Araştırma Makalesi
Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi
16 (2022) 204-218

RECURRENT NEURAL NETWORK BASED MODEL DEVELOPMENT FOR WHEAT YIELD FORECASTING

Halit ÇETİNER[1*], Burhan KARA[2]

[1]Isparta University of Applied Sciences, Vocational School of Technical Sciences, Isparta, Turkey
[2]Isparta University of Applied Sciences, Agriculture Faculty, Department of Field Crops, Isparta, Turkey
Geliş Tarihi/Received Date: 17.02.2022 Kabul Tarihi/Accepted Date: 11.04.2022 DOI: 10.54365/adyumbd.1075265

ABSTRACT

In the study carried out in line with the stated purposes, monthly rain, humidity and temperature data, wheat production amount, and wheat productivity data of Konya province between 1980-2020 were used. Using these data, wheat productivity estimation was performed with (Gated Recurrent Units) GRU and Long Short Term Memory (LSTM) methods, which are Recurrent Neural Network (RNN) based algorithms. When wheat productivity estimation performance was examined with the implemented GRU-based model, 0.9550, 0.0059, 0.0280, 0.0623, 7.45 values were obtained for the R^2 score, MSE, RMSE, MAE and MAPE values, respectively. In the performance results obtained with the LSTM method, which is another RNN-based method, 0.9667, 0.0054, 0.0280, 0.0614, 7.33 values were obtained for the R^2 score, MSE, RMSE, MAE and MAPE values, respectively. Although the LSTM method gave better results than the GRU method, the training modelling time of the LSTM method took longer than that of the GRU method.

Keywords: *Wheat yield, wheat production, GRU, LSTM, regression analysis*

BUĞDAY VERİM TAHMİNİ İÇİN YENİLEMELİ SİNİR AĞI TABANLI MODEL GELİŞTİRME

ÖZET

Bu çalışmada 1980-2020 yılları arasında Konya ilinin aylık yağış, nem ve sıcaklık verileri, buğday üretim miktarı ve buğday verimlilik verileri kullanılmıştır. Bu veriler kullanılarak Recurrent Neural Network (RNN) tabanlı algoritmalar olan (Gated Recurrent Units) GRU ve Long Short Term Memory (LSTM) yöntemleri ile buğday verimlilik tahmini yapılmıştır. Gerçekleştirilen GRU tabanlı model ile buğday verimliliği tahmin performansları incelendiğinde R^2 puan, MSE, RMSE, MAE ve MAPE değerleri için sırasıyla 0.9550, 0.0059, 0.0280, 0.0623, 7.45 değerleri elde edilmiştir. RNN tabanlı bir diğer yöntem olan LSTM yöntemiyle elde edilen performans sonuçlarında ise R^2 puan, MSE, RMSE, MAE ve MAPE değerleri için sırasıyla 0.9667, 0.0054, 0.0280, 0.0614, 7.33 değerleri elde edilmiştir. LSTM yöntemi, GRU yönteminden daha iyi sonuçlar vermesine rağmen LSTM yönteminin eğitim modelleme süresi GRU yönteminden daha fazla sürmüştür.

Anahtar Kelimeler: *Buğday verimi, buğday üretimi, GRU, LSTM, regresyon analizi*

1. Introduction

Agriculture is one of the most important areas in the development of countries around the world. The agricultural sector directly affects nutrition and food statistics in the world beyond development.

* e-posta[1] : halitcetiner@isparta.edu.tr ORCID ID: https://orcid.org/0000-0001-7794-2555 (Sorumlu Yazar)
e-posta[2] : burhankara@isparta.edu.tr ORCID ID: https://orcid.org/0000-0002-4207-0539

Continuous and regular crop production is a major challenge for agricultural farmers. Achieving standard crop production is not possible due to changing weather, water, and soil conditions. Many factors affect crop production in a wide range, such as the type of land desired to be cultivated, the harsh or hot climate of planting. For the reasons stated, researchers around the world are trying to find methods that can accurately predict farmers' crop productivity by overcoming the stated challenges [1–3].

There has been a great increase in food insecurity worldwide since 2015 [4]. Cao et al., Dodds and Bartram, Gorelick et al. estimate that two billion people will be added to the current food-supplied population in about three decades [3, 5]. In this situation, it is expected that the demand for food will double compared to today [6]. Approximately 40% of the world's population lives on wheat. To meet the food needs of the growing population in Turkey, as in the world, the production of the 6.43 million hectares of winter wheat production area should be planned so as to maintain wheat productivity [7]. However, due to various reasons, wheat cultivation in the world has stopped in agricultural areas. In this situation, wheat producers and consumers are believed to have difficulties meeting the increasing demand [8]. It has been determined that the wheat yields of different countries have stopped for many years [9]. For this reason, accurate and timely estimation of wheat productivity has gained importance to protect the interests of agricultural producers worldwide and to ensure global food security.

In addition to the difficulties mentioned above, there are irregularities in temperature and rainfall distributions due to climate change. Due to these irregularities and uncertainties, agricultural production and food security are affected [10, 11]. Taking into account studies that state that the global temperature will increase by 2.5 °C in approximately 30 years, the increase in temperatures is expected to affect semi-arid regions such as Turkey more. The increase in temperature in such regions is believed to reduce crop production [11, 12]. The increase in global temperatures has begun to become more noticeable today [2, 13]. In projection studies carried out to determine the future increase in global warming, it is expected that the temperature in Turkey will increase by 2-3 °C in about 20 years [14]. It is stated that after this increase in temperature, serious risks will occur in crop production, especially wheat, in countries affected by climate change, especially in Turkey [15]. The wheat crop is an important food crop, mainly grown in rainy conditions [16]. According to the information reported by TUIK, there has been a decrease in tons of wheat production, especially in the production of barley, rye and oats [17]. According to the Cline study, it is stated that there will be a decrease in rain rates and an increase in temperature in Turkey in the next 50 years [18]. As a result of the increase in temperature, a yield loss of around 6.0% is expected in the wheat crop, which is deprived of genetic improvement and effective adaptation process [19]. It is determined that there will be a shortening in the duration of the growing season after increasing temperature, as a result of which there will be a decrease in wheat grains and a decrease in yield [20, 21].

In many literature studies, machine learning techniques such as random forest, artificial neural networks, multivariate regression methods have been used. The developed machine learning models take air components and soil conditions as inputs to obtain crop yield. Yield estimation studies in the literature are grouped under two different headings. While yield estimation is performed with machine learning approaches in the first title, it is seen that yield estimations are made with deep learning approaches in the second title. Jeong et al. [22] used the random forest algorithm to estimate yields for wheat, maize, and potato crops. It is stated that the random forest algorithm achieves a better result than the machine learning method called linear regression in wheat, corn and potato crops. Shahhosseini et al. [23] predicted corn yield and nitrate loss using a random forest machine learning algorithm like Jeong et al. [22] in their studies. Jiang et al. [24] have estimated the yield of wheat crop based on climate data with artificial neural network and multilinear methods. As a result of the study, it is stated that the artificial neural network model outperforms the multiple linear regression method in the prediction of wheat yield in the northern China region. It is reported that deep learning methods with layers with different properties perform well compared to artificial neural network models with a single hidden layer used in yield estimation [25]. In the literature, it is emphasized that the layers that do not use hand-made features in the deep learning models used in yield estimation contribute to higher accuracy. Deep learning methods have learning methods that will increase the level of representation of a raw input [26].

In the estimation of crop yield, which is extremely difficult, problems can arise due to crop genotype and environmental factors. To overcome the stated difficulties, a study based on CNN and RNN yield estimation is presented in the literature [25]. In this study, yield estimation was carried out using data from corn and soybeans in 2016, 2017 and 2018 in the United States. You et al. [27] developed a soybean yield prediction model based on CNN and RNN. Khaki and Wang [28] developed a model based on data from 2008 to 2016 to estimate corn yields. It is reported that the model gives better results than classical machine learning methods such as regression tree. Kim et al. [29] developed a deep neural network model for crop yield prediction using a meteorological dataset from 2006 to 2015. Tian et al. [30] developed an LSTM model using meteorological data from the People's Republic of China. With the model they developed, they estimate the wheat yield. The model they propose gives a better result than the SVM method, which is one of the classical machine learning methods.

Due to various reasons such as global warming, irregularity in rainfall, population growth, it has been decided to support the crop production planning of farmers living around the world. For this purpose, the main contributions of the study carried out for an accurate and timely estimation of wheat productivity in effective production planning that will support global food security are listed below.

- Data normalization was applied to evaluate the monthly rain, humidity and temperature data, wheat production amount, and wheat productivity data of Konya province between 1980-2020 accurately and quickly.
- Two different deep learning models based on LSTM and GRU have been developed to accurately analyse the wheat yield estimation with normalized data.
- The R^2 score, MSE, RMSE, MAE, and MAPE measurement metrics are presented comparatively to evaluate the performance of two different models proposed in the estimation of wheat yield.
- When the measurement metrics were compared, in the test processes, 0.9667, 0.0054, 0.0280, 0.0614, 7.33 and 0.9550, 0.0059, 0.0280, 0.0623, 7.45 R^2 scores, MSE, RMSE, MAE and MAPE measurement metrics were obtained in the LSTM and GRU models, respectively.

The sections after this step of the article are planned as follows. In the second section, we introduce the data set prepared for wheat productivity and deep learning models that perform regression analysis on this data set. In the third section, comparatively, the results of the R^2 score, MAPE, RMSE, MSE, MAE performance obtained from LSTM, and GRU based deep learning models are presented. In the last section, the study is concluded.

2. Material and Methods

In this section of the study, the dataset used to predict wheat productivity and the basis of two different deep learning architectures that enable regression-based analysis are explained.

2.1. Material

To perform regression analysis in the study, a data set consisting of monthly rain, humidity and temperature data, wheat production amount, and wheat productivity data was created between 1980 and 2020 in Konya province. In the created data set, there are monthly average relative humidity (%), monthly average temperature (°C) and monthly total rainfall (mm=kg÷m²). These values were obtained from the T.C. Ministry of Agriculture and Forestry. Wheat productivity data were obtained from the Turkish Statistical Institute (TUIK) central distribution system.

The data used in the regression analysis were first normalized to the 0-1 range to perform faster processing. Year, temperature, humidity, rain, cultivated area, and production amount values were used as input attributes in the study. The obtained wheat yield value was used as the output value. The inputs

from the deep learning model used as input attributes are given in Table 1. All data presented in Table 1 were used in the study. In order not to get a different result when running each deep learning model, the data set is divided into two separate parts as training and testing according to the K-fold 5 value.

Table 1. Features used in wheat yield estimation

Raw Inputs Year	Inputs					Outputs
Year	Temperature	Humidity	Rain	Cultivated Area	Production Amount	Yield
1980	11.349	62.216	32.06	9,209.930	1,878.825	204
1981	12.108	61.116	28.091	9,262.190	1,713.505	185
….	….	…	…	…	….	…
1990	10.658	59.366	19.25	8,890.250	1,813.611	204
….	….	…	…	…	….	…
2000	10.858	58.483	21.541	7,959.120	1,806.615	237
….	….	…	…	…	….	…
2007	12.608	66.525	21.808	6,751.320	1,026.565	343
….	….	…	…	…	….	…
2017	11.775	59.399	25.983	7,468.193	2,192.574	1.339
….	….	…	…	…	….	…
2020	13.116	56.674	24.308	6,202.606	1,921.433	1.355

The year, input and output values of the input data defined in Table 1 are shown. Temperature, humidity, and rainfall data from these input data includes annual average data. The average data was obtained by taking the average of the data collected from January to December of temperature, humidity, and rainfall data. In the specified structure, there are 5 inputs under normal conditions: temperature, humidity, rainfall, cultivated area, and production amount. Considering that the temperature, humidity, and rainfall data are the averages of the data for all months from January to December, each input has 12 values. By adding monthly inputs of temperature, humidity and rainfall data, cultivated area, and production amount inputs, it reaches 38 input values in total.

Table 2. Monthly average temperature data

Raw Inputs Year	Temperature Inputs (°C)											
Year	January	February	March	April	May	June	July	August	September	October	November	December
1980	-1.4	-0.3	4.8	10.1	15.3	20.8	25.4	22.4	17.1	12.7	6.5	2.8
1981	1.5	2.0	8.0	11.0	13.3	20.4	23.1	21.7	19.1	15.0	4.8	5.4
….	….	…	…	…	….	….	…	…	…	….	….	…
1990	-4.4	-0.4	5.9	10.0	13.4	19.2	23.6	21.3	17.3	12.2	7.4	2.4
….	….	…	…	…	….	….	…	…	…	….	….	…
2000	-5.2	-2.3	3.8	12.7	14.6	19.3	26.0	22.6	19.0	11.5	7.2	1.1
….	….	…	…	…	….							
2007	0.3	0.3	6.6	8.9	19.1	22.7	25.3	25.6	20.1	14.0	7.1	1.3
….	….	…	…	…	….	….	…	…	…	….	….	…
2017	-4.8	-1.5	6.9	10.8	15.4	20.4	25.3	24.3	22.4	12.5	6.2	3.4
….	….	…	…	…	….	….	…	…	…	….	….	…
2020	0.4	2.8	7.2	10.8	15.9	20.3	25.5	24.2	22.6	17.1	6.0	4.6

In Table 2, the monthly temperature values that enable the annual average temperature values in Table 1 to be obtained are shown. The temperature input is obtained by calculating the average of these temperature values.

Table 3. Monthly average relative humidity

Raw Inputs Year	Humidity Inputs (%)											
Year	January	February	March	April	May	June	July	August	September	October	November	December
1980	77.5	80.1	72.4	66.8	57.8	46.0	39.3	45.4	50.3	63.3	73.6	74.1
1981	83.7	76.9	65.2	50.1	55.7	51.3	45.6	48.8	47.6	62.5	71.5	74.5
….	….	…	…	…	….	…	…	…	….	….	…	…
1990	76.3	77.8	55.4	58.0	61.2	47.4	42.9	40.6	47.9	55.5	71.6	77.8
….	….	…	…	…	….	…	…	…	….	….	…	…
2000	75.6	77.2	60.6	58.0	59.5	46.7	29.3	46.6	44.2	60.9	62.3	80.9
….	….	…	…	…	….	…	…	…	….	….	…	…
2007	84.0	86.7	74.8	69.5	59.5	52.0	37.8	43.1	46.8	69.3	85.2	89.6
….	….	…	…	…	….	…	…	…	….	….	…	…
2017	85.6	77.8	63.6	53.0	58.1	54.7	35.3	45.2	31.7	53.4	73.6	80.8
….	….	…	…	…	….	…	…	…	….	….	…	…
2020	77.0	72.2	67.6	59.5	53.6	47.9	36.4	31.4	42.6	46.4	68.1	77.4

Table 3 shows the monthly relative humidity values, which allow the annual average relative values in Table 1 to be obtained. By calculating the average of these relative humidity values, the relative humidity input values are obtained.

Table 4. Monthly total rainfall

Raw Inputs Year	Rain Inputs (mm)											
Year	January	February	March	April	May	June	July	August	September	October	November	December
1980	33.3	33.0	42.0	73.5	57.6	21.2	0.8	0	3.9	69.7	33.9	15.9
1981	112.2	32.5	19.3	18.2	40.7	23.5	13.9	0.3	0.0	10.2	15.1	51.2
….	….	…	…	…	….	…	…	…	….	….	…	…
1990	9.2	24.9	3.0	17.1	41.4	8.0	0.2	0	25.7	27.3	22.5	51.7
….	….	…	…	…	….	…	…	…	….	….	…	…
2000	30.1	15.2	11.2	38.7	56.2	17.6	0	4.4	4.5	32.3	26.2	22.1
….	….	…	…	…	….	…	…	…	….	….	…	…
2007	20.9	19.3	15.4	16.1	16.3	15.9	0.4	6.0	4.1	25.5	68.0	53.8
….	….	…	…	…	….	…	…	…	….	….	…	…
2017	30.6	2.4	61.4	33.9	45.6	22.6	0.0	19.3	3.9	14.6	62.8	14.7
….	….	…	…	…	….	…	…	…	….	….	…	…
2020	48.7	36.5	51.8	35.3	43.5	23.9	0.9	0.4	6.9	4.1	19.6	20.1

In Table 4, the monthly total rainfall values that provide the annual total rainfall in Table 1 are shown. By calculating the average of these total rainfall values, the total rainfall input

values are obtained. To reduce the number of entries from 38 to 5, a data set entry is defined, which is obtained from annual averages by calculating the average temperature, humidity, and total rainfall values. When the annual average values are obtained, it is ensured that the annual crop yield estimation is made from the inputs given to the model.

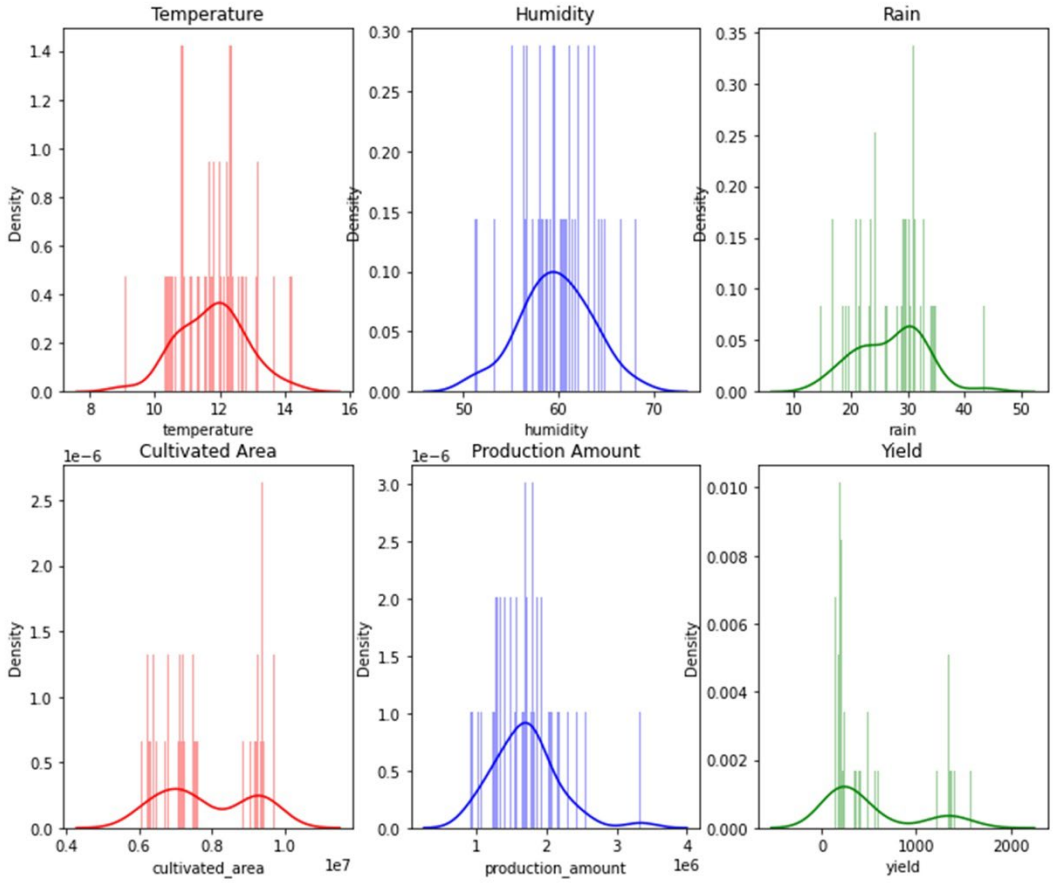

Figure 1. Input and output parameters in the used data set

2.2. Methods

In the study, methods based on the Recurrent Neural Network (RNN) architecture, which are popular recently and are frequently used in the analysis of serial data, are used. For this purpose, a two-door Gated Recurrent Units (GRU) structure was used for wheat productivity analysis. After the GRU structure, the LSTM structure was used, which resolved the gradient burst problems in the RNN structures. It is difficult to train models consisting of RNN structure in training datasets consisting of long-range data [31]. To overcome this training difficulty, models such as GRU and LSTM, which can be trained with long-range data, are preferred.

2.2.1. GRU

Structures that do not contain memory units to control information flow in RNN structures are called GRU structures. In this structure, all confidential situations can be used without any information flow control. Compared to RNN models, models based on the GRU structure have fewer parameters. As a result of fewer parameters, the processing load is reduced and faster training is realized. In addition, generalizations can be made with very little data. The success of retrospective transactions made from

the current step to the previous steps is reported to be low [32]. However, it is preferred because it can generalize with fewer parameters and performs fast training modeling.

There are two gates for the modeling of GRU structures [33]. One of these gates is the reset gate, which decides how a new input is combined with the memory from the previous step. The second gate of the GRU structure determines how long the state before the current step will remain in memory. The equation defining the gates of the GRU structure is given in detail below.

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (1)$$

$$Update\ gate\ (z_t) = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (2)$$

$$\tilde{h}_t = \tanh\big(W_h x_t + r_t \odot (U_h h_{t-1})\big) + b_h \quad (3)$$

$$Reset\ gate(r_t) = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (4)$$

The expressions defined in Equations 1-4 protect important features. At the same time, data protection is required for long-term transfers of these features [34]. Among the symbols in Equations 1-4, W represents the weight at time t, and h_{t-1} represents the values of the hidden layer at time $t-1$. U and σ represent the cell units in GRU and the sigmoid activation function, respectively [33]. Equation 1 shows the linear interpolation of $\tilde{h}_t$ with new regression information to h_{t-1}, the situation at time $t-1$. z_t, which is one of the important gates of the GRU structure, decides how much new information can be added and how much old information can be kept. x_t is the vector at time t given as input to the GRU structure. The r_t is a gate that checks how much the state at time $t-1$ affects the current state. The small r_t value indicates that less information is retained at time t than at time $t-1$.

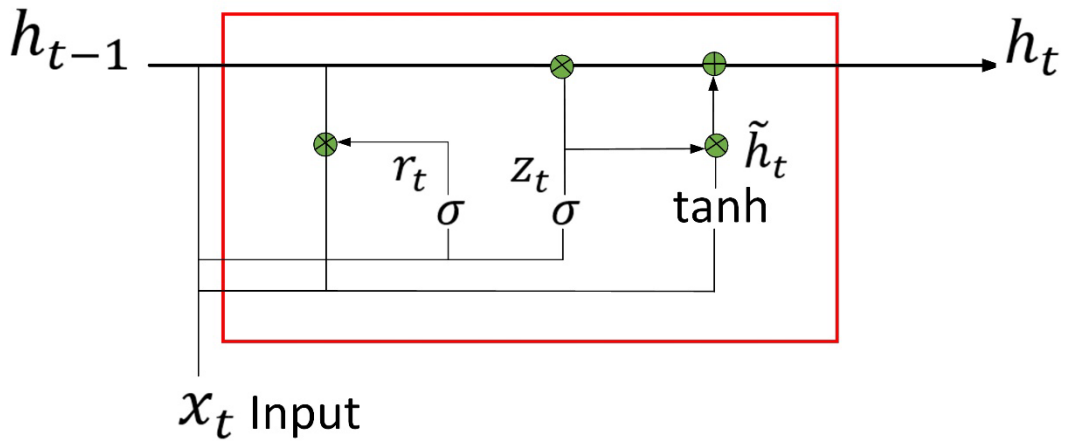

Figure 2. GRU structure internal diagram

2.2.2.LSTM

Depending on the input depth, the training process of recurrent neural networks can take a lot of time. Loss functions in these neural networks can also have varying sensitivities. Depending on the variability of the losses in different layers, different gradient values can be obtained in different layers [35]. Problems with vanishing and gradient bursts due to variable gradient values are frequently encountered in RNN structures [36]. This problem, which occurs as a result of multiplying the weight matrices one after another, arises in the backpropagation stage of the RNN algorithm with the disappearance and explosion of the gradient. Despite the mentioned drawbacks, RNN constructs give good success in non-long-term transactions between the current step and pre-current step [37]. However, the mentioned disadvantages affect recall in long-term transactions. The LSTM method has been developed to reduce forgetting by increasing recall in long-term processes [38, 39]. LSTM is a basic

format of RNN structures with sequential data entry. The LSTM structure maintains a chain structure that animates time steps in time series [3].

In LSTM structures, h_t showing the current hidden state, h_{t-1} representing the state of the previous step, and x_t depending on the input from outside. LSTM structures have three doors and one layer [40]. These doors and the interior details of the layer are shown in Figure 3. The terms f_t, i_t, and o_t for forgetting, entry and exit gates at time *t*, respectively, are shown in Figure 3. The state layer is also shown with the g_t symbol.

$$f_t = \sigma\left(W_f X_t + U_f h_{t-1} + b_f\right) \quad (5)$$

$$i_t = \sigma(W_i X_t + U_i h_{t-1} + b_t) \quad (6)$$

$$o_t = \sigma(W_o X_t + U_o h_{t-1} + b_o) \quad (7)$$

$$g_t = tanh\left(W_g x_t + R_g h_{t-1} + b_g\right) \quad (8)$$

In the equations in Figures 5-8, the weight, repetitive weight, and bias values are indicated by the symbols W, R, and b, respectively.

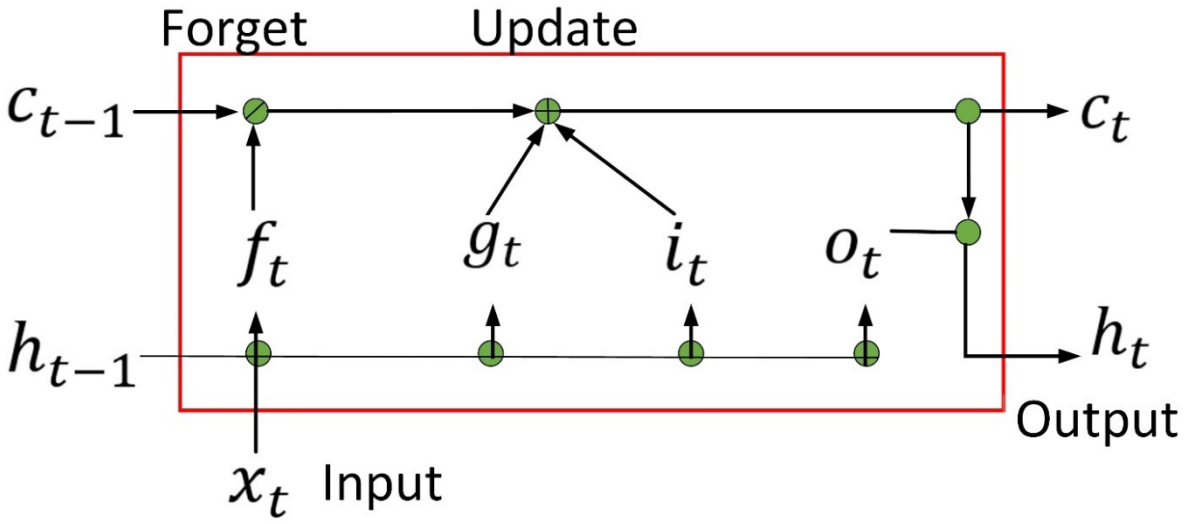

Figure 3. LSTM cell structure

3. Experimental Results

In the regression analysis, the data set consisting of monthly rain, humidity, temperature data, wheat production amount, and wheat productivity data was first normalized to the 0-1 range. The normalized data are divided into two parts, training and testing, to model according to the K-fold 5 value. The training data are modeled with a structure based on the GRU algorithm detailed in Figure 2. The parameters of the GRU algorithm used in training modeling are presented in Table 5.

Table 5. Used GRU model parameters

Parameter	Value
Layers	1, 5, 10
Loss	Mean absolute error
Optimizer	Adam, Adamax, RMSprop

Epochs	50, 75, 100
Batch size	4, 32, 74
Activation	Tanh, ReLU

In this article, MAPE, MAE, R^2, MSE, and RMSE performance measurement metrics presented in Equations 9-13 were used to measure the success of models developed with GRU and LSTM, which are deep learning models proposed to predict wheat productivity [40].

$$MAPE = \frac{100}{m}\sum_{i=1}^{m}\left[\frac{Y_i - \hat{Y}_i}{Y_i}\right] \tag{9}$$

$$MAE = \frac{100}{m}\sum_{i=1}^{m}\left[\frac{Y_i - \hat{Y}_i}{Y_i}\right] \tag{10}$$

$$R^2 = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2} \tag{11}$$

$$MSE = \frac{1}{m}\sum_{i=1}^{m}\left(Y_i - \hat{Y}_i\right)^2 \tag{12}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(Y_i - \hat{Y}_i\right)^2} \tag{13}$$

In Table 5, the recommended GRU structures recommended for wheat productivity are given. As an optimization method, Adamax, RMSprop, and Adam optimization methods have been tested. After the test processes, the best performance values for R^2, MSE, RMSE, and MAPE were seen to be obtained by the Adam optimization method. 50, 75, and 100 steps were tested separately in the testing of the proposed GRU method. To provide a comprehensive overview of these steps, the graphs obtained as a result of 100-step tests are presented. The best results were obtained in the 10-layer structure of the model, which was created from 1, 5, and 10-layer GRU structures in series. To measure the performance of the proposed model in wheat productivity, the performance measurement values of R^2, MSE, RMSE, and MAPE, which are widely used in the literature, were used.

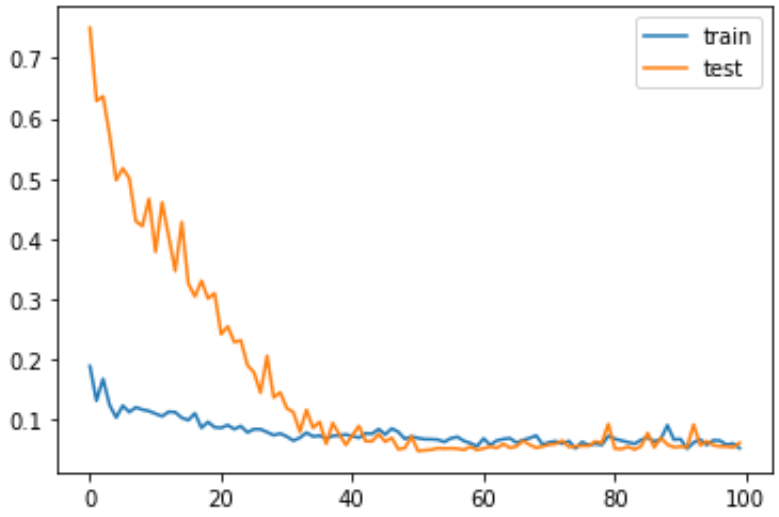

Figure 4. GRU model training and test loss graph

Training and test loss graphs are given in Figure 4 to support these results. Although there is a large difference between the training and test loss values at the beginning, after the 38th step, the training and test loss continue to equalize. In Figure 5, the graph showing the actual and predicted values

obtained as a result of the regression analysis of wheat productivity is shown. Although the difference is high in the first stages of the test process, it is seen that the difference decreases towards the last stage of the iteration.

Table 6. Performance results of the GRU model

Algorithm	R^2 Score	MSE	RMSE	MAE	MAPE
GRU Model with Adam (Training)	0.9403	0.0092	0.0369	0.0736	8.52
GRU Model with Adam (Testing)	0.9550	0.0059	0.0280	0.0623	7.45

In Table 6, the performance results obtained in both the training and testing phases are presented. Based on these performance results presented, the performance results of the proposed GRU model are at a satisfactory level.

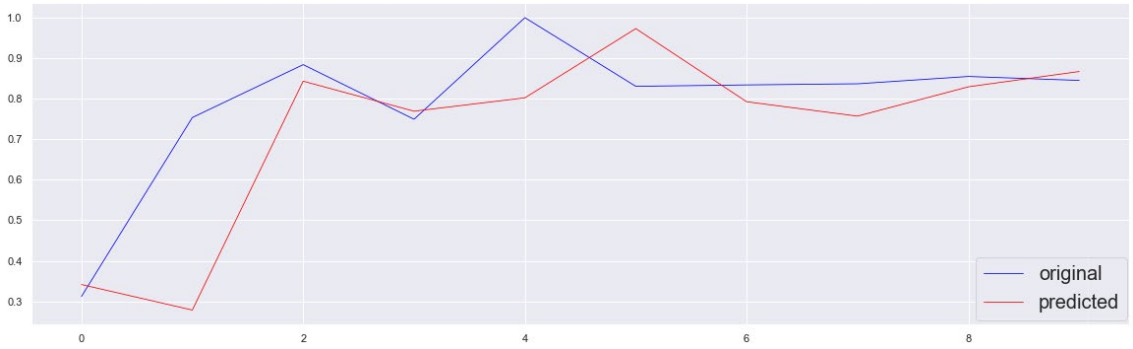

Figure 5. Proposed GRU model performance output

In the next stage of the study, the LSTM algorithm was carried out. The adjustment table used to perform the test operations with the LSTM algorithm is given in Table 7.

Table 7. Used GRU model parameters

Parameter	Value
Layers	1, 8, 10
Loss	Mean absolute error
Optimizer	Adam, Adamax, RMSprop
Epochs	25, 50, 100
Batch size	4, 16, 64
Activation	ReLU

In general, results similar to the results obtained from the GRU method were also obtained in the LSTM method. It was seen that the best R^2, MSE, RMSE, and MAPE performance values were obtained by the Adam optimization method in the test processes of the LSTM algorithm. In testing the proposed

LSTM method, steps 25, 50, and 100 were tested separately. To provide a comprehensive overview of these steps, the graphs obtained as a result of 100-step tests are presented. The best results were obtained in the 10-layer structure of the model, which was created from 1, 8, and 10-layer LSTM structures in series. In the proposed LSTM model, as in the GRU model, the performance measurement values of R^2, MSE, RMSE, and MAPE, which are widely used in the literature, were used to measure wheat productivity performance.

Table 8. Performance results of the LSTM model

Algorithm	R^2 Score	MSE	RMSE	MAE	MAPE
LSTM Model with Adam (Training)	0.9510	0.0080	0.0346	0.0722	7.80
LSTM Model with Adam (Testing)	0.9667	0.0054	0.0280	0.0614	7.33

In Table 8, both the training and test performance results obtained from the proposed LSTM method are presented. Based on these performance results presented, it can be stated that the performance results of the proposed LSTM model are at a satisfactory level.

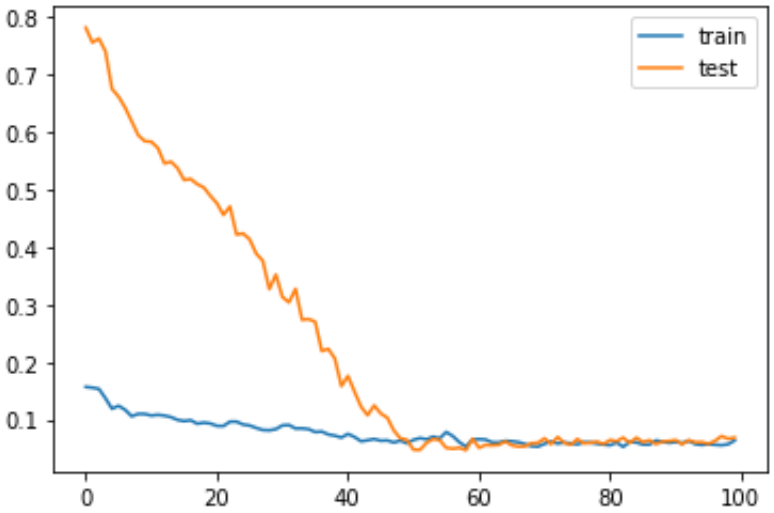

Figure 6. GRU model training and test loss graph

Training and test loss graphs are given in Figure 6 to support these results. Although there is a big difference between the training and test loss values at the beginning, after the 48th step, the training and test loss continues to be equalized. In Figure 7, the graph shows the actual and predicted values obtained as a result of the regression analysis of wheat productivity. Although the difference is high in the first stages of the test process, it is seen that the difference decreases towards the last stage of the iteration.

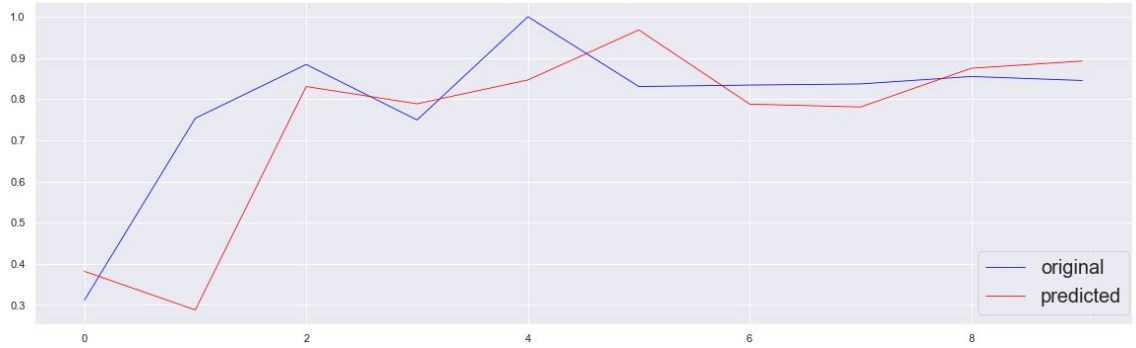

Figure 7. Proposed LSTM model performance output

The data set used in this article is original. It has not been used in any previous study in the literature. It has not been used in any previous study in the literature. For the stated reason, deep learning models focusing on different wheat crop yield estimations were preferred to compare the produced models more effectively and accurately. The results of the comparison obtained with these models are presented in Table 9. In Table 9, the comparison results performed on the wheat yield estimation data set in a different region than the Konya region, which constitutes the data set we used, are presented. R^2 and RMSE performance metrics obtained from Cao et al. [3]'s DNN (Deep Neural Network), CNN, and LSTM deep learning models are presented. The DNN model is a deep learning model consisting of feed-forward, fully interconnected neural networks. In Cao et al. [3]'s study, when the DNN, CNN, and LSTM models were evaluated within themselves, it was determined that the LSTM model was more successful than the CNN models consisting of 1-dimensional convolution, maximum pooling, and fully connected layers, and the five-layer DNN models with 528, 384, 128, 64 and 32 neurons. As the RMSE value decreases, the success rate increases. When the model comparison is made according to these two rules, it can be said that the study has academic innovation and effectiveness.

Table 9. Model comparisons in different wheat datasets

Algorithm	R^2 Score	MSE	RMSE	MAE	MAPE
GRU Model with Adam (Testing)	0.9550	0.0059	0.0280	0.0623	7.45
LSTM Model with Adam (Testing)	0.9667	0.0054	0.0280	0.0614	7.33
[3]'s DNN Model	0.85	-	742.49	-	-
[3]'s CNN Model	0.86	-	767.43	-	-
[3]'s LSTM Model	0.87	-	657.91	-	-
[30]'s LSTM Model	0.83	-	812.83	-	-

In another study on the estimation of wheat yield in the literature, Tian et al. [30] developed the LSTM model. With this model, artificial neural networks and SVM, which are machine learning methods, are compared for wheat yield estimation. As a result of the comparison processes, the artificial neural network reached a success rate of RMSE=812.83, 0.42 for R^2. The SVM method, on the other hand, reached a success rate of RMSE=867.70, 0.41 for R^2. On the other hand, in the LSTM model, RMSE = 812.83 gives a performance result of 0.83 for R^2. The results obtained as a result of the comparison information given show that the deep learning models proposed for wheat yield prediction

are reliable and promising. The results of the studies in the literature show that LSTM-based models can give more accurate results in yield estimations.

4. Conclusion and Discussion

To realize food production according to the increase in the world population, the crop to be produced must be planned in advance, at the right time, and reliably. In this study, two different deep learning models based on GRU and LSTM, which make productivity estimation on wheat data of Konya province, are proposed to meet the stated requirement. Since the two proposed deep learning models are based on RNN [41–44], both the performance and training times of the models were compared. As a result of the comparison processes, it is seen that the results of the LSTM model are slightly better than the GRU model. However, the training time of the GRU model was much shorter than the training time of the LSTM model. While the GRU model performs training modeling in 5.43 seconds, the LSTM model completes the training modeling in 6.62 seconds. In future studies, it is planned that more detailed yield analyses can be carried out in studies to be carried out using monthly or daily data obtained from all parameters in the data set.

References

[1] Vanli, Ö., Ustundag, B. B., Ahmad, I., Hernandez-Ochoa, I. M., & Hoogenboom, G. (2019). Using crop modeling to evaluate the impacts of climate change on wheat in southeastern turkey. *Environmental Science and Pollution Research*, *26*(28), 29397–29408. https://doi.org/10.1007/s11356-019-06061-6.

[2] Asseng, S., Cammarano, D., Basso, B., Chung, U., Alderman, P. D., Sonder, K., … Lobell, D. B. (2017). Hot spots of wheat yield decline with rising temperatures. *Global Change Biology*, *23*(6), 2464–2472. https://doi.org/https://doi.org/10.1111/gcb.13530.

[3] Cao, J., Zhang, Z., Luo, Y., Zhang, L., Zhang, J., Li, Z., & Tao, F. (2021). Wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine. *European Journal of Agronomy*, *123*, 126204. https://doi.org/https://doi.org/10.1016/j.eja.2020.126204.

[4] FAO, I. (2017). WFP (2015). The state of food insecurity in the World. Meeting the 2015 international hunger targets: taking stock of uneven progress. Rome, FAO.

[5] Dodds, F., & Bartram, J. (2016). *The water, food, energy, and climate Nexus: Challenges and an agenda for action*. Routledge.

[6] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/https://doi.org/10.1016/j.rse.2017.06.031.

[7] Vanli, Ö., Ahmad, I., & Ustundag, B. B. (2020). Area Estimation and Yield Forecasting of Wheat in Southeastern Turkey Using a Machine Learning Approach. *Journal of the Indian Society of Remote Sensing*, *48*(12), 1757–1766. https://doi.org/10.1007/s12524-020-01196-3.

[8] He, Z., Xia, X., & Zhang, Y. (2010). Breeding Noodle Wheat in China. In *Asian Noodles: Science, Technology, and Processing* (pp. 1–23). https://doi.org/10.1002/9780470634370.ch1.

[9] Chen, Y., Zhang, Z., Tao, F., Wang, P., & Wei, X. (2017). Spatio-temporal patterns of winter wheat yield potential and yield gap during the past three decades in North China. *Field Crops Research*, *206*, 11–20. https://doi.org/https://doi.org/10.1016/j.fcr.2017.02.012.

[10] Ahmad, I., Saeed, U., Fahad, M., Ullah, A., Habib ur Rahman, M., Ahmad, A., & Judge, J. (2018). Yield Forecasting of Spring Maize Using Remote Sensing and Crop Modeling in Faisalabad-Punjab Pakistan. *Journal of the Indian Society of Remote Sensing*, *46*(10), 1701–1711. https://doi.org/10.1007/s12524-018-0825-8.

[11] Nasim, W., Amin, A., Fahad, S., Awais, M., Khan, N., Mubeen, M., … Jamal, Y. (2018). Future risk assessment by estimating historical heat wave trends with projected heat accumulation using

SimCLIM climate model in Pakistan. *Atmospheric Research*, *205*, 118–133. https://doi.org/https://doi.org/10.1016/j.atmosres.2018.01.009.

[12] Ben-Asher, J., Yano, T., Aydın, M., & Garcia y Garcia, A. (2019). Enhanced Growth Rate and Reduced Water Demand of Crop Due to Climate Change in the Eastern Mediterranean Region (pp. 269–293). https://doi.org/10.1007/978-3-030-01036-2_13.

[13] Ahmad, I., Wajid, S. A., Ahmad, A., Cheema, M. J. M., & Judge, J. (2019). Optimizing irrigation and nitrogen requirements for maize through empirical modeling in semi-arid environment. *Environmental Science and Pollution Research*, *26*(2), 1227–1237. https://doi.org/10.1007/s11356-018-2772-x.

[14] Belhouchette, H., Nasim, W., Shahzada, T., Hussain, A., Therond, O., Fahad, E., … Wery, J. (2017). Economic and environmental impacts of introducing grain legumes in farming systems of Midi-Pyrenees region (France): a simulation approach.

[15] Dogan, H. G., & Karakas, G. (2018). The effect of climatic factors on wheat yield in Turkey: a panel DOLS approach. *Fresenius Environ Bull*, *27*, 4162–4168.

[16] Dudu, H., & Cakmak, E. H. (2018). Climate change and agriculture: an integrated approach to evaluate economy-wide effects for Turkey. *Climate and Development*, *10*(3), 275–288.

[17] TÜİK. (2021). TÜİK. Retrieved from https://data.tuik.gov.tr/.

[18] Cline, W. R. (2007). *Global warming and agriculture: End-of-century estimates by country*. Peterson Institute.

[19] Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., … Ciais, P. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences*, *114*(35), 9326–9331.

[20] Asseng, S., Ewert, F., Martre, P., Rötter, R. P., Lobell, D. B., Cammarano, D., … White, J. W. (2015). Rising temperatures reduce global wheat production. *Nature climate change*, *5*(2), 143–147.

[21] Ahmed, I., Ullah, A., Rahman, M. H. ur, Ahmad, B., Wajid, S. A., Ahmad, A., & Ahmed, S. (2019). Climate change impacts and adaptation strategies for agronomic crops. In *Climate change and agriculture* (pp. 1–14). IntechOpen London, UK.

[22] Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., … Kim, S.-H. (2016). Random Forests for Global and Regional Crop Yield Predictions. *PLOS ONE*, *11*(6), e0156571. Retrieved from https://doi.org/10.1371/journal.pone.0156571.

[23] Shahhosseini, M., Martinez-Feria, R., Hu, G., & Archontoulis, S. (2019). Maize yield and nitrate loss prediction with machine learning algorithms. *Environmental Research Letters*, *14*. https://doi.org/10.1088/1748-9326/ab5268.

[24] Jiang, D., Yang, X., Clinton, N., & Wang, N. (2004). An artificial neural network model for estimating crop yields using remotely sensed information. *International Journal of Remote Sensing*, *25*(9), 1723–1732. https://doi.org/10.1080/0143116031000150068.

[25] Khaki, S., Wang, L., & Archontoulis, S. V. (2020). A CNN-RNN Framework for Crop Yield Prediction. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.01750.

[26] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539.

[27] You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep gaussian process for crop yield prediction based on remote sensing data. In *Thirty-First AAAI conference on artificial intelligence*.

[28] Khaki, S., & Wang, L. (2019). Crop Yield Prediction Using Deep Neural Networks. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.00621.

[29] Kim, N., Ha, K.-J., Park, N.-W., Cho, J., Hong, S., & Lee, Y.-W. (2019). A Comparison Between Major Artificial Intelligence Models for Crop Yield Prediction: Case Study of the Midwestern United States, 2006–2015. *ISPRS International Journal of Geo-Information*, *8*, 240. https://doi.org/10.3390/ijgi8050240.

[30] Tian, H., Wang, P., Tansey, K., Zhang, J., Zhang, S., & Li, H. (2021). An LSTM neural network for improving wheat yield estimates by integrating remote sensing data and meteorological data in the Guanzhong Plain, PR China. *Agricultural and Forest Meteorology*, *310*, 108629.

https://doi.org/https://doi.org/10.1016/j.agrformet.2021.108629.
[31] Jayaraman, A. K., Murugappan, A., Trueman, T. E., & Cambria, E. (2021). Comment toxicity detection via a multichannel convolutional bidirectional gated recurrent unit. *Neurocomputing*, *441*, 272–278. https://doi.org/https://doi.org/10.1016/j.neucom.2021.02.023.
[32] Wang, J., Zhang, Y., Yu, L.-C., & Zhang, X. (2022). Contextual sentiment embeddings via bi-directional GRU language model. *Knowledge-Based Systems*, *235*, 107663. https://doi.org/https://doi.org/10.1016/j.knosys.2021.107663.
[33] Hu, L., Wang, C., Ye, Z., & Wang, S. (2021). Estimating gaseous pollutants from bus emissions: A hybrid model based on GRU and XGBoost. *Science of The Total Environment*, *783*, 146870.
[34] Chen, J. X., Jiang, D. M., & Zhang, Y. N. (2019). A Hierarchical Bidirectional GRU Model With Attention for EEG-Based Emotion Classification. *IEEE Access*, *7*, 118530–118540. https://doi.org/10.1109/ACCESS.2019.2936817.
[35] Aggarwal, C. C. (2018). *Neural Networks and Deep Learning*. *Neural Networks and Deep Learning*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-94463-0.
[36] Liu, G., & Guo, J. (2019). Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing*, *337*, 325–338. https://doi.org/https://doi.org/10.1016/j.neucom.2019.01.078.
[37] Pang, Z., Niu, F., & O'Neill, Z. (2020). Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons. *Renewable Energy*, *156*, 279–289. https://doi.org/https://doi.org/10.1016/j.renene.2020.04.042.
[38] Wang, J. Q., Du, Y., & Wang, J. (2020). LSTM based long-term energy consumption prediction with periodicity. *Energy*, *197*, 117197.
[39] Srinivasu, P. N., SivaSai, J. G., Ijaz, M. F., Bhoi, A. K., Kim, W., & Kang, J. J. (2021). Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM. *Sensors*. https://doi.org/10.3390/s21082852.
[40] Çetiner, H., & Çetiner, İ. (2021). Analysis of Different Regression Algorithms for the Estimate of Energy Consumption. *European Journal of Science and Technology*, (31), 23–33. https://doi.org/10.31590/ejosat.969539.
[41] ArunKumar, K. E., Kalaga, D. V, Kumar, C. M. S., Kawaji, M., & Brenza, T. M. (2022). Comparative analysis of Gated Recurrent Units (GRU), Long Short-Term Memory (LSTM) cells, Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA) for forecasting COVID-19 trends. *Alexandria Engineering Journal*.
[42] Ahmadzadeh, E., Kim, H., Jeong, O., Kim, N., & Moon, I. (2022). A Deep Bidirectional LSTM-GRU Network Model for Automated Ciphertext Classification. *IEEE Access*.
[43] Bhadouria, S. S., & Gupta, S. (2022). Combined LSTM GRU Model for Prediction of Congestion State in QUIC Protocol. In *Proceedings of International Conference on Computational Intelligence and Emerging Power System* (pp. 123–131). Springer.
[44] Li, W., Wu, H., Zhu, N., Jiang, Y., Tan, J., & Guo, Y. (2021). Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (GRU). *Information Processing in Agriculture*, *8*(1), 185–193.