 초매개변수 설정 등에 따라 예측력이 결정된다. 그리고 최적화 설정값은 별도로 정의된 것이 없고, 분석 데이터 및 여건 등에 따라 수동으로 조정해야 한다(Le et al., 2011). 따라서 딥러닝 모델 구축 시 일반적으로 사용되는 전략은 다양한 구조의 시나리오 분석을 통해 도출된 결과를 비교하여 최적의 모델을 선택하는 것이다. 본 연구에서는 LSTM과 GRU 모델에 Table 4의 초매개변수를 동일하게 적용하고 두 모델의 예측력을 비교하였다.

모델 구조 중 입력층은 총 17가지의 독립변수를 사용하기 때문에 총 17개의 노드로 구성하였다. 은닉층은 모두 3개의 층으로 구성하였으며, 각 층은 256개, 10개, 그리고 4개의 노드로 구성하였다. 최종적으로는 하나의 노드로 구성된 출력층으로 예

Table 4. Settings of hyper parameter

Layer Structure	Input layer	Hidden layer	Output layer
	17	256-10-4	1
Optimizer algorithm	Adam		
epochs	1,000		

측된 도매가격 값이 반환된다.

최적화 알고리즘으로 사용한 Adam (adaptive moment estimation)은 경사 하강법(gradient descent)을 발전시킨 것으로, Kingma에 의해 제안되었다. AdaGrad와 RMSProp의 결합으로 형성되었으며 대규모의 고차원 머신러닝 학습에 맞게 확장이 가능한 다목적 알고리즘이다(Kingma and Ba, 2015). Epoch은 학습 알고리즘이 실행되는 횟수를 정의하는 초매개변수이다(Brownlee, 2018). Epochs은 구체적으로 딥러닝 모델의 학습 반복 횟수를 의미한다. 본 연구에서는 모델의 오차가 최소화될 때까지 학습하기 위해 epoch의 수를 1,000회로 설정하였다.

학습 데이터(training data)와 테스트 데이터(test data) 비율은 97:3으로 설정하였다. 총 191기의 자료 중 앞의 185기 변수를 학습 데이터로 그리고 뒤의 6기 변수를 테스트 데이터로 활용하였다. 학습, 테스트 데이터의 비율을 7:3 혹은 8:2 등으로 설정하는 것이 일반적이지만, 본 연구에서는 분석 대상이 되는 시계열 자료의 수가 다소 부족하여 충분한 모델 학습을 위해 학습 데이터의 비중을 상대적으로 높게 설정하였다. 그리고 독립변수의 시차는 t-12기 자료부터 t-1기 자료를 이용하여 t기의 종속변수를 예측하도록 모델을 구성하였다.

예측력 평가 지표

분석 결과에 대한 예측력 평가 지표로는 식 (6)과 같이, MAPE를 사용하였다.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - X_i}{Y_i} \right| \times 100(\%) \dots\dots\dots (6)$$

여기서, Y_i 는 실제값 그리고 X_i 는 예측값을 각각 의미한다. MAPE는 오차의 상대적인 지표로, MAE, MSE 등의 지표와는 달리 직관적인 해석이 가능하여 서로 다른 데이터를 이용하는 예측 모델의 성능 비교 시 자주 사용된다(Goodwin and Lawton, 1999; de Myttenaere et al., 2016). 하지만 Y_i 와 X_i 가 0에 가까울수록 불안정해지며, 실제값이 0이면 계산이 불가능하다는 단점 등이 있다(Voyant et al., 2017). 본 연구의 분석 대상 변수들은 이러한 단점에 구애받지 않을 만큼 값의 단위가 크기

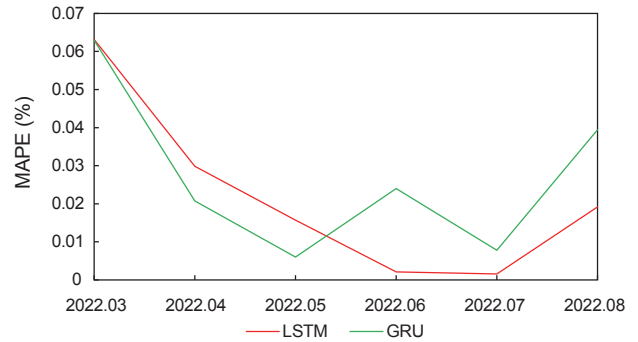


Fig 6. MAPE comparison of LSTM and GRU. MAPE, Mean average percent error; LSTM, Long-short term memory; GRU, Gated recurrent units.

때문에 직관적인 예측력 파악 및 선행연구들과의 비교를 위해 MAPE를 사용하였다.

결 과

본 연구에서는 선정한 17가지 변수를 모두 사용하여 모델별 예측력 비교를 수행하였다. 다양한 변수 조합으로 예측해 보았지만 17가지 변수를 모두 사용한 예측 결과가 가장 우수하였기에 최종적으로 모든 변수를 모델에 투입하였다. LSTM과 GRU로 훈련한 모델로 예측한 결과 및 MAPE는 Table 5와 Fig. 6에서 보는 바와 같다. 우선, LSTM 모델의 최대 오차는 2022년 3월의 954.02원, 최소 오차는 2022년 7월의 24.24원으로 나타났으며, 이에 대한 MAPE는 각각 6.32%와 0.15%로 계산되었다. GRU 모델의 경우, 최대 오차는 2022년 3월의 950.41원, 최소 오차는 2022년 5월의 93.48원으로 나타났으며, 이에 대한 MAPE는 각각 6.29%와 0.60%로 계산되었다. LSTM과 GRU의 6개월 평균 MAPE는 각각 2.19%와 2.68%로 나타나 결과적으로 LSTM을 이용한 예측의 정확도가 상대적으로 높은 것으로 평가되었다.

각각의 예측치를 구체적으로 살펴보면, 예측력이 상대적으로

Table 5. Forecast results and MAPE of LSTM and GRU

	Actual value (won)	LSTM			GRU		
		Forecasted value (won)	Error (won)	APE (%)	Forecasted value (won)	Error (won)	APE (%)
2022.03	15,100	16,054.02	954.02	6.32	16,050.41	950.41	6.29
2022.04	15,406	15,866.92	460.92	2.99	15,726.84	320.84	2.08
2022.05	15,719	15,966.10	247.10	1.57	15,625.52	93.48	0.60
2022.06	16,000	15,966.42	33.58	0.21	15,616.33	383.67	2.40
2022.07	16,438	16,413.76	24.24	0.15	16,310.83	127.17	0.78
2022.08	16,250	16,561.18	311.18	1.91	16,891.13	641.13	3.95
MAPE				2.19			2.68

MAPE, Mean average percent error; LSTM, Long-short term memory; GRU, Gated recurrent units.

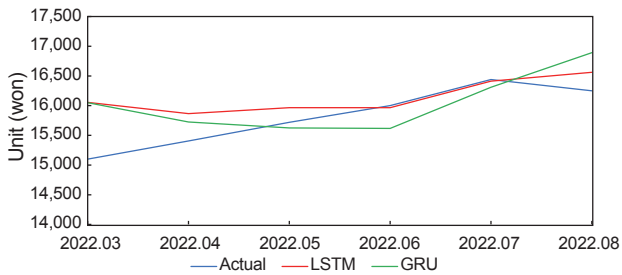


Fig 7. Comparison of forecasted value of LSTM and GRU and actual values. LSTM, Long-short term memory; GRU, Gated recurrent units

떨어지는 시점인 2022년 3월의 경우, 이전의 3월 도매가격들과 달리 높은 수준의 가격이 형성되어 있으며 전월 대비 상승하지 않고 하락한 시점이다. 2022년 8월도 이전의 여름철과 달리 가격이 크게 하락하지 않고 높은 상태로 유지된 시점이다. 이처럼 예측력이 떨어지는 시점은 이례적인 가격 변화 추이를 보인다는 특성이 있다. 하지만 전체적인 추세가 비교적 정확히 예측되었고 두 모델의 오차율이 모두 3% 이하로 평가되었다(Fig. 7).

고 찰

수산물 도매가격의 중요성이 대두되는 가운데 넘치는 우리나라의 대표적인 양식어종임에도 불구하고 도매가격 예측에 대한 연구가 상대적으로 부족한 실정이다. 특히 최근 농·축산물 식품 가격 예측 등 다양한 분야에 활발히 이용되고 있는 딥러닝 기법 기반의 시계열 예측은 넘치의 도매가격 예측뿐만 아니라 수산물 분야 전반적으로 활용도가 낮다.

본 연구에서는 17가지 변수를 포함한 LSTM과 GRU 모델을 구축하여 양식산 넘치의 도매가격을 예측하고, 모델별 예측력을 평가·비교해 보았다. 2006년 10월부터 2022년 8월까지 총 191기의 월별 자료 중 185기의 자료로 모델을 학습하고 나머지 6기의 자료로 학습된 모델을 테스트해 본 결과, 모델별 평균 MAPE가 LSTM은 2.19%, 그리고 GRU는 2.68%로 나타나 LSTM의 예측력이 상대적으로 우수한 것으로 평가되었다.

본 연구에서는 특히 수산물 가격 예측을 위해 다변량 딥러닝 모델을 구축 및 활용하여 단변량 딥러닝 모델 혹은 전통적 시계열 모형 등을 이용한 선행연구들의 한계점을 보완하였다. Nam and Jeong (2017)의 연구에서는 전통적 시계열 모형을 이용하여 넘치의 산지가격을 예측하고 MSE와 MAE로 정확도를 평가하였는데, 각 모델의 예측치로 다시 MAPE를 계산해 본 결과 가장 예측력이 좋은 ARMA (1,1) 모형의 MAPE는 7.48%로 계산되었다. Woo and Shin (2022)에서 LSTM과 MLP 모델을 단변량으로 구축하여 조피볼락 도매가격을 예측한 결과, 두 모델의 MAPE는 각각 32%와 17%로 평가되었다. 반면, 본 연구에서는 다양한 투입 변수를 활용하여 모델을 구축하여 약 2% 수

준의 MAPE를 도출하였기에 예측 오차율이 큰 선행연구의 한계점을 다소 보완하였다고 볼 수 있다. 또한 향후 수산물 가격 예측 분야에서 전통적 시계열 모형의 한계점을 보완하는 다변량 LSTM 및 GRU 모델의 충분한 활용 가능성을 보였으며, 특히 다양한 딥러닝을 더욱 활발히 활용할 필요가 있을 것이다.

양식산 수산물의 도매가격은 매우 중요한 정보임에도 불구하고 아직 정확한 예측이 이루어지지 않고 있다. 양식산 수산물의 도매가격에는 산지가격, 생산량, 환경적 요인 그리고 대체 어종의 수급 상황 등이 복합적으로 영향을 미친다. 따라서 도매가격 예측 시 이들 요인에 대한 고려가 수반되어야 하며, 변수의 종류나 양적인 제한 없이 사용 가능한 딥러닝 모델을 활용한다면 보다 다양한 요인을 고려한 상대적으로 정확한 예측이 가능할 수 있을 것으로 기대된다. 이러한 예측의 결과는 생산자 소비자, 정부를 포함한 모든 시장 및 산업 구성원에게 유용한 정보로 활용될 수 있다. 예를 들어, 생산자는 연구에서 활용된 변수에 해당하는 다양한 요소들이 수산물 가격에 직접적인 영향을 미침을 고려하여 계획 경영을 수행하고, 유통업자 및 소비자는 예측된 가격을 기반으로 합리적인 소비를 행할 수 있다. 또한, 정부 및 지자체에서는 수산물 수급 조절이나 지원 정책과 같은 정책 수립 시 예측된 가격을 기반으로 보다 타당한 정책을 수립할 수 있을 것이다.

본 연구의 한계점으로는 우선 활용 가능한 자료가 월별 자료로 한정되어 있어 모델 학습이 충분하지 못한 점이다. 그리고 2020년 이후와 같이 가격 변화 추이가 이례적이었던 시기에 대한 정확한 원인을 규명하지 못했다. 이에 대해 본 연구에서는 COVID-19 상황을 더미 변수화하여 활용하였지만, 이를 더욱 세분화된 시계열 자료로 보완하거나 양식경영 여건 변화 및 국내의 정세 변화를 종합적으로 고려하여 모델에 반영할 수 있을 것이다. 따라서 향후 보다 발전된 수산물 가격 예측을 위해서는 다량의 시계열 데이터를 확보하여 예측 모델을 고도화하고, 보다 다양한 변수들을 포함시켜야 할 것이다.

References

- Adebiyi A, Adewumi A and Ayo K. 2014. Comparison of arima and artificial neural networks models for stock price prediction. *J Appl Math* 2014, 614342. <https://doi.org/10.1155/2014/614342>.
- Bae K and Kim C. 2016. An agricultural estimate price model of artificial neural network by optimizing hidden layer. *J Korean Ins Inf Technol* 14, 161-169. <https://doi.org/10.14801/jkiit.2016.14.12.161>.
- Brownlee J. 2018. What is the Difference Between a Batch and an Epoch in a Neural Network?. Retrieved from <https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/> on Sep 1, 2022.
- Cho K, Van M, Gulcehre C, Bahdanau D, Bougares F, Schwenk H and Bengio Y. 2014. Learning phrase representations

- using RNN encoder-decoder for statistical machine translation. arXiv:1406.1078, 1-15. <https://doi.org/10.48550/arXiv.1406.1078>.
- Cooper J. 1999. Artificial neural networks versus multivariate statistics: An application from economics. *J Appl Stat* 26, 909-921. <https://doi.org/10.1080/02664769921927>.
- de Myttenaere A, Golden B, Le Grand B and Rossi F. 2016. Mean absolute percentage error for regression models. *Neurocomputing* 192, 38-48. <https://doi.org/10.1016/j.neucom.2015.12.114>.
- Fu R, Zhang Z and Li L. 2016. Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction. 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, China, 324-328. <https://doi.org/10.1109/YAC.2016.7804912>.
- Goodwin P and Lawton R. 1999. On the asymmetry of the symmetric MAPE. *Int J Forecast* 15, 405-408. [https://doi.org/10.1016/S0169-2070\(99\)00007-2](https://doi.org/10.1016/S0169-2070(99)00007-2).
- Gordon D. 2017. Price modelling in the canadian fish supply chain with forecasts and simulations of the producer price of fish. *Aquac Econ Manag* 21, 105-124. <https://doi.org/10.1080/13657305.2017.1265021>.
- Guttormsen A. 1999. Forecasting weekly salmon prices: Risk management in fish farming. *Aquac Econ Manag* 3, 159-166. <https://doi.org/10.1080/13657309909380242>.
- Hochreiter S and Schmidhuber J. 1997. Long short-term memory. *Neural Comput* 9, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Hwang K, Choi J and Oh T. 2012. Forecasting common mackerel auction price by artificial neural network in Busan cooperative fish market before introducing TAC system in Korea. *J Korean Soc Fish Ocean Technol* 48, 72-81. <https://doi.org/10.3796/KSFT.2012.48.1.072>.
- Im J, Kim W, Byoun W and Shin S. 2018. Fruit price prediction study using artificial intelligence. *J Converg Cult Technol* 4, 197-204. <https://doi.org/10.17703/JCCT.2018.4.2.197>.
- Kang SK. 2015. The causality and volatility spillover between farming fish species in consumption replacement relation. *J Fish Bus Adm* 46, 119-127. <https://doi.org/10.12939/FBA.2015.46.3.119>.
- Kingma D and Ba L. 2015. Adam: A method for stochastic optimization. International Conference on Learning Representations (ICLR). arXiv:1412.6980, 1-15. <https://doi.org/10.48550/arXiv.1412.6980>.
- Kim N and Kang K. 2019. 2018 Fisheries Consumption Behavior. Korea Maritime Institute, Busan, Korea, 15-30.
- KMA (Korea Meteorological Administration). 2022. National Climate and Weather Database. Retrieved from <https://data.kma.go.kr> on Sep 1, 2022.
- KMI (Korea Maritime Institute). 2022. Fisheries Outlook Statistics. Retrieved from <https://www.foc.re.kr> on Sep 1, 2022.
- Ko BH. 2009. Price volatility, seasonality and day-of-the week effect for aquacultural fishes in Korean fishery markets. *J Fish Bus Adm* 40, 49-70.
- KOSIS (Korean Statistical Information Service). 2022. Survey on Fishery Processing Industry. Retrieved from https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1EW0001 on Sep 1, 2022.
- KREI (Korea Rural Economic Institute). 2022. 2020 Food Balance Sheet. KREI, Naju, Korea.
- KSC (Korea Customs Service). 2022. Trade Statistics. Retrieved from <https://unipass.customs.go.kr> on Sep 1, 2022.
- Le Q, Ngiam J, Coates A, Lahiri A, Prochnow B and Ng A. 2011. On optimization methods for deep learning. In: Proceedings of the 28th International Conference on International Conference on Machine Learning. Bellevue, Washington D.C., U.S.A.
- Lee DC, Kim DH, Kim SM, Kang MS, Hong MJ, Kim HJ and Park SI. 2002. Effects of stress induced by changes of water temperature on the non-specific defense mechanism in cultured olive flounder, *Paralichthys olivaceus*. *J Fish Pathol* 15, 65-75.
- Lee GH and Kim DH. 2022. A study on forecasting the producer price of farmed Korean rockfish using deep learning models. *Korean J Food Mark Econom* 39, 21-41. <https://doi.org/10.47085/KJFME.39.4.2>.
- Lee HD and Ma CM. 2020. Asymmetric transmission between producer and wholesale prices in farmed olive flounder market. *J Fish Bus Adm* 51, 69-83. <https://doi.org/10.12939/FBA.2020.51.4.069>.
- Lee JM and Kim KS. 2010. An empirical study on asymmetric price transmissions in the distribution channels of fisheries market. *J Fish Bus Adm* 41, 59-78.
- MOF (Ministry of Oceans and Fisheries). 2021. 2020 A Survey on the Status of Marine Products Production and Distribution Industry. MOF, Busan, Korea.
- Mun SE, Jang SB, Lee JH and Lee JS. 2016. Machine learning and deep learning technology trends. *Inform Commun Magaz* 33, 49-56.
- Najafabadi M, Villanustre F, Khoshgoftaar T, Seliya N, Wald R and Muharemagic E. 2015. Deep learning applications and challenges in big data analytics. *J Big Data* 2, 1. <https://doi.org/10.1186/s40537-014-0007-7>.
- Nam JO, Baek EY and Noh SG. 2014. Forecast for laver producer price using time series models. *Ocean Pol Res* 29, 271-303. <https://doi.org/10.35372/kmiopr.2014.29.2.010>.
- Nam JO and Jeong MJ. 2017. Comparison on forecasting power and monthly producer price forecast of Jeju olive flounder. *Ocean Pol Res* 32, 1-22. <https://doi.org/10.35372/kmiopr.2017.32.2.001>.
- Selvin S, Vinayakumar R, Gopalakrishnan E, Menon V and So-man K. 2017. Stock price prediction using LSTM, RNN and CNN-sliding window model. In: 2017 International Conference on Advances in Computing, Communications and In-

- formatics (ICACCI), Udupi, India, 1643-1647. <https://doi.org/10.1109/ICACCI.2017.8126078>.
- Shahid F, Zameer A and Muneeb M. 2020. Predictions for covid-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos Solit Fractals* 140, 110212. <https://doi.org/10.1016/j.chaos.2020.110212>.
- Shin S, Lee M and Song S. 2018. A prediction model for agricultural products price with LSTM network. *J Korea Con Assoc* 18, 416-429. <https://doi.org/10.5392/JKCA.2018.18.11.416>.
- Song YN. 2021. A study on predicting the producer price of farmed olive flounder in Jeju Island using artificial neural network models. M.S. Thesis, Pukyong National University, Busan, Korea.
- Voyant C, Notton G, Kalogirou S, Nivet M, Paoli C, Motte F and Fouilloy A. 2017. Machine learning methods for solar radiation forecasting: A review. *Renew Energ* 105, 569-582. <https://doi.org/10.1016/j.renene.2016.12.095>.
- Wang H, Lei Z, Zhang X, Zhou B and Peng J. 2019. A review of deep learning for renewable energy forecasting. *Energy Conv Manag* 198, 111799. <https://doi.org/10.1016/j.enconman.2019.111799>.
- Wohlgenant M. 1985. Competitive storage, rational expectations, and short-run food price determination. *Am J Agric Econ* 67, 739-748. <https://doi.org/10.2307/1241813>.
- Woo KW and Shin YM. 2022. Prediction of rockfish price using ARMA model and machine learning technique. *J Fische Mar Sci Educ* 34, 415-426. <https://doi.org/10.13000/JFMSE.2022.6.34.3.415>.
- Yun SJ, Lee CS and Yang SR. 2016. Development of price forecast models for international grains using artificial neural networks. *Korean J Agricult Econ* 57, 83-101.