J. Integr. Sci. Technol. 2024, 12(1), 710 Article

Journal of Integrated
SCIENCE & TECHNOLOGY

Integrated Science

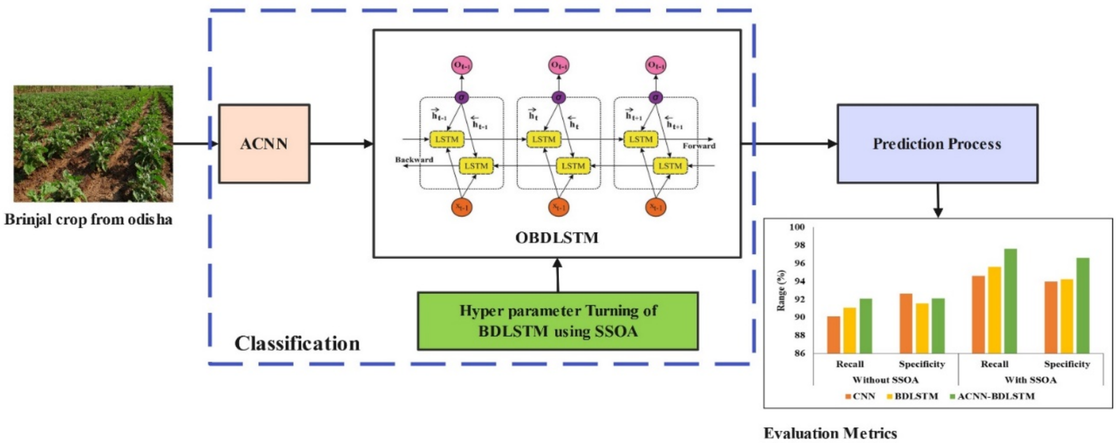
Brinjal crop yield prediction using shuffled shepherd optimization algorithm based ACNN-OBDLSTM model in smart agriculture

M Venkateswara Rao,[1] Y. Sreeraman,[2] Srihari Varma Mantena,[3] Venkateswarlu Gundu,[4] D. Roja,[5] Ramesh Vatambeti[6]*

[1]Department of CSE, NRI Institute of Technology, Vijayawada 521212, India. [2]Department of CSE, School of Technology, The Apollo University, Chittoor 517127, India. [3]Department of Computer Science and Engineering, Sagi Rama Krishnam Raju Engineering College, Bhimavaram 534204, India. [4]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India. [5]Department of CSE-Data Science, Chalapathi Institute of Technology, Guntur 522016, India. [6] School of Computer Science and Engineering, VIT-AP University, Vijayawada-522237, India.

Received on: 15-Jun-2023, Accepted and Published on: 21-Aug-2023

ABSTRACT



The need to ensure food security in the face of growing environmental concerns like climate change and natural catastrophes is raising the need for accurate crop output predictions. Predicting agricultural output is difficult because of the various non-linear interactions involved. So, instead of using traditional statistical tools, many researchers are turning to deep learning approaches to investigate these connections. Since brinjal is so important to the diets of Indians, protecting their ability to eat is of paramount importance. To this end, the attention-based convolution neural network with optimised bidirectional long short-term memory (ACNN-OBDLSTM) model was used to analyse and determine brinjal forecasts. The shuffling shepherd optimisation method (SSOA) is used for hyperparameter tuning of the BDLSTM model, which improves detection performance. The designed approach is applicable to lead brinjal producing states.

Keywords: Shuffled shepherd optimization algorithm; Bidirectional Attention-based convolution neural network; Crop yield forecasting; Brinjal Crop; Agriculture.

INTRODUCTION

Global warming and climatic variability in particular are key causes for alarm because of their potential negative effects on agricultural production in the future.[1] Because of the complex interplay between climate, weather, soil, fertiliser use, and seed type, this challenge necessitates the use of many datasets.[2] Predicting agricultural production typically requires using laborious statistical models.[3] Machine Learning and other cutting-edge data analysis techniques have become increasingly popular this decade because of the proliferation of "big data." Depending on the nature of the research challenge and the questions being asked, a machine learning model may be descriptive or predictive.[4] Predictive models, as opposed to descriptive ones, are used to generate predictions about the future based on the acquired data and explanations of the past.[5] In recent years, it has been put to use in fields as diverse as medicine, biology, finance, and even agriculture. Predicting agricultural yields, and therefore helping farmers decide what to plant and how to tend to it during the growing season, is a key use of machine learning.[6]

*Corresponding Author: Ramesh Vatambeti
Email: v2ramesh634@gmail.com
URN:NBN:sciencein.jist.2024.v12.710

India's agricultural sector is often considered the country's backbone. Ancient people were able to adjust to changing conditions since they farmed crops on their own land.[7] Consequently, many different animals and birds, as well as people, cultivate and consume natural crops. The creature's consumption of the earth's verdant produce has ensured its continued excellent health and well-being.[8] Since the introduction of cutting-edge technology and methods, agriculture has been in steady decline. As a result of these numerous innovations, people are increasingly investing their time and energy into the production of false commodities, or hybrid products, which contributes to an unhealthy way of life.[9] Most people nowadays don't know how crucial it is to plant crops there. These farming practises not only affect seasonal climatic conditions but also pose a threat to basic resources like soil, water, and air, leading to food insecurity.[10] We have evaluated these problems and obstacles, including the weather, temperature, and other factors, and have come to the conclusion that there is no technology or solution that can help us. Numerous strategies exist in India for fostering agricultural economic growth.[11] The quantity and quality of harvested crops may be increased and enhanced in a variety of ways. Predicting agricultural yields is another use for data mining. The overarching goal of data mining is to take raw data and use various analysis techniques to extract useful insights.[12]

The ability to examine data from several angles, classify it, and summarise the correlations discovered is made possible by data mining tools. Data mining is a technique for discovering patterns or connections across several fields in large relational databases.[13] Information may be gleaned from the patterns, correlations, and interactions present in all of this data. Data may be mined for insights into both the past and the future. For instance, crop loss may be identified and avoided in the future with the use of summaries of production data.[14] In agriculture, estimating future crop yields is a major challenge. How much of a harvest one can anticipate is a topic of constant interest among farmers. In the past, a farmer's familiarity with a given crop was taken into account when making yield predictions. Factors such as climate, animals, and harvest readiness all influence crop production.[15] Having reliable data on past crop yields is essential for agricultural risk management. As a result, this research suggests a means of calculating the yield of the police officer. The agricultural production per acre will be checked before planting begins.

These days, the data science paradigm is dominated by Deep Learning (DL) algorithms. Financial and economic time series forecasting are among its many uses.[16] There is a comparison between statistical models and machine learning methods. Artificial neural networks, generalised neural networks, long-short-term memory networks, recurrent neural networks, etc. are all popular DL methods. All of these methods learn the data's stochastic dependence and are thus data-driven nonparametric approaches. In order to estimate brinjal yield, this study employs an ACNN-OBDLSTM, an attention-based convolutional neural network with memory. The BDLSTM model's hyperparameters are tuned using the SSOA, which improves its detection performance.

RELATED WORK

Using data from the MARS Crop maintained by the European Commission's Joint Research Centre, Paudel et al.[17] have assessed the efficacy and understandability of neural network models for crop yield forecasting. Both the LSTM recurrent neural network and the one-dimensional (1DCNN) were chosen because of their ability to process either sequential or time-series input. A linear trend model and a GBDT model, both trained on human-created features, were evaluated in terms of performance. Experts in crop yield modelling and agronomy analysed the feature significance scores calculated from the input variables using feature attribution techniques. For soft wheat in Germany, LSTM models outperformed GBDT models, although both models performed similarly in all other experiments. The influence of severe temperature and moisture conditions on crop production was poorly recorded by LSTM models, although the effect of yield trend and static factors (such as elevation, soil water holding capacity, and biomass parameters) was effectively captured. Highlighting the necessity and difficulty of integrating human stakeholders in analysing model interpretability, our study demonstrates the potential of deep learning to identify characteristics and create trustworthy crop output estimates.

For the Rwandan district of Musanze, Kuradusenge et al.[18] utilise data mining techniques to forecast future agricultural harvests based on weather and yield historical data. This research uses machine learning methods to forecast agricultural yields in the face of varying weather conditions and disseminate this knowledge regarding production patterns. Ireland's potato and maize harvest totals and weather information were compiled from a number of sources. Random Forest, Polynomial Regression, and the Support Vector Regressor were used to analyse the data. Predictions of precipitation and temperature were made. Both trained and tested models were used. Results show that Random Forest is the most effective model, with R2 values of 0.875 and 0.817 for potato and maize, respectively, and root mean square errors of 510.8 and 129.9, respectively. Each crop's optimal weather conditions for maximum harvest were calculated. Based on the findings, Random Forest appears to be the best model for making accurate early-season crop output predictions. The results of this research will greatly help low- and medium-income nations like Rwanda rely more heavily on statistics when making choices about agriculture and climate change.

Jhajharia et al.[19] have used a number of machine learning methods to predict the harvest for five crops grown in Rajasthan. The findings show that at 0.963 R2, 0.035 RMSE, and 0.0251 MAE, the random forest approach outperformed the other algorithms used (Support Vector Machine, Gradient Descent, Long Short-Term Memory, and Lasso Regression). Cross-validation and error were used to verify the results. The purpose of this work is to assist farmers in increasing their crop yields by applying the crop selection approach.

The complete crop yield prediction system developed by Vignesh et al.[20] uses a mix of data mining and deep learning to bridge the gap between raw data and forecasted crop yields. The proposed research proposes replacing the modified chick swarm optimisation strategy with a Discrete Deep belief network using the VGG Net classification method to predict agricultural yield. Each layer of the Network received the data parameters in turn. The network architecture is used to build a prediction environment for crop production based on the input parameters. The best features of the input data are selected using the modified chick swarm optimisation method, and the resulting data is then fed into the classification procedure. Data classification and crop yield predictions are made using a Discrete Deep belief network trained with a Visual Geometry Group Net classifier. The projected model outperforms the state-of-the-art alternatives while keeping the baseline data distribution intact, allowing for 97% accurate predictions of crop yield.

Using previous information to describe dynamic temporal linkages and geographic topological structures, Qiao et al.[21] propose a novel information-guided Spatial-Temporal Attention Graph Network (KSTAGE). The proposed KSTAGE begins by embedding the first spectral feature using a 3D CNN. In order to create temporal attention weights automatically from a self-attention mechanism under the supervision of past information, KSTAGE relies on a brand-new information-guided Temporal Multi-Head Attention Algorithm (KTMA). Additionally, a novel method is devised for adjusting self-attention ratings so that they are consistent with the preceding distribution. In order to aggregate the spatial neighbourhood information for the ultimate yield prediction, a location-aware Spatial Attention Graph is presented, which makes use of geospatial knowledge. The proposed KSTAGE makes notable gains over the baselines, as evidenced by experimental findings (CONUS).

Ed-Daoudi et al.[22] investigated how climate, soil moisture, and precipitation may all be included in Machine Learning models to better predict agricultural yields. Traditional statistical methods for crop prediction are contrasted with the results of other algorithms, such as Decision Trees, Random Forests, and Neural Networks. The research showed that when it came to forecasting agricultural yields, Machine algorithms were superior to Statistical models. Mean squared errors ranged from 0.10 to 0.23, while those for statistical models were between 0.16 and 0.24, with corresponding values for the coefficient algorithm outperforming the other two Machine Learning algorithms with its 0.10 mean squared error and 0.90 R2 score. These findings point towards the possibility of enhanced food security and more efficient resource distribution for farmers in Morocco as a consequence of the use of Machine Learning algorithms to increase the accuracy of crop output projections.

Using DCNN and Radial Basis Feed Forward Neural Networks (RBFNN), S. Abisha et.al.[23] reported a novel approach for detecting and categorising contaminated brinjal leaves. We gathered 1100 photographs of brinjal leaf illness from five distinct species (Pseudomonas, Alternaria melongenea, Pythium aphanidermatum, and Tobacco Mosaic Virus), as well as 400 images from the same sector. To begin, a Gaussian filter is applied to the raw picture of the plant leaf to remove unwanted noise and boost the image's quality. Next, we use an EM-based segmentation technique to isolate the affected areas of the leaf. After that, the pictures are processed by the discrete Shearlet transform to pull out important details like texture, colour, and structure, which are then combined to create vectors. Finally, brinjal leaves are disease-type classified using DCNN and RBFNN. In terms of diagnosing leaf illnesses, the DCNN outperformed with a mean accuracy of 93.30% (with fusion), respectively.

Research Gap

The literature review shows that there are limitations to the ability to anticipate brinjal output with the currently available models.[23] The detection of brinjal leaves is the main work covered in recent studies.[24,25] This is what prompted the research summarised below to use deep learning to forecast brinjal harvests.

Designed System

Soil factors, including type, pH, fertility, and water-holding capacity, are fed into the crop selection model. Suitable crops for the land are selected using a combination of soil and expected weather criteria. Selecting crops can be done on a yearly basis or according to the seasons. The algorithm recommends one or more crops that will thrive in a given season, as well as other details about the crops, such as when they should be sown and how much water they will need, based on weather forecasts.

Materials and Methods

From AGMARKNET (https://agmarknet.gov.in/)[26], we have compiled the daily wholesale prices of Brinjal across seventeen marketplaces in Odisha, India, from January 1, 2015, to May 31, 2021. The Indian government's Directorate of Marketing and Inspection is responsible for maintaining the site. The "Agmark" programme is a proprietary application used by the agricultural produce marketplaces to enter data. The missing data was imputed using appropriate statistical methods before analysis.

In AGMARKNET,[26] we find statistical compilation of all available pricing information. This data shows that while the average price in Dhenkanal has remained high, in Bargarh it has fallen. During the research period, the highest price per quintal in Angul was Rs 7,500, while the lowest price was Rs 250. Most markets have a platykurtic distribution because kurtosis is high. The coefficient of variation (CV) for the price series reveals a spread from a low of 25.25% in Hinjilicut to a high of 95.41% in Athagarh Market. All market prices were found to be non-normal according to the Jarque-Bera normalcy test. Demand and supply may be gauged by looking at the price and availability of brinjal in different marketplaces across the state of Odisha. Potential exists for the cultivation of all tropical, subtropical, and temperate vegetable species in the area. The handling capability of any given commodity in a given market also plays a role in its entry into that market. It's a reflection of the region's commodity output as a whole. Farmers need to pay attention to these signals so that they can maximise the market value of the resources they produce. The biggest amount of brinjal is shipped to Bahadajholla (220.76 metric tonnes), followed by Angul (211.24 metric tonnes), Hinjilicut (188.95 metric tonnes), Jaleswar (117.34 metric tonnes), and Sarankul (103.17 metric tonnes). In addition to these markets, we also have statistics on the arrival of 3.70 and 3.92 metric tonnes of brinjal in the Khunthabandha and Boudh markets, respectively. It follows that the farmers of Odisha will be able to sell their produce in local farmers' markets such as Bahadajholla, Angul, Hinjilicut, and Jaleswar. A seasonal examination of brinjal arrivals in key markets in north-eastern India reveals lower arrivals in the market in the autumn and winter months of October and December. The number of brinjals available in the Angul market dropped to its lowest point in April.

Production Detection Using ACNN-OBDLSTM Model

The ACNN-OBDLSTM model is used to accurately identify brinjal production. For effective feature extraction and improved prediction accuracy, a new time series forecast network approach known as ACNN-OBDLSTM may be merged with a BLSTM network and a lightweight (ECA) constituent [27] to form a unified structure. By making full use of available data, the given method automatically learns and extracts local feature series, reducing the burden of models. The attention procedure has also developed, allowing for the extraction of more crucial characteristics. The basic structure of the BDLSTM method is exposed in Figure 1.

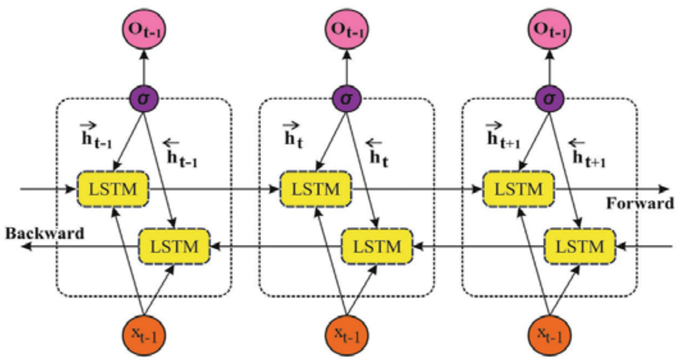


Figure 1: Architecture of BDLSTM model

Finally, the forecasting tasks were executed using the dense approach, which consists of several fully connected (FC) layers. In this instance, we used CNN to extract useful characteristics from the data. Alike to the standard neural network (NN) structure, CNN utilises local connections between neurons to reduce the overall number of connection layer parameters. In particular, it is the link between CNN's n-minus-one and n layers. The BDLSTM network was used to construct a more precise technique of forecasting; it works as a forward and backward LSTM network on all of the sequences that were trained. Both LSTM networks feed into the same output layer, which provides comprehensive context information at every sequence point.

There is much promise in the Channel Attention (CA) technique for improving the effectiveness of DCNNs. However, in order to achieve maximum efficiency, one of the offered ways is dedicated to constructing additional challenging attention components, which increases the complexity and computing weight of the procedure. A lightweight and trouble-free component called ECA was designed to prevent overfitting of the approach and reduce the calculation. The ECA was able to not only provide weights for each channel but also discover relationships between them. The time series data has been analysed by assigning more importance to the most important features and less importance to the least important features [28]. Therefore, ECA concentrates on the proper data, which increases the network's sensitivity to critical characteristics. Global Average Pooling (GAP) in channels is primarily the responsibility of the ECA. After that, ECA makes use of every channel and the k closest channels to them to record the local cross-channel connections. Fast 1D convolutions are performed by the ECA to add weight to the channel.

$\omega = \sigma(C1D_k(y))$ (1)

where C1D is a one-dimensional convolutional and k is the one-dimensional convolutional kernel size. In order to prevent the need for manual k adjustment, ECA uses a channel-dimensional adaptive mapping method to find an optimal value for k. The corresponding relationship was shown to be since the kernel size k of 1D unswervingly proportionate to the channel dimensionality C:

$C = \phi(k) = 2^{(\gamma * k - b)}$ (2)

So, to deliver the C, the kernel scope k is adjustably distinct as follows:

$k = \psi(C) = \left|\frac{log_2(C)}{\gamma} + \frac{b}{\gamma}\right|_{odd}$ (3)

where $| |_{odd}$ implies the adjacent odd sum. Both c and b are set in stone at their respective constants of [2, 1]. The high-dimensional channel has a shorter interface range than the low-dimensional channel does with non-linear mapping.

Hyperparameter Tuning for Bidirectional LSTM model

The SSOA technique is used to select the best possible value for the hyperparameter in the BDLSTM model.

Shuffled shepherd optimization procedure

In this subsection, we will primarily focus on expanding the use of the SSOA[29] meta-heuristic method. Each possible solution X_i with multiple variables (i.e. X_i= X_(i,j)) is a "sheep" in SSOA. Sheep are sorted into groups according to the values of the goal function. Select sheep are referred to as "shepherds," while those with superior goal function are referred to as "horses" in the herd. That's why each shepherd can have his or her own share of horses and sheep. The shepherd shifts positions by jumping onto one of the sheep or the horse in an effort to herd the sheep towards the horse. This is done for two reasons: (i) switching to a different, potentially better, member stimulates exploration, and (ii) switching to a different, potentially poorer member causes exploitation.[30-34] When the novel goal function is not worse than the previous one, the shepherd's site is updated, creating a hierarchy in the algorithm.

Following is a synopsis of the SSOA process.

1) The SSOA parameters $\alpha_0, \beta_0, \beta_{max}, iter_{max}, h, s$ are set. Where 'h' is the total number of flocks,'s' is the total sum of sheep, and 'iter max' is the maximum sum of iterations allowed.

2) The following equation (Eq. (4)) determines the beginning position in an m-dimensional search space[35]:

$X_i^0 = X_{min} + rand^\circ(X_{max} - X_{min}), i = 1,2,\ldots,n$ (4)

where X_i^0 is the initial key vector of the ith sheep, X_{max} and X_{min} are the limits of the design space, rand is a vector whose elements are in the range [0,1], and the sum of its machineries is equal to the sum of design variables, and n is the sum of sheep (n = h s). and the symbol ">" represents a multiplier that works element by element.

3) Each sheep's objective function is calculated, and the flock is then ranked from highest to lowest. Spread the sheep around so you can grow your herd. The first h sheep are chosen at random and distributed throughout the flocks (one sheep per flock). Then gather the second group of h sheep together. This procedure is repeated until each flock of sheep has a designated leader.

4) Pick out each individual sheep in a flock. The chosen sheep are called shepherds, while the best of the herd is referred to as horses. Choose at random a horse and a flock of sheep; the step size for each shepherd may be found by

$$Stepsize_i = \beta \times rand^\circ(X_d - X_i) + a \times rand^\circ(X_j - X_i) \quad (5)$$

where X_i, X_d, X_j in an m-dimensional search space; rand is a random vector, whose components are in the intermission [0,1] and whose number of mechanisms we know from the sum of components of solution vectors; and are computed using Eq. (6) and Eq. (7).

$$\propto = a_0 - \frac{a_0}{Iter_{max}} \times iteration \quad (6)$$

$$\beta = \beta_0 + \frac{\beta_{max} - \beta_0}{iter_{max}} \times iteration \quad (7)$$

Since the first selected sheep in a herd cannot have a better offspring than itself, the first period of the step size is zero, and the last picked sheep cannot have a worse offspring than itself, hence the size is also zero.

5) The temple each sheep calculates by the subsequent equation (Eq. (8)):

$$X_i^{temple} = X_i^{old} + stepsize_i \quad (8)$$

If ancient objective function, then the site of the sheep is changed, so we have $X_{new}^i = X_i^{temple}$, then the position of the shepherd is not changed and we have $X_{new}^i = X_i^{old}$. When the location of all sheep has been updated, the flocks should be combined for communication purposes.

Step 3 of the optimisation process is repeated up to the maximum allowed number of times, which is the termination criterion.

Results and Discussion

The testing was conducted on hardware comprising a PC running Linux with a Core i5 CPU and 8 Gigabytes of Random Access Memory. Here at Anaconda [30], we use Python 3.8 as our main operating system. The Pandas library was utilised for this purpose. Using the Pandas and Numpy tools, we performed data preprocessing to convert the raw data into matrix data. Using train_test_split from the scikit-learn module, we arbitrarily divided the data into a 90% training set and a 10% testing set. We take advantage of the fact that the scikit-learn package has all the necessary functions for implementing our machine learning algorithms. Accuracy and loss on training and validation data are shown in Figures 2 and 3, respectively.

Figure 2: Accuracy graph analysis of projected technique

Figure 3: Loss graph analysis of projected technique

Table 1: Results of the analysis of proposed technique under different measures

Techniques	Accuracy	Precision	Recall	Specificity	F-Score
Without SSOA					
CNN	90.10	92.64	91.39	92.80	92.01
BDLSTM	91.10	91.58	92.80	91.39	92.18
ACNN-BDLSTM	92.07	92.11	92.10	92.10	92.10
With SSOA					
CNN	94.60	93.95	92.33	93.88	94.64
BDLSTM	95.60	94.25	94.88	95.33	96.56
ACNN-BDLSTM	97.60	96.60	95.60	96.60	97.60

In the above Table 1, the Consequences of the analysis of the proposed technique under different measures are characterised. In this analysis without SSOA, the CNN model reached the accuracy rate of 90.1, the precision value of 92.64, the recall value of 91.39, the specificity of 92.80, and finally the F-score value of 92.01, respectively. Then the BDLSTM model reached the accuracy rate of 91.10, the precision value of 91.58, the recall value of 92.80, the specificity of 91.39, and finally the F-score value of 92.18, respectively. Formerly, the ACNN-BDLSTM model reached the accuracy rate of 92.07, the precision value of 92.11, the recall value of 92.10, the specificity of 92.10, and finally the F-score value of 92.10. And with SSOA analysis, the CNN model reached the accuracy rate of 94.60, the precision value of 93.95, the recall value of 92.33, the specificity of 93.88, and finally the F-score value of 94.64, respectively. At that moment, the BDLSTM model reached the accuracy rate of 95.60, the precision value of 94.25, the recall value of 94.88, the specificity of 95.33, and finally the F-score value of 96.56, respectively. At that point, the ACNN-BDLSTM model reached the accuracy rate of 97.60, the precision value of 96.60, the recall value of 95.60, the specificity of 96.60, and finally the F-score value of 97.60, respectively.

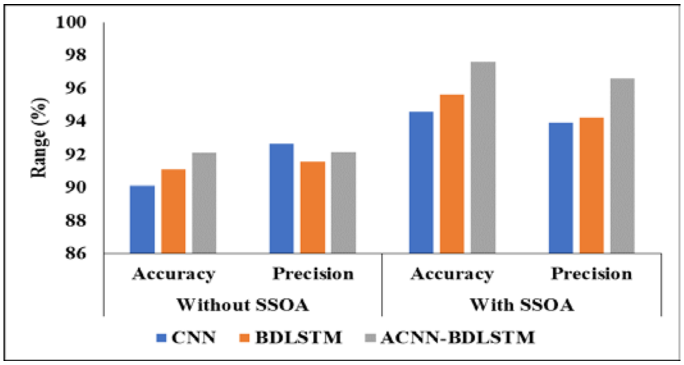


Figure 4: Validation of proposed model in terms of various metrics

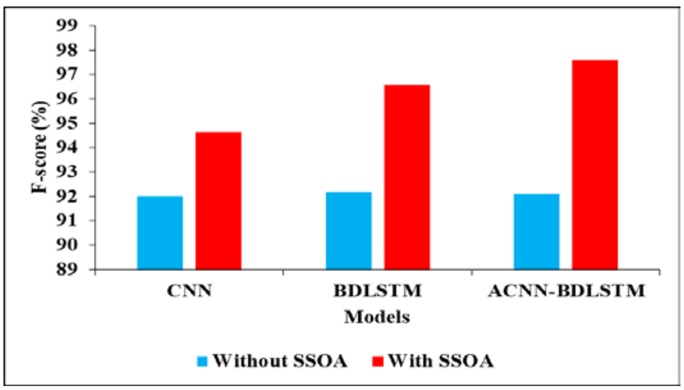


Figure 5: F-score analysis

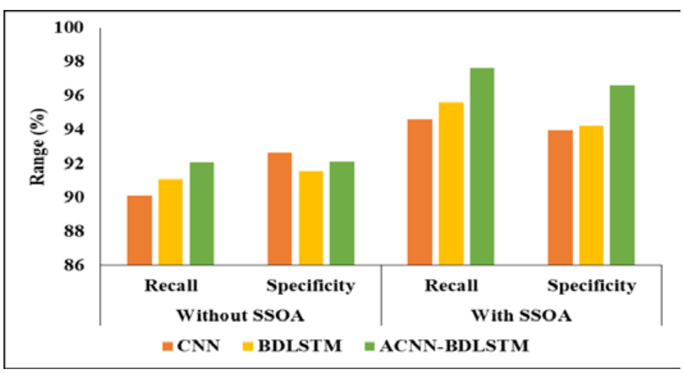


Figure 6: Comparison of SSOA model in Different Classifiers

In the Table 2, the Comparative analysis results of the projected technique are compared against existing approaches. In this investigation, we have used different methods to evaluate the presentation of the model in the DBN model, which reached an accuracy rate of 0.9101, a sensitivity value of 0.8923, and finally a Specificity value of 0.9042. After the AE model reached the accuracy rate of 0.9106, the sensitivity value of 0.8931, and finally the Specificity value of 0.9115, Then the CNN model reached the accuracy rate of 0.9272, the sensitivity value of 0.9128, and finally the Specificity value of 0.9016, respectively. Then the LSTM model reached an accuracy rate of 0.9380, a sensitivity value of 0.9320, and finally a Specificity value of 0.9476, respectively. In the BDLSTM model, the accuracy rate was 0.9579, the sensitivity value was 0.9493, and the Specificity value was 0.9658. And the ACNN-OBDLSTM model reached an accuracy rate of 0.9760, a sensitivity value of 0.9660, and finally a Specificity value of 0.9860, respectively.

Table 2: Comparative analysis results of proposed technique against existing approaches

Methods	Accuracy	Sensitivity	Specificity
DBN	0.9101	0.8923	0.9042
AE	0.9106	0.8931	0.9115
CNN	0.9272	0.9128	0.9016
LSTM	0.9380	0.9320	0.9476
BDLSTM	0.9579	0.9493	0.9658
ACNN-OBDLSTM	0.9760	0.9660	0.9860

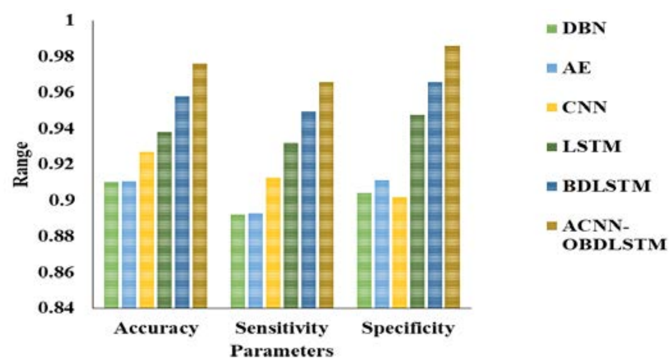


Figure 7: Graphical Comparison of proposed model with various DL classifiers

Conclusion

The use of deep learning techniques has been widely adopted because of their efficacy in a variety of fields, including agriculture. Predicting crop yields is important in many facets of the economy since it gives information on which to base decisions. However, most agricultural lands are still underdeveloped because ecosystem control methods have not been widely used. Using deep learning methods, the agricultural industry may anticipate the crop from a given data set and so avoid this issue. In this study, the hyper-parameter tuning of BDLSTM is optimised using the SSOA model, and then the ACNN-OBDLSTM is utilised to estimate the yield of brinjal. In this study, we studied market arrival and pricing data from 17 of Odisha's largest marketplaces. According to the seasonal index, brinjal is most expensive in the Dhenkanal market in July, then in October, and finally in November. Between Athagarh and Delhi, the market prices are at their lowest between September and December. If farmers know what the market will pay for their crops, they may take advantage of higher prices in surrounding markets and increase their profits. Limitations as accuracy of any machine learning model heavily depends on the quality and quantity of the data used for training and testing. If the dataset used for training is limited, noisy, or unrepresentative of real-world conditions, the model's predictions may not be reliable.

A future study shall focus on optimising the efficiency of the suggested algorithm through careful examination of the complete dataset. This method of predicting is not limited to use in the agricultural sector. The farmer's financial situation might also be taken into account in order to suggest the most lucrative crop.

Conflict of Interest statement

Authors declared no conflict of interest is there for this work.

References

1. S. Coulibaly, B. Kamsu-Foguem, D. Kamissoko, D. Traore. Deep learning for precision agriculture: A bibliometric analysis. *Intell. Syst. with Appl.* **2022**, 16, 200102.
2. M. Kuradusenge, E. Hitimana, D. Hanyurwimfura, et al. Crop Yield Prediction Using Machine Learning Models: Case of Irish Potato and Maize. *Agric.* **2023**, 13 (1), 225.
3. R. Ed-Daoudi, A. Alaoui, B. Ettaki, J. Zerouaoui. Improving Crop Yield Predictions in Morocco Using Machine Learning Algorithms. *J. Ecol. Eng.* **2023**, 24 (6), 392–400.
4. A. Oikonomidis, C. Catal, A. Kassahun. Deep learning for crop yield prediction: a systematic literature review. *New Zeal. J. Crop Hortic. Sci.* **2023**, 51 (1), 1–26.
5. A. Morales, F.J. Villalobos. Using machine learning for crop yield prediction in the past or the future. *Front. Plant Sci.* **2023**, 14, 1128388.
6. M.K. Dharani, R. Thamilselvan, P. Natesan, P.C.D. Kalaivaani, S. Santhoshkumar. Review on Crop Prediction Using Deep Learning Techniques. *J. Phys. Conf. Ser.* **2021**, 1767 (1), 12026.
7. M.S. Rao, A. Singh, N.V.S. Reddy, D.U. Acharya. Crop prediction using machine learning. In *Journal of Physics: Conference Series*; IEEE, **2022**; Vol. 2161, pp 926–932.
8. T. van Klompenburg, A. Kassahun, C. Catal. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* **2020**, 177, 105709.
9. S. Agarwal, S. Tarar. A hybrid approach for crop yield prediction using machine learning and deep learning algorithms. *J. Phys. Conf. Ser.* **2021**, 1714 (1), 12012.
10. J. Ranjani, V.K.G. Kalaiselvi, A. Sheela, D. Deepika Sree, G. Janaki. Crop Yield Prediction Using Machine Learning Algorithm. In *Proceedings of the 2021 4th International Conference on Computing and Communications Technologies, ICCCT 2021*; IEEE, **2021**; pp 611–616.
11. G. Lei, W. Zeng, J. Yu, J. Huang. A comparison of physical-based and machine learning modeling for soil salt dynamics in crop fields. *Agric. Water Manag.* **2023**, 277, 108115.
12. D. Elavarasan, P.M. Durairaj Vincent. Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications. *IEEE Access* **2020**, 8, 86886–86901.
13. R. Dhanapal, A. Ajanraj, S. Balavinayagapragathish, J. Balaji. Retraction: Crop price prediction using supervised machine learning algorithms. *J. Phys. Conf. Ser.* **2021**, 1916 (1), 12042.
14. P. Patil, V. Panpatil, S. Kokate. Crop Prediction System Using Machine Learning Algorithm. *J. Xidian Univ.* **2020**, 14 (6).
15. J. Pant, R.P. Pant, M. Kumar Singh, D. Pratap Singh, H. Pant. Analysis of agricultural crop yield prediction using statistical techniques of machine learning. *Mater. Today Proc.* **2021**, 46, 10922–10926.
16. S.M. Pande, P.K. Ramesh, A. Anmol, et al. Crop Recommender System Using Machine Learning Approach. In *Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021*; IEEE, **2021**; pp 1066–1071.
17. D. Paudel, A. de Wit, H. Boogaard, et al. Interpretability of deep learning models for crop yield forecasting. *Comput. Electron. Agric.* **2023**, 206, 107663.
18. M. Kuradusenge, E. Hitimana, D. Hanyurwimfura, et al. Crop Yield Prediction Using Machine Learning Models: Case of Irish Potato and Maize. *Agric.* **2023**, 13 (1), 225.
19. S. Iniyan, V. Akhil Varma, C. Teja Naidu. Crop yield prediction using machine learning techniques. *Adv. Eng. Softw.* **2023**, 175, 406–417.
20. K. Vignesh, A. Askarunisa, A.M. Abirami. Optimized Deep Learning Methods for Crop Yield Prediction. *Comput. Syst. Sci. Eng.* **2023**, 44 (2), 1051–1067.
21. M. Qiao, X. He, X. Cheng, et al. KSTAGE: A knowledge-guided spatial-temporal attention graph learning network for crop yield prediction. *Inf. Sci. (Ny).* **2023**, 619, 19–37.
22. R. Ed-Daoudi, A. Alaoui, B. Ettaki, J. Zerouaoui. Improving Crop Yield Predictions in Morocco Using Machine Learning Algorithms. *J. Ecol. Eng.* **2023**, 24 (6), 392–400.
23. S. Abisha, A.M. Mutawa, M. Murugappan, S. Krishnan. Brinjal leaf diseases detection based on discrete Shearlet transform and Deep Convolutional Neural Network. *PLoS One* **2023**, 18 (4 April), 284021.
24. S.F. Syed, G. Phonsa. Machine Learning Approaches for Brinjal Leaf Disease Detection. In *2nd International Conference on Sustainable Computing and Data Communication Systems, ICSCDS 2023 - Proceedings*; IEEE, **2023**; pp 226–234.
25. S. Natarajan, V. Ponnusamy. Classification of Organic and Conventional Vegetables Using Machine Learning: A Case Study of Brinjal, Chili and Tomato. *Foods* **2023**, 12 (6), 1168.
26. R.K. Paul, M. Yeasin, P. Kumar, et al. Machine learning techniques for forecasting agricultural prices: A case of brinjal in Odisha, India. *PLoS One* **2022**, 17 (7 July), 270553.
27. A.S. Almasoud. Intelligent Deep Learning Enabled Wild Forest Fire Detection System. *Comput. Syst. Sci. Eng.* **2023**, 44 (2), 1485–1498.
28. Y. Chen, R. Fang, T. Liang, et al. Stock Price Forecast Based on CNN-BiLSTM-ECA Model. *Sci. Program.* **2021**, 2021 (1), 1–20.
29. A. Kaveh, A. Zaerreza. Shuffled Shepherd Optimization Method: A New Meta-Heuristic Algorithm. In *Studies in Systems, Decision and Control*; Springer, Cham, **2023**; Vol. 463, pp 11–52.
30. A. Kaveh, M. Kamalinejad, K. Biabani Hamedani. Enhanced versions of the shuffled shepherd optimization algorithm for the optimal design of skeletal structures. *Structures* **2021**, 29, 1463–1495.
31. L. Raamesh, S. Radhika, S. Jothi. Generating Optimal Test Case Generation Using Shuffled Shepherd Flamingo Search Model. *Neural Process. Lett.* **2022**, 54 (6), 5393–5413.
32. T. Daniya, V. Srinivasan. Shuffled shepherd social optimization based deep learning for rice leaf disease classification and severity percentage prediction. *Concurr. Comput. Pract. Exp.* **2023**, 35 (4).
33. B.K. Pandey, V. M. R. M. , S. Ahmad, C. Rodriguez, D. Esenarro. ExpSSOA-Deep maxout: Exponential Shuffled shepherd optimization based Deep maxout network for intrusion detection using big data in cloud computing framework. *Comput. Secur.* **2023**, 124 (102975).
34. M. Jayabalan, E. Srinivas, F.H. Shajin, P. Rajesh. On Reducing Test Data Volume for Circular Scan Architecture Using Modified Shuffled Shepherd Optimization. *J. Electron. Test. Theory Appl.* **2021**, 37 (5–6), 577–592.
35. K.S. Bhanumathi, D. Jayadevappa, S. Tunga. Feedback Artificial Shuffled Shepherd Optimization-Based Deep Maxout Network for Human Emotion Recognition Using EEG Signals. *Int. J. Telemed. Appl.* **2022**, 2022.