RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF
MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt24616

Predictive Modeling of Crop Yield Using Deep Learning Based Transformer with Climate Change Effects

S. Yash Pravesh [a], Nakshatra Garg [a], Ravik Arora [a], Sudhanshu Singh [a], S. Siva Sankari [a,*]

[a] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore-632014, Tamil Nadu, India
* Corresponding Author Email: sivasankari.s@vit.ac.in
DOI: https://doi.org/10.54392/irjmt24616
Received: 19-10-2023; Revised: 04-11-2024; Accepted: 18-11-2024; Published: 30-11-2024

Abstract: Climate change is a significant global challenge concerning agriculture and food security. The understanding of climate change effects on crop production is necessary for developing an effective adaptation strategies and predicting a crop yield accurately. This paper suggests the combined Clustering Long Short Term Memory Transformer (CLSTMT) model for crop yield prediction. CLSTMT is a hybrid model that integrates clustering, deep learning based LSTM and Transformer techniques. The outliers from the historical crop and climate data are removed using k-means clustering. Followed by, the crop yield is predicted using Transformer-based neural network with LSTM layers and feed-forward neural network (FNN) components. The model design effectively captures climate-influenced patterns, enhances the precision and comprehensiveness of crop yield prediction. The experiment is conducted using the dataset with crop yield, climate, and pesticide details over 101 countries collected from 1990 to 2013. The comparative analysis reveals that the CLSTMT model outperforms other regression models such as SGDRegressor (SGDR), Lasso Regression (LR), Support Vector Regression (SVR), ElasticNet (EN) and Ridge Regression (RR). The proposed design effectively captures climate-influenced patterns, enhancing the precision and comprehensiveness of crop yield predictions. The findings indicate that the proposed model provides an accurate prediction of crop yield with high R^2 of 0.951 and lesser Mean Absolute Percentage Error (MAPE) of 0.195. This value suggests a minimal average percentage deviation between the actual and predicted yields. The findings indicate that the CLSTMT model provides more accurate crop yield prediction compared to others.

Keywords: Agriculture, Attention Mechanism, Crop Yield Prediction, Deep Learning, Encoders, Sustainable Agricultural Practices.

1. Introduction

Now a days, the crop yield prediction confer a significant benefit to policymakers and farmers. It emphasizes the necessity of customized adaptive strategies to tackle regional disparities in crop reactions to climate change [1]. The issue of climate change presents a substantial and difficult challenge to agricultural systems on a global scale, and it also represents a notable and concerning threat to the security of food supplies [2-4]. Therefore, there exists an urgent requirement for accurate and dependable forecasts pertaining to the consequence of change in climate on crop production. The ongoing climate changes are crucial to comprehend the complex correlation between climatic factors and agricultural productivity in various geographical areas [5]. This understanding is essential for developing successful adaptation strategies [6, 7]. The incorporation of deep learning techniques, specifically Transformer Encoders, into the modeling procedure offers a robust mechanism for capturing intricate patterns within the dataset, leading to enhanced precision and comprehensiveness in predicting crop yields. By utilizing the computational power of Transformer Encoders to effectively handle lengthy sequences and integrate attention mechanisms, the model is capable of identifying complex patterns that are influenced by changes in climate, thereby improving its predictive abilities [8]. The objective of this investigation is to provide the valuable insights related to intricate relationship between climate change and the crop production. This research examines the influence of various climatic factors, including temperature and precipitation, on crop productivity, with a particular focus on the adverse effects of higher temperatures and reduced precipitation levels. In this research, the dataset undergoes preprocessing steps including handling null values, removing duplicates, and converting categorical variables using one-hot encoding. After that clustering is performed to eliminate outliers. Then the dependent and independent variables are identified. A hybrid deep learning model, combining an LSTM and a Transformer encoder, is then employed. Integrating hybrid deep learning models into the modeling process provides a robust approach for capturing complex patterns in the dataset, resulting in improved accuracy and depth in crop yield predictions. To ensure the sustainability of agricultural practices among the climate change challenges, it is imperative to implement tailored adaptation strategies that are suitable for specific local circumstances.

The findings provide a vital basis for the development of proactive strategies aimed at mitigating the detrimental impacts of climate change on agricultural productivity [9, 10]. The need for comprehensive strategies that integrate climate science, agricultural practices, and policy frameworks is evident due to the challenges posed by climate change [11, 12]. This study provides evidence-based insights that can be utilized by policymakers and farmers to develop adaptive strategies that are specifically tailored to different regions and crops. By doing so, these strategies can help to enhance agricultural resilience and ensure the security of global food supplies [13]. The present study demonstrates the capability of advanced machine learning (ML) techniques, particularly deep learning (DL) with Transformer Encoders, in analyzing the complex relationship between climate change and crop production. The integration of climate science, data analytics and agricultural expertise presents a viable approach to ensure food security and reducing the effects of climate changes on susceptible communities. The global acknowledgement of climate change is imperative to prioritize data-driven and proactive approaches informed by research findings.

The remaining sections are structured as follows: Section 2 offers the review of previous research endeavors and methodologies adopted to tackle the complexities of forecasting crop yields amidst climate change and agricultural productivity challenges. Section 3 discusses CLSTMT prediction model architecture and the evaluation metrics employed. Section 4 describes the dataset, discusses the study's findings, including a detailed examination of the results obtained. Finally, Section 5 concludes the study and outlines the potential future research directions.

2. Related Works

This section explores the related research work undertaken by many researchers in the field of crop yield prediction. The primary concern pertaining to food security resides in the ramifications of climate change, specifically with regards to the impact on agricultural productivity [14]. To address this issue, researchers are employing deep learning techniques to examine extensive datasets and generate forecasts regarding the consequences of changes in climate on the agricultural sector. In recent years, there has been a growing acceptance and adoption of this particular approach, which holds promise for improving our understanding and preparedness for the future consequences of climate change on worldwide food systems. In [15], the researcher utilized publicly available crop yield data acquired from the Indian government to ingress the model's effectiveness. The author suggests that in order to make progress in the field of study, it is necessary to improve the existing model by incorporating supplementary data, exploring different regression techniques, and considering alternative evaluation metrics. Chandraprabha and Dhanaraj [16] presented the study that investigates the utilization of data mining and machine learning methodologies on agricultural data analysis. Crop production forecasting utilizes various algorithms such as k-Nearest Neighbors (KNN), Bayesian Networks, Support Vector Machines (SVM) and K-Means Clustering. Significant advancements have been achieved in the pursuit of attaining elevated levels of precision and obtaining favorable results, thereby facilitating enhanced decision-making capabilities for agricultural practitioners and ultimately leading to improved crop productivity.

Shakoor *et al.,* [17] improved the agricultural productivity by employing an intelligent information prediction analysis within the domain of farming. The research included two algorithms, specifically Decision Tree Learning- K-Nearest Neighbors Regression (KNNR) and ID3 (Iterative Dichotomiser3) to generate forecasts on agricultural yield rates. The study's results revealed that the decision tree, particularly the ID3 algorithm, had higher performance in comparison to the KNNR technique in most cases. This phenomenon was clearly demonstrated by its capacity to produce more accurate forecasts while simultaneously reducing the occurrence of errors. Reddy *et al.,* [18] implemented a smart irrigation system that utilizes ML algorithms and Internet of Things (IoT). The structure employs temperature, humidity, and moisture sensors and utilizes the Decision Tree algorithm to anticipate the crop water requirements. The system facilitates the transmission of email notifications to farmers, thereby assisting in the proactive management of water resources and the optimization of water utilization. The aforementioned technology possesses the capacity to greatly augment agricultural productivity while simultaneously preserving water resources. Additional empirical research is required to substantiate its efficacy across diverse agricultural contexts.

Agarwal and Tarar [19] introduced an innovative methodology for forecasting agricultural yields, which involves the incorporation of DL and ML algorithms. The model incorporates Support Vector Machines (SVM), Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) algorithms. By examining numerous variables including temperature, precipitation, pH levels, relative humidity, and land area, the model possesses the capacity to forecast the optimal crop for a given set of conditions. This phenomenon leads to improved precision in comparison to existing models that are presently accessible. The system provides farmers with valuable information, enabling them to make data-driven decisions and optimize profits while minimizing expenses. The consideration of artificial intelligence (AI) techniques in the agricultural sector demonstrates considerable potential in enhancing and streamlining productivity within the farming sector. Elavarasan and Vincent [20] presented a comprehensive framework based on deep reinforcement learning for the motive of crop yield prediction in sustainable agriculture. The study highlights the significance of predicting the crop yields in sustainable agriculture. The application of the deep reinforcement learning model is employed to address the challenges posed by unexpected weather patterns, fluctuations in soil composition and other variables that influence crop output. The study results offer substantiation for the effectiveness of the suggested model in properly predicting crop production, thus indicating its potential usefulness in the domain of intelligent agriculture. This study makes a valuable contribution to the progress of agrarian practices and establishes the groundwork for the development of more sustainable and efficient strategies for crop management.

Palanivel and Surianarayanan [21] predicted agricultural productivity using machine learning algorithms and large-scale datasets. The study highlights the importance of precise yield prediction within the agricultural industry, particularly in relation to the challenges posed by water scarcity and unpredictable weather conditions. This research investigates a range of artificial neural networks and machine learning algorithms, such as Stochastic Gradient Descent regressor (SGDRegressor) and support vector machines, for predicting crop yield. The methodology involves the collection and preparation of data related to soil fertility, weather patterns, and various other factors that exert influence on crop productivity. According to the authors, both SVM and artificial neural network (ANN) exhibit favorable outcomes in the prediction of crop yield. In addition, the authors suggest conducting a research study to examine the effects of employing big data methodologies at various stages of the prediction process in order to improve accuracy. This study significantly contributes to the domain of smart farming by offering valuable assistance to farmers in their decision-making procedures, ultimately resulting in improved productivity and economic advancement.

Durai and Shamili [22] utilized the deep learning and machine learning methodologies within the scope of intelligent agriculture. The research highlights the importance of employing cutting-edge technologies to improve agricultural methodologies. By utilizing these methodologies, agricultural practitioners can leverage data analysis to enhance crop productivity, reduce resource consumption, and enhance overall operational effectiveness. The results demonstrate how deep learning and machine learning models can transform traditional agriculture through the adoption of a smarter and more environmentally sustainable approach. This study contributes to the development of intelligent farming solutions, enhancing the efficiency and productivity of the agricultural industry. Cedric *et al.* [23] presented a comprehensive analysis of prediction of crop yield using ML algorithms. The study emphasizes the potential of these models in effectively predicting crop yields, a factor that can have a substantial influence on agricultural planning and food security within the area. The results indicate that the utilization of machine learning methodologies exhibits potential in offering significant contributions to the understanding of crop productivity. This has implications for policymakers and farmers, allowing them to make informed decisions to improve agricultural practices. Further research and increased collaboration among stakeholders are necessary to improve and expand the capabilities of these models. These endeavors will greatly contribute to the sustainable development of agriculture in West Africa.

Ang and Seng [24] highlighted the benefits of employing hyperspectral and multispectral information processing systems and technologies in order to improve agricultural productivity and methods, thereby providing valuable insights to farmers and crop managers. These technologies have been widely utilized in various agricultural applications such as crop yield prediction, crop management, crop disease detection, monitoring of land utilization, soil conditions, and water resources. However, the integration of hyperspectral data in the field of agriculture presents notable obstacles concerning Big Data, as a result of the considerable volume of spatial and spectral data that is encompassed. Kalimuthu *et al*. [25] contributed to the field of precision agriculture by addressing the challenges faced by farmers due to unpredictable climate variations. The integration of machine learning techniques within the proposed crop prediction system offers a potentially effective strategy for accurately predicting forthcoming crop yields. This approach considers pertinent environmental and soil factors to enhance the precision of the forecasts. The combination of a naive Bayes Gaussian classifier and the boosting algorithm has demonstrated its efficacy as a dependable approach for attaining a high level of accuracy in crop prediction.

Deforce *et al.* [26] suggested to employ Temporal Fusion Transformers (TFTs) as a means of predicting soil moisture levels in the context of smart agriculture. TFTs represent advanced time-series forecasting models that possess the ability to incorporate static features, such as soil characteristics, as well as future information, alongside historical data. The architecture of thin-film transistors (TFTs) includes variable selection networks, static feature encoders, LSTM encoders and decoders, gating mechanisms, and temporal multi-head attention. These elements enable the capture of intricate interrelationships among various input features. Nagini *et al*., [27] developed a model to predict agricultural yield for different crops in the states of Andhra Pradesh and Telangana through the utilization of exploratory data analysis and diverse predictive models. The analysis considered various factors such as water availability, nitrogen levels, weather patterns, soil properties, crop rotation practices, soil moisture content, surface temperature, and rainfall. The study's findings demonstrate the effectiveness of different predictive models, including Linear, Multiple Linear, and Non-SGDRegressor models. The aforementioned models demonstrated the capability to generate precise predictions regarding crop yield by utilizing the provided input parameters. The formulas pertaining to two and three predictors have been identified as valuable tools for predicting crop production within the field of Agriculture.

Junankar *et al.* [28] presented the Temporal Fusion Transformers (TFT) as an innovative methodology for predicting time series data across different time intervals. TFT is characterized by the combination of advanced predictive modeling techniques and the capability to interpret temporal patterns. The model incorporates a gated mechanism, variable selection networks, static covariate encoders, and interpretable multi-head attention to achieve superior performance on real-world datasets, surpassing current methods. Numerous scholarly investigations have been conducted to explore diverse Deep Learning and Machine Learning models in the context of predicting crop yield. The temporal aspect of this problem remains highly significant. Temporal Fusion Transformers (TFTs) have been introduced to address the existing gap in the agricultural sector regarding decision-making for farmers. TFTs provide a reliable and easily understandable approach to predict crop yield, thereby enhancing performance in this domain. This contribution is considered valuable. Ang et al. [29] designed an improved oil palm yield prediction model through the utilization of multiple sources of data, the application of feature selection techniques, and the implementation of advanced DL and ML algorithms. The utilization of feature selection methodologies was employed to ascertain the most significant predictors, thereby enhancing the precision of crop yield forecasting. The incorporation of climatic variables acquired from satellites, such as CHIRPS-derived rainfall data and Landsat-derived Land Surface Temperature (LST) data, significantly improved the precision of yield forecasting.

Goel *et al.* [30] utilized the VGG19 model and logistic regression as the classifier. They achieved impressive results marked by a high level of accuracy. The previously mentioned combination exhibited superior performance compared to alternative classifiers, such as neural network, decision tree, random forest, SVM, and AdaBoost. The utilization of the VGG19 convolutional neural network (CNN) in conjunction with Logistic Regression yielded the most favorable results across all evaluation metrics. Gurrapu *et al.* [31] presented a novel conceptual framework known as DeepAg, which integrates DL and econometrics methodologies to ingress the consequences of outlier events on agricultural production. The objective of the authors is to forecast commodity production by incorporating widely utilized financial indices, such as the Dow Jones, with agricultural products such as Milk and Cheese. In order to attain precise predictions, researchers employ Isolation Forests and Long Short-Term Memory (LSTM) networks for the motive of identifying outliers. El Hachimi *et al.* [32] offered significant contributions by offering valuable insights into the utilization of DL and ML methodologies in precision agriculture. This study establishes a foundation for improved resource management and climate change mitigation in the agricultural sector by offering accurate crop recommendations and weather forecasts. The potential implementation of the FLA7A (decision support system) platform would represent a significant progress in addressing the intricate challenges stemming from the growing global population and the increasing imperative to guarantee food security. The authors intend to employ the weather forecasting model in subsequent research endeavors with the aim of augmenting the decision support platform through the estimation of crop yield.

Parasuraman *et al.* [33] designed a novel framework that utilizes machine vision technologies, drone data, and Internet of Things (IoT) for making informed decisions regarding crop cultivation by farmers. The framework achieves a crop detection accuracy rate of 99.96%.This paper highlights the significance of utilizing advanced technologies in order to optimize agricultural practices and enhance crop yield, all while mitigating environmental consequences and reducing reliance on detrimental pesticides. According to the research findings, the proposed framework provides the positive implications for the agriculture sector and facilitate the adoption of sustainable farming practices. The accuracy of deep learning methods can be improved with feature selection and optimization techniques [34, 35]. The prediction of plant diseases also will help an effective food management and food security [36]. Most of the studies have made a region specific study on precision agriculture and used datasets specific to a region. The various algorithms mentioned have not tried to build advanced machine learning techniques or deep learning techniques. The research done does not emphasizes the importance of tailored adaptive strategies to address regional disparities in crop responses to climate change.

3. Methodology

The objective of the paper is to perform predictive analysis on crop yield by utilizing the pesticide and climate features along with crop yield feature. The procedural steps necessary for implementing the model involve the process of acquiring the statistical summary, unique values and dataset information. The data undergoes a sequence of preprocessing steps, which encompass data cleaning, null value validation and normalization. Then the outlier is eliminated by grouping similar samples and finding the outlier samples using k-means clustering. Kmeans is also utilized to uncover underlying patterns in the dataset, with a special focus on traits that are predicted to have a major impact on crop yield. Then the columns pertaining to area and item have been subjected to the process of one-hot encoding. After that, the integration of model and dataset is achieved by defining the configuration of the neural network architecture. Subsequently, the model undergoes training using the specified training dataset, followed by validation using the validation dataset. After that the test data is used to generate predictions. Several evaluation metrics are calculated. Following this, a comprehensive examination is conducted utilizing diverse models, subsequently accompanied by the visualization of data in order to forecast crop yield based on empirical outcomes.

3.1 CLSTMT Architecture

The CLSTMT model is a Transformer-based neural network that incorporates LSTM layers and feed-forward neural network (FNN) components in its architecture. The model is designed to accept input data that represents the characteristics of the training dataset. An Embedding layer converts the 2D input to a 3D representation. The key component of the design is the Transformer Encoder block, which has two major components: the Attention mechanism and the Feed Forward (FF) section. The FF component analyses and integrates this information back into the network while the attention mechanism assists the model in focusing on pertinent information from the input sequences helping in processing the data. An embedding layer transforms the 2D input into a 3D representation at the base of the design. The embeddings are stored in the embedded inputs variable. The input data is processed in a succession of Transformer Encoder blocks after the first LSTM layers, with each block being in charge of processing the data in a self-attention and feed-forward fashion. The hyper parameter controls the quantity of Transformer Encoder blocks. The output of the Transformer Encoder blocks is then globally averaged and the data is processed and transformed further using multiple Dense layers with ReLU activation. In order to prevent overfitting during training, dropout layers are used. Adam optimizer with a learning rate of 0.001 is utilized as an optimizer and mean absolute error (MAE) is utilized as a loss function for building the model. Figure 1 depicts the proposed CLSTMT prediction model architecture.

The architectural design addresses the regression task by synergistically integrating the capabilities of Transformers for sequence manipulation and attention mechanisms with LSTM layers and feed-forward neural networks. To enhance the model's performance with respect to specific datasets and workloads, it is possible to modify the hyper parameters of the model. In CLSTMT model, the manual tuning of the parameters are done. Finally, the design of the transformer model is set as follows. Head size as 64, number of heads as 2, feed forward dimension as 64, number of transformer blocks as 4, MLP units as [128, 64] and dropout rate as 0. The design of LSTM layer is set as follows. LSTM units as 128, Optimizer as Adam with a learning rate of 0.001, losss function as mean absolute error, epochs as 20 and batch size as 64. The data exploration process begins by importing necessary libraries and loading the dataset using Python libraries such as pandas, numpy, matplotlib, and seaborn. The next step in data wrangling includes identifying and addressing missing values, performing data cleaning, and removing duplicates. Various data visualizations, such as histograms, line plots, bar plots, and scatter plots, are used to analyze relationships between variables and identify trends and patterns in crop-related factors. Different regression models, including CLSTMT, SGDRegressor, Ridge Regression, Lasso Regression, Support Vector Regression (SVR), and Elastic Net are designed for predicting the crop yield. Each model performance is evaluated using R^2 and MAPE.

This study investigates the integration of a clustering and Transformer model with LSTM layers as a novel approach to improve predictive accuracy. The Transformer architecture is utilized, which includes embedding, multi-head attention, and feed-forward layers. The model undergoes training, evaluation, and its performance is compared to that of traditional regression models. This study presents a comprehensive analysis of the results, including scatter plots that demonstrates the alignment between actual and predicted crop yields. The evaluation of the model demonstrates their efficacy in forecasting crop yields. The CLSTMT model starts with preprocessing of the input data, followed by the clustering and removal of the outliers. Then it is applied to an input layer, followed by an embedding layer that converts the data into dense vectors. These vectors are processed by a Long Short-Term Memory (LSTM) block, capturing temporal dependencies through multiple LSTM layers. A Transformer encoder block is also applied, incorporating layer normalization, multi-head attention, dropout, residual connections, and feed-forward networks to capture complex dependencies within the data. The outputs from the LSTM and Transformer blocks are merged using a concatenate layer. This combined output is then passed through fully connected layers (MLP) with ReLU activation and dropout layers for regularization. The final output layer generates the predicted crop yield.

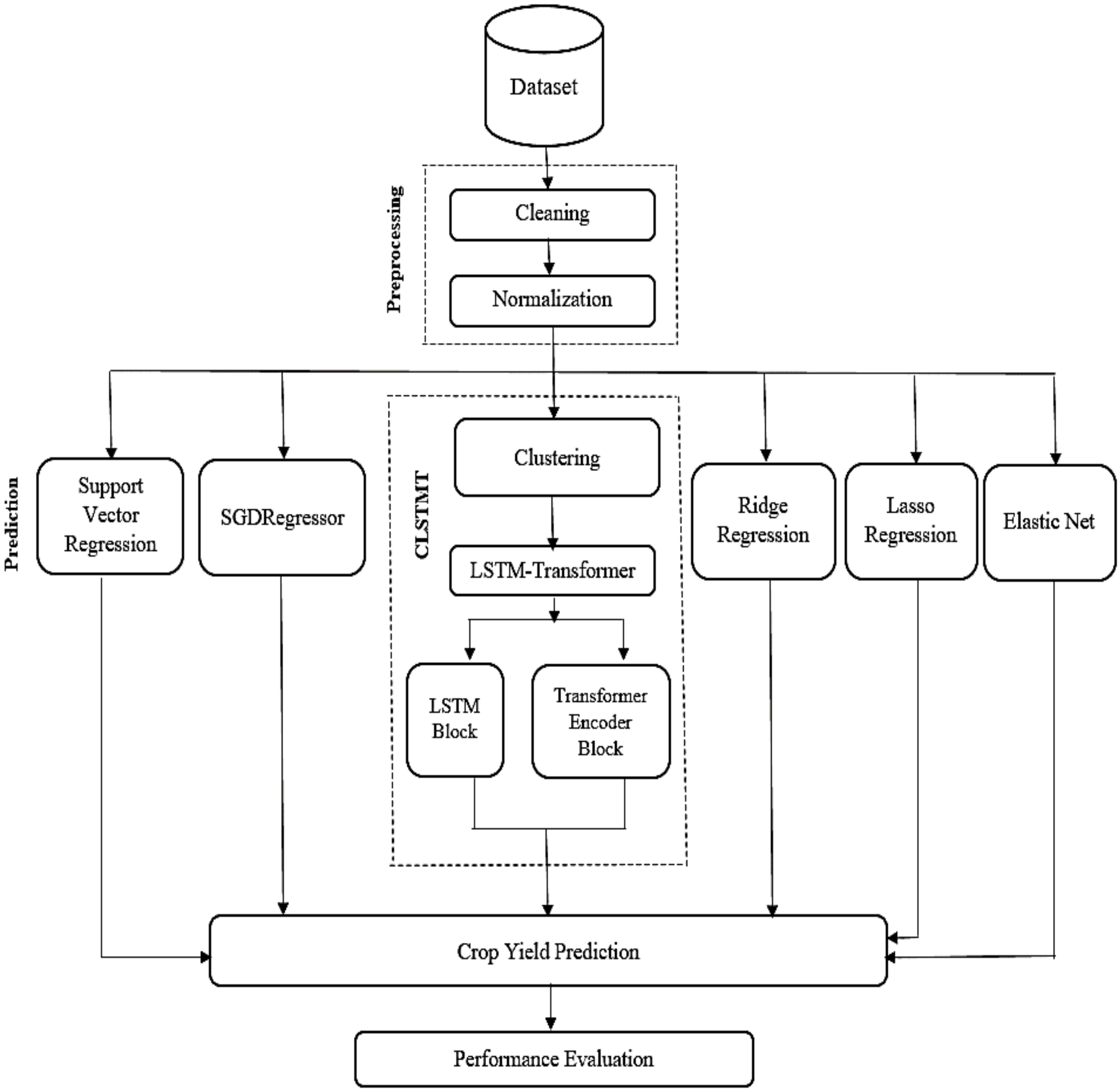

Figure 1. CLSTMT Model Architecture

Its performance is evaluated using R^2 and Mean Absolute Percentage Error (MAPE). The following section discusses the functioning of the transformer encoder block and the LSTM model.

3.1.1 Transformer Encoder Block

The core of the model architecture lies in the definition of the Transformer encoder block, encapsulated within the transformer encoder function. This function integrates normalization, multi-head self-attention, and feed-forward layers into a cohesive block. Initially, layer normalization is applied to the inputs to stabilize and expedite the training process. The multi-head self-attention mechanism enables the model to focus concurrently on various segments of the input sequence, significantly enhancing its ability to capture intricate dependencies within the data. Dropout layers are introduced for mitigating the overfitting issues by randomly nullifying a fraction of input units during the training phase. Subsequently, a feed-forward network, consisting of additional layer normalization and two convolutional layers, processes the data. The first convolutional layer employs a ReLU activation function to introduce non-linearity and the second layer restores the input dimensions. Residual connections, adding the input to the output, are used throughout to facilitate the training of deeper networks by alleviating the vanishing gradient problem.

3.1.2 Transformer LSTM Model

The transformer model function delineates the comprehensive architecture of the Transformer model. The model receives input features with a specified dimensionality of the input embedding. Initially, an embedding layer converts these input features into dense vectors, which are then processed by two Long

Short-Term Memory layers. These LSTM layers are pivotal in capturing temporal dependencies within the data. The embedded input is subsequently fed through multiple Transformer encoder blocks, with the number of blocks determined by the number of transformer blocks parameter, set to four in this case. Following this, a global average pooling layer condenses the data dimensions, effectively summarizing the information from the entire sequence. The processed data is then passed through a series of dense (fully connected) layers, specified by the multilayer perceptron units parameter, which includes layers with 128 and 64 units. In order to prevent the overfitting, each dense layer is followed by a dropout layer. The final output layer is a dense layer with a linear activation function, generates a single value, suitable for the regression task of predicting crop yield.

3.2 Performance Evaluation

The model performance is evaluated in numerous ways using error measures [37, 38]. In the present paper the evaluation is done by using coefficient of determination (R^2) and MAPE. The R^2, also known as the coefficient of determination. It quantifies the portion of the variance in dependent variable that can be considered by the independent variables within a regression model. A greater R^2 value signifies a more robust correspondence between the model and the data, whereas a smaller R^2 value implies that the model inadequately captures the variability present in the data. R^2 is computed as follows,

$$R^2 = 1 - \frac{\sum_1^n (\hat{y}_t - y_t)^2}{\sum_1^n (y_t - \bar{y}_t)^2} \quad (1)$$

Where the number of data, actual, mean of actual and forecast values are denoted by n, 'y_t', '$\bar{y}_t$' and $\hat{y}_t$ respectively [39].

MAPE computes the mean percentage deviation of predicted values from actual values for 'n' samples as follows.

$$\text{MAPE} = \frac{1}{n}\sum_{t=0}^{n}\left|\left(\frac{actual_t - predict_t}{actual_t}\right) * 100\right| \quad (2)$$

Where 'actual$_t$' is the actual observed value at period't' and 'predict$_t$' is the predicted value at period 't' generated by the model [40, 41].

4. Experimental Results

The dataset utilized for the experimental purpose encompassing agricultural data from a diverse array of countries, including Albania, Algeria, Angola, Argentina, and more, reflecting a global scope of agricultural practices and conditions. Within this framework, various crops are represented, ranging from staple grains like maize, rice, and wheat to root vegetables such as potatoes, sweet potatoes, and yams. This comprehensive dataset facilitates the exploration of agricultural trends and practices across different regions and crop types. It is collected from Kaggle database. The dataset contains 28242 samples and 7 attributes such as area, crop (item), year, rain, pest, temp and crop yield. It consists of samples of 10 crops from 101 countries from the year 1990 to 2013. Table 1 shows the statistical details like avg_rain (average rain), pest_ton (Pesticide Tones), avg_temp(average temperature) and hg/ha_yield (Crop Yield) for different geographical area.

The dataset appears to contain information relating to climate factors and pesticide usage for different crops grown in various countries. This comprehensive dataset facilitates the exploration of agricultural trends and practices across different regions and crop types. It is examined to uncover trends in crop production across various countries and to understand how environmental factors influence crop yields. Recognizing the crucial role of agriculture in the global economy, it's essential to grasp worldwide crop yield patterns. This knowledge is vital for tackling food security issues and mitigating the effects of climate change. Crop yield is heavily influenced by weather conditions such as rainfall and temperature, as well as pesticide usage. Having accurate historical data on crop yields is crucial for making informed decisions regarding agricultural risk management.

The dataset will be divided into separate sets for training and testing purposes. Around 70% of the data will be allocated for training the models, 15% for validation and remaining 15% reserved for evaluating their performance. This division guarantees that the models are trained on a significant portion of the dataset and assessed thoroughly on unseen data to assess their ability to generalize.

Table 1. Statistical characteristics of sample data from different area

Area	Item	Year	avg_rain	pest_ton	avg_temp	hg/ha_yield
Albania	Maize	1990	1485	121	16.37	36613
India	Sorghum	1991	1083	72133	26.54	6553
Italy	Rice, paddy	1992	832	88227.6	12.89	58758
Japan	Potatoes	1993	1668	79821.18	14.68	304856
Kenya	Yams	1994	630	3469	16.58	74000

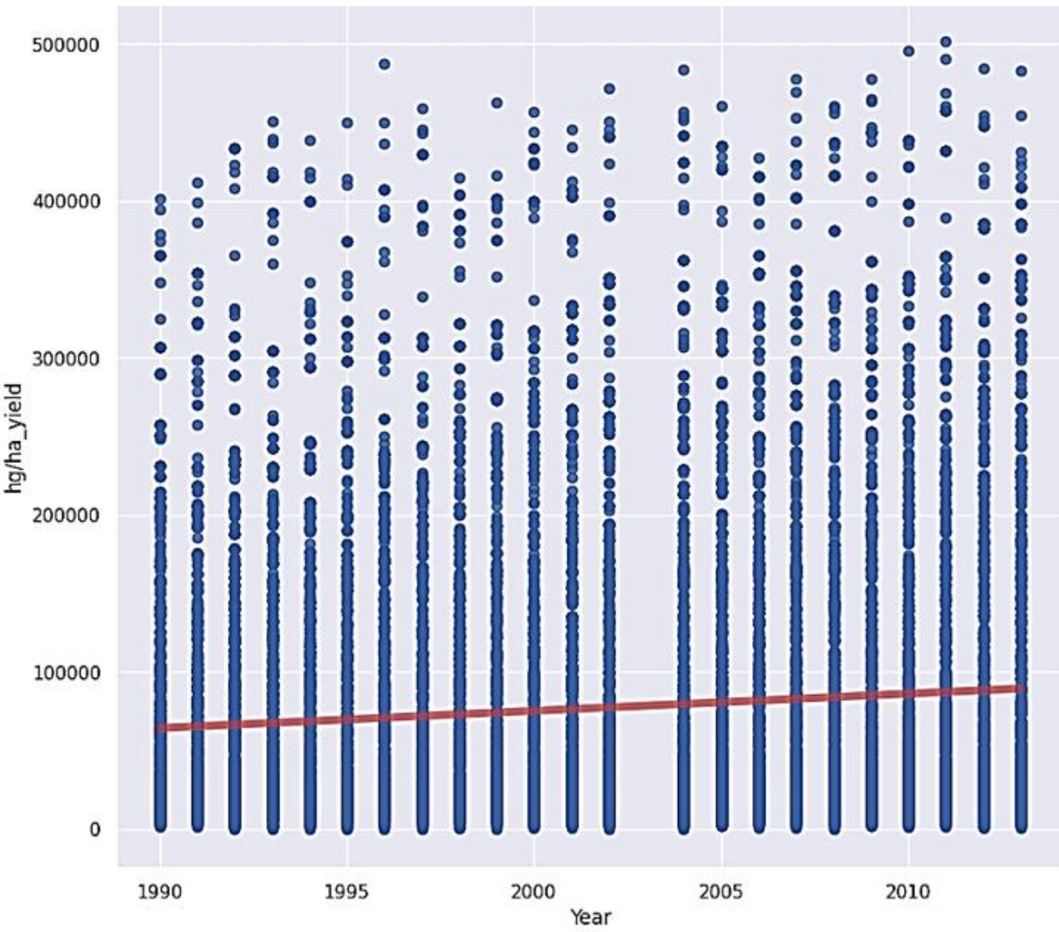

Figure 2. Correlation between year and yield

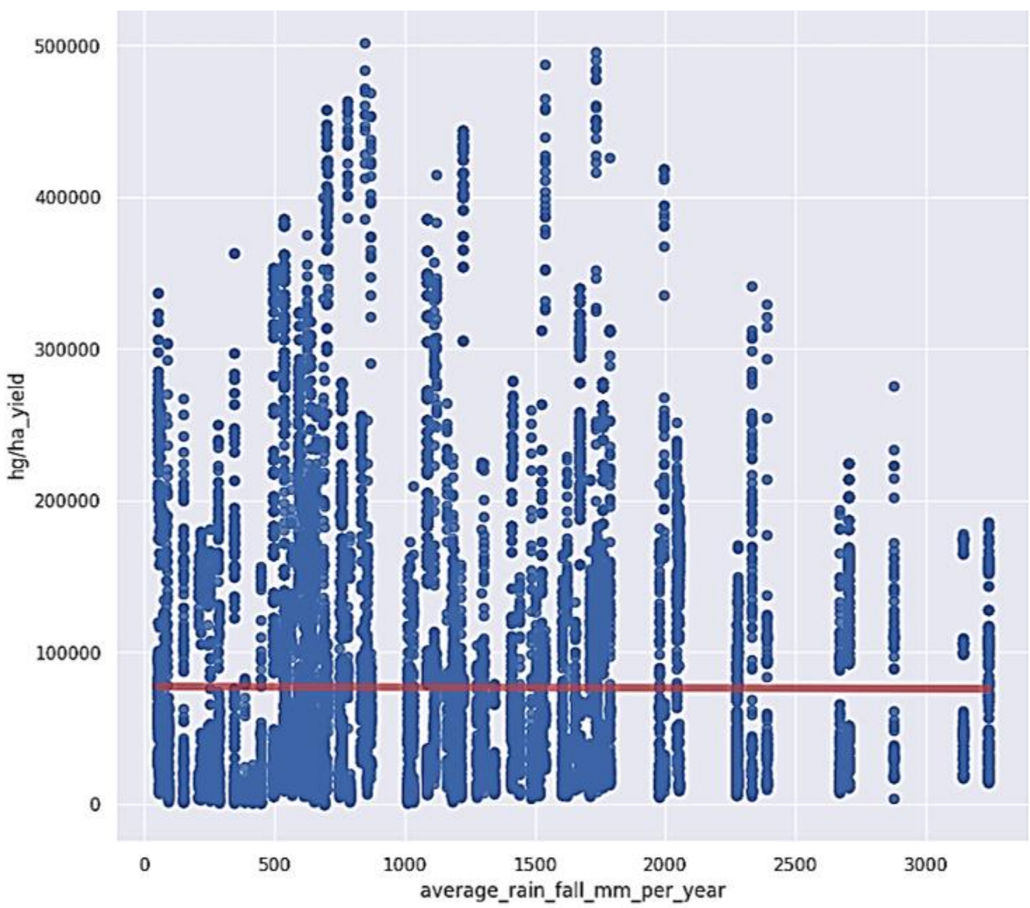

Figure 3. Correlation between average rainfall and yield

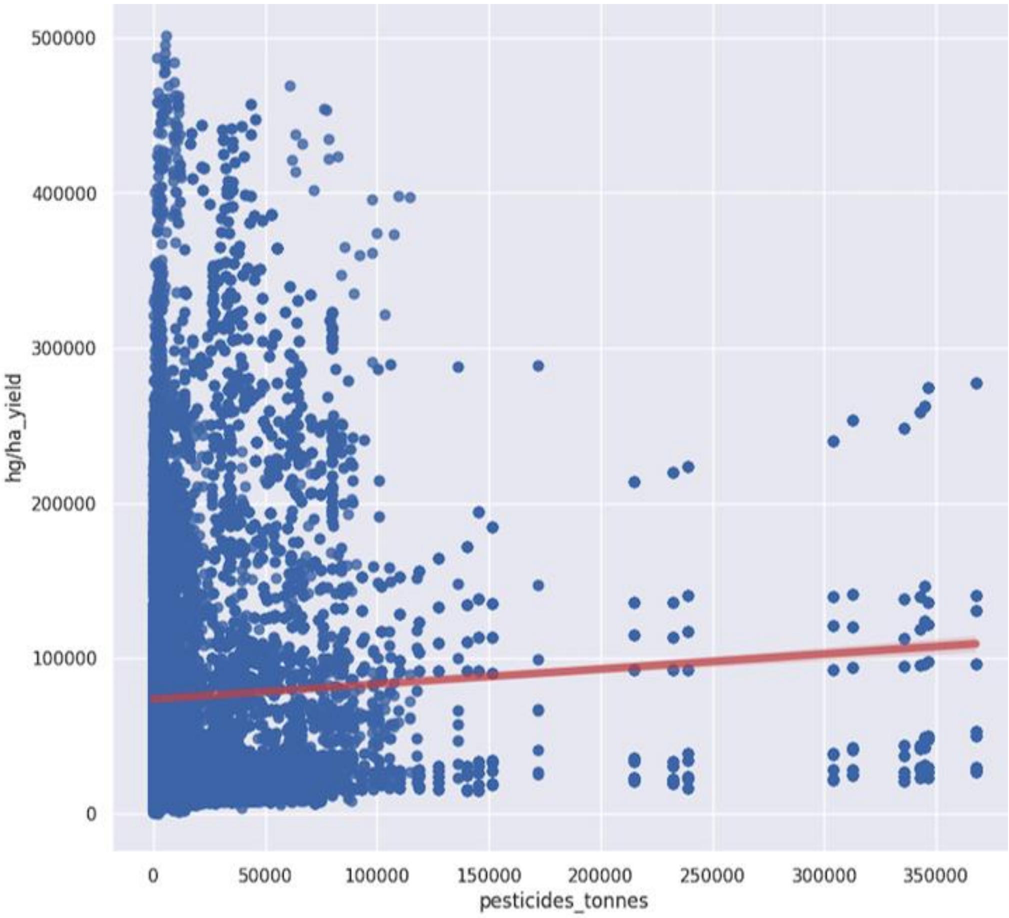

Figure 4. Correlation between pesticide tonnes and yield

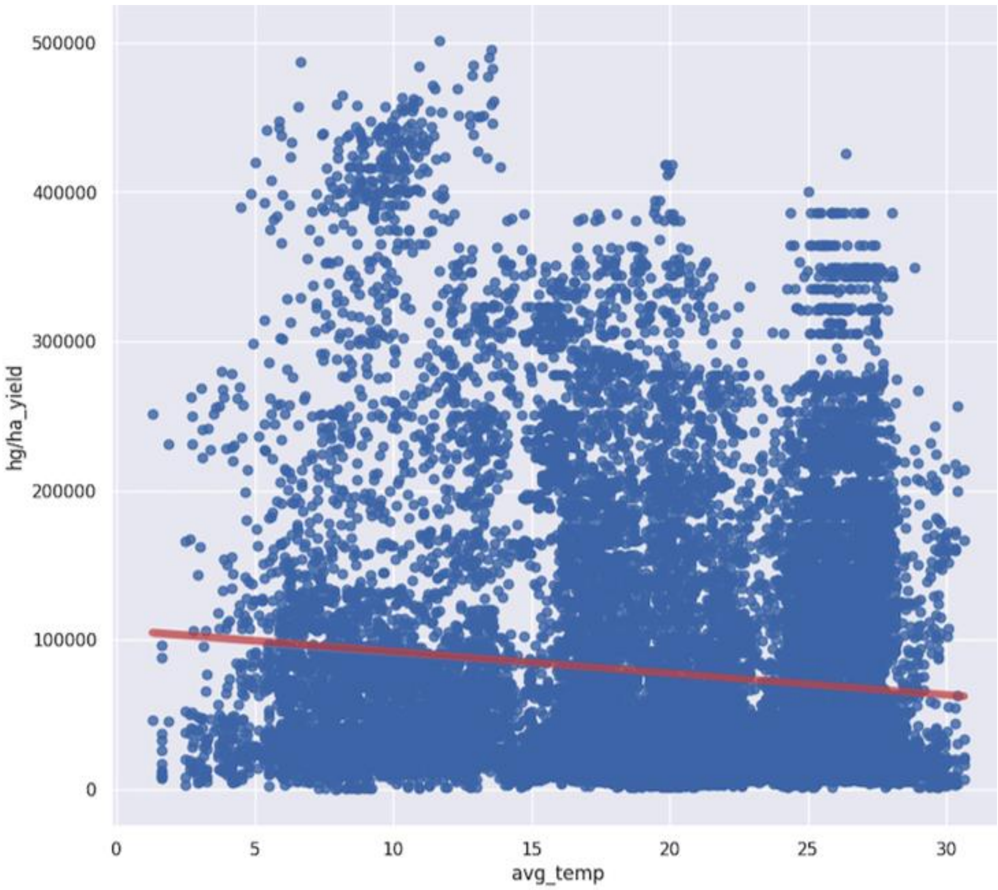

Figure 5. Correlation between average temperature and yield

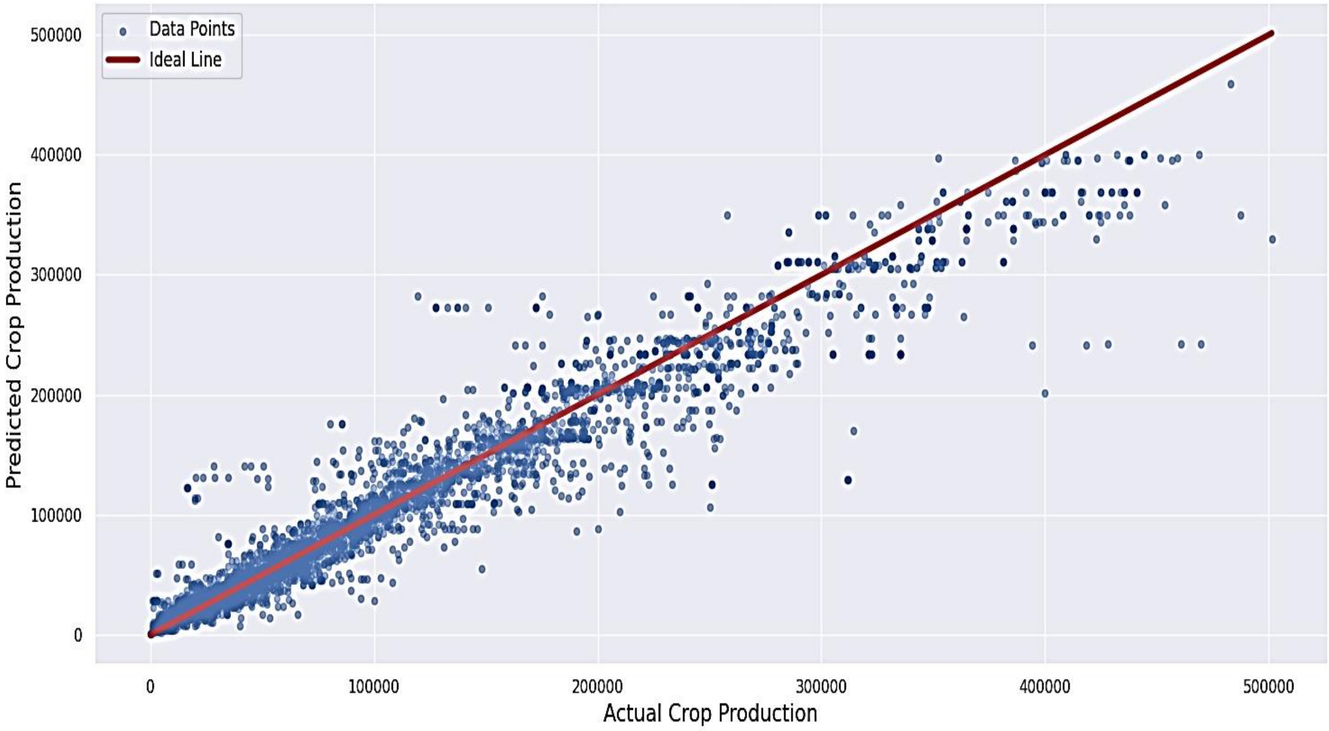

Figure 6. CLSTMT: Predicted crop yield vs. Actual crop yield

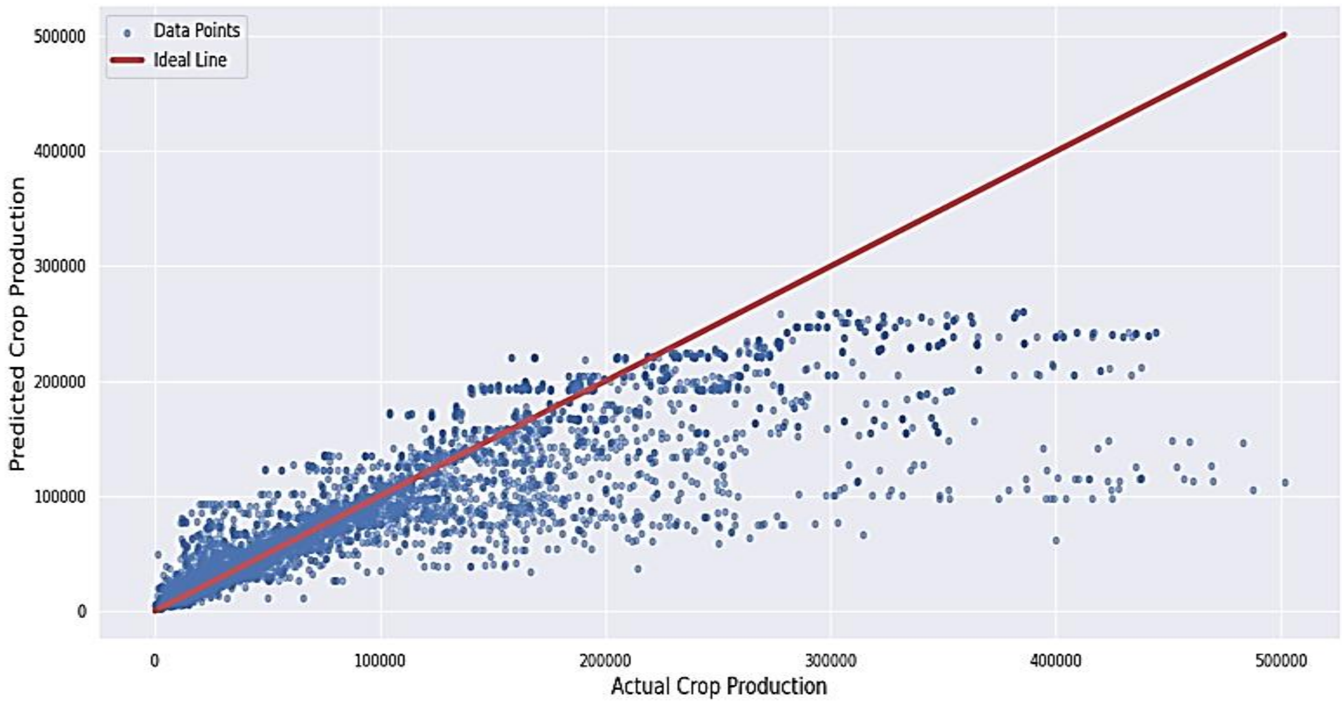

Figure 7. Support Vector Regression: Predicted crop yield vs. Actual crop yield

During the Exploratory data analysis the scatter plots are produced to show the correlation between year, avg_rain (average rain), pest_ton(Pesticide Tones), avg_temp(average temperature) with hg/ha_yield (Crop Yield). It is shown in Figure 2 to Figure 5. The regression line in each plot illustrates the general trend among two variables.

This study suggests creating a tailored deep learning model using Transformer Encoders to predict crop yields by utilizing past climate data. The proposed model integrates LSTM layers and Transformer Encoders to effectively capture intricate patterns influenced by climatic fluctuations and enhance crop yield predictions accuracy. The primary objective of the Transformer Encoder block is to integrate attention mechanisms and effectively handle lengthy sequences, thereby improving the predictive capabilities of the model. The model comprises multiple dense layers with varying sizes and ReLU activation functions.

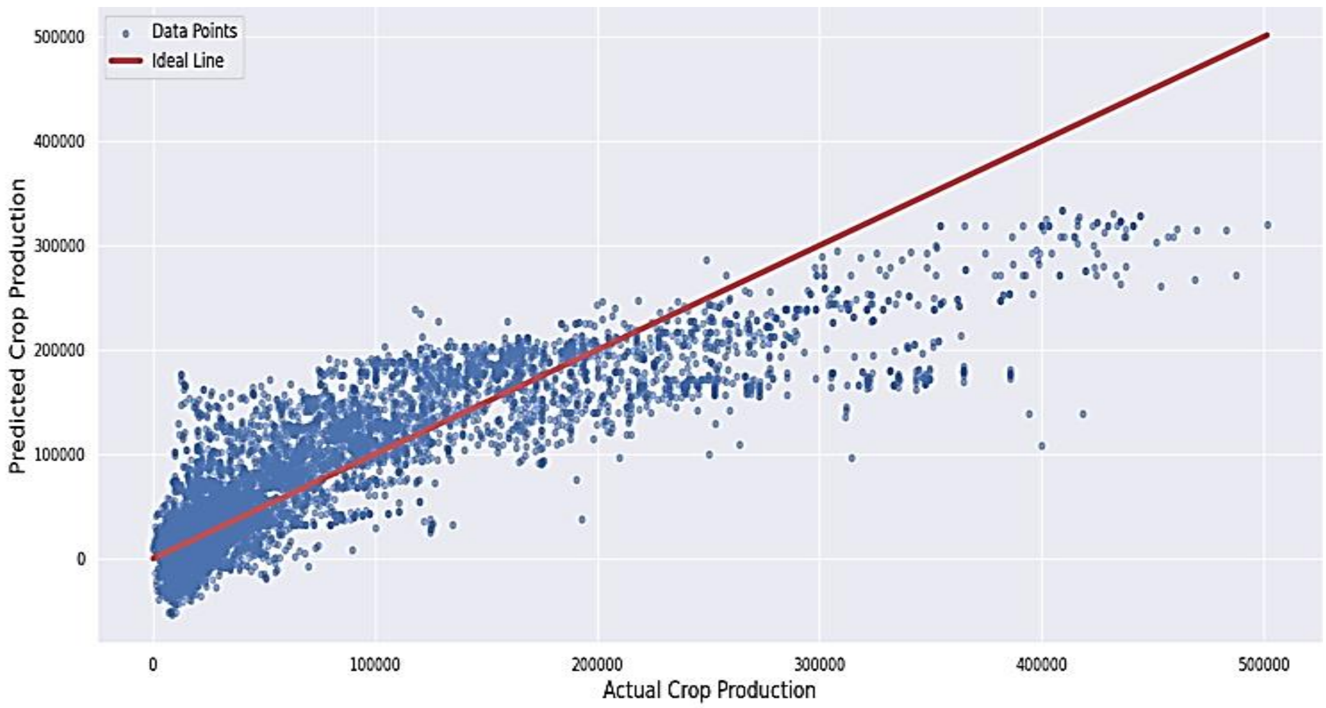

Figure 8. SGDRegressor : Predicted crop yield vs. Actual crop yield

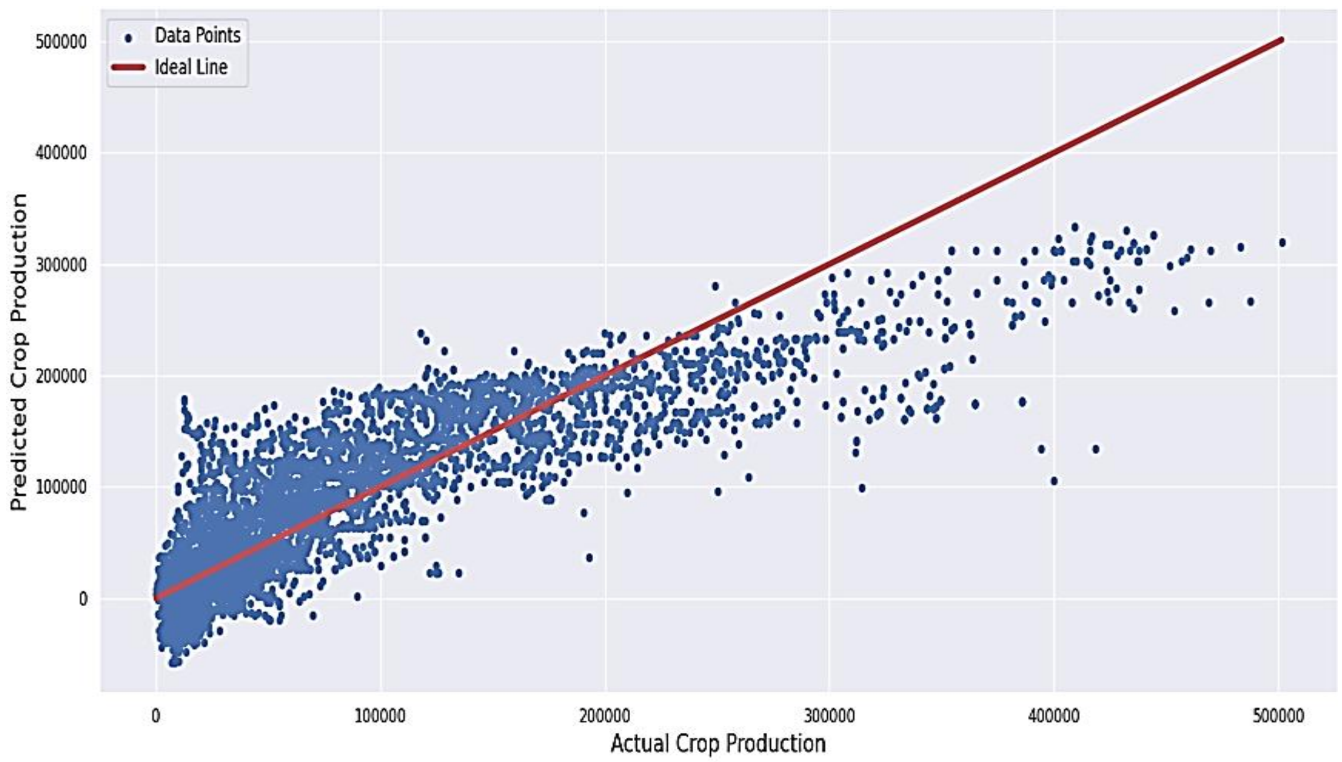

Figure 9. Ridge Regression: Predicted crop yield vs. Actual crop yield

It concludes with a dense layer that utilizes a linear activation function to generate the output. The approach used involves employing the Adam optimizer and the MAE loss function for training the model. The model is compiled and trained by utilizing scaled target variables, thereby ensuring a stable optimization process. During the training process, the model undergoes 20 iterations, known as epochs, with a batch size of 64. The purpose of these iterations is to optimize the model's parameters and minimize the MAE loss function. The performance of the model is validated using the validation data. Testing dataset containing the independent variables (features) for making predictions and the corresponding target variable (crop yield), to spot any overfitting issues that could have occurred. Followed by the training phase the model generates predictions for the test dataset.

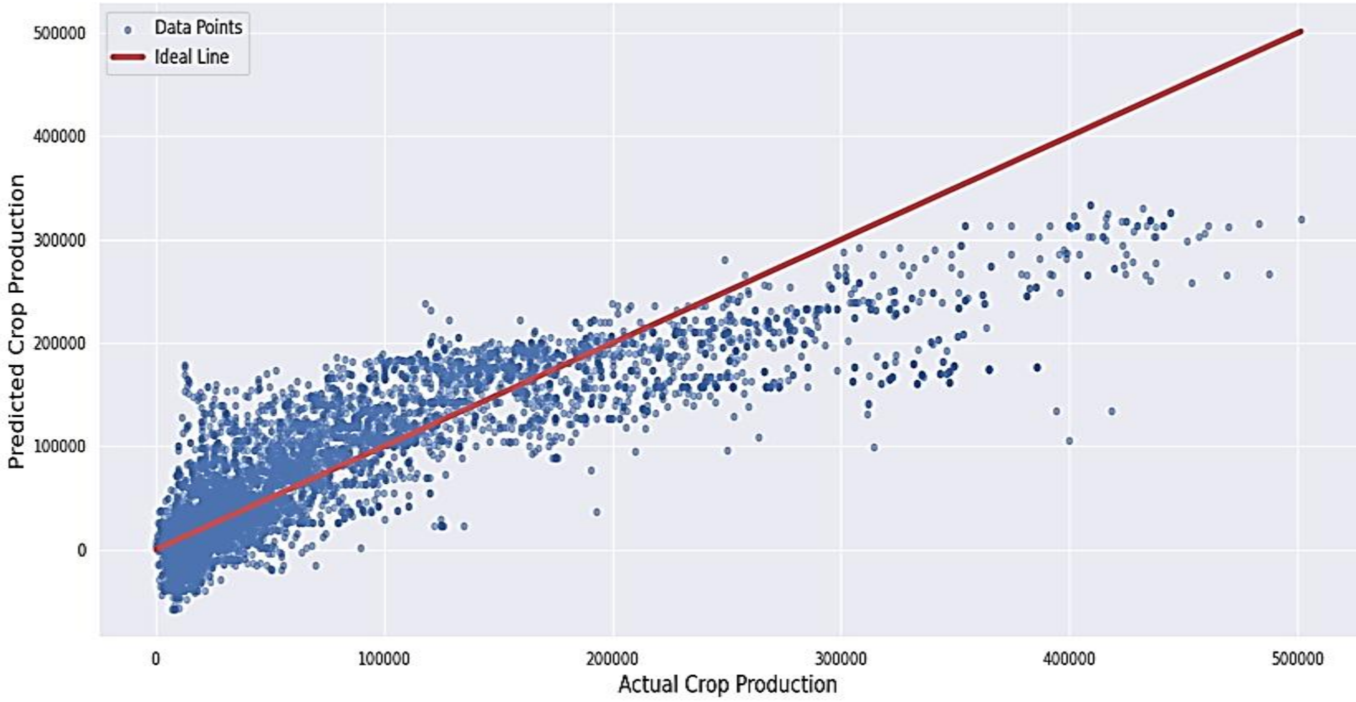

Figure 10. Lasso Regression: Predicted crop yield vs. Actual crop yield

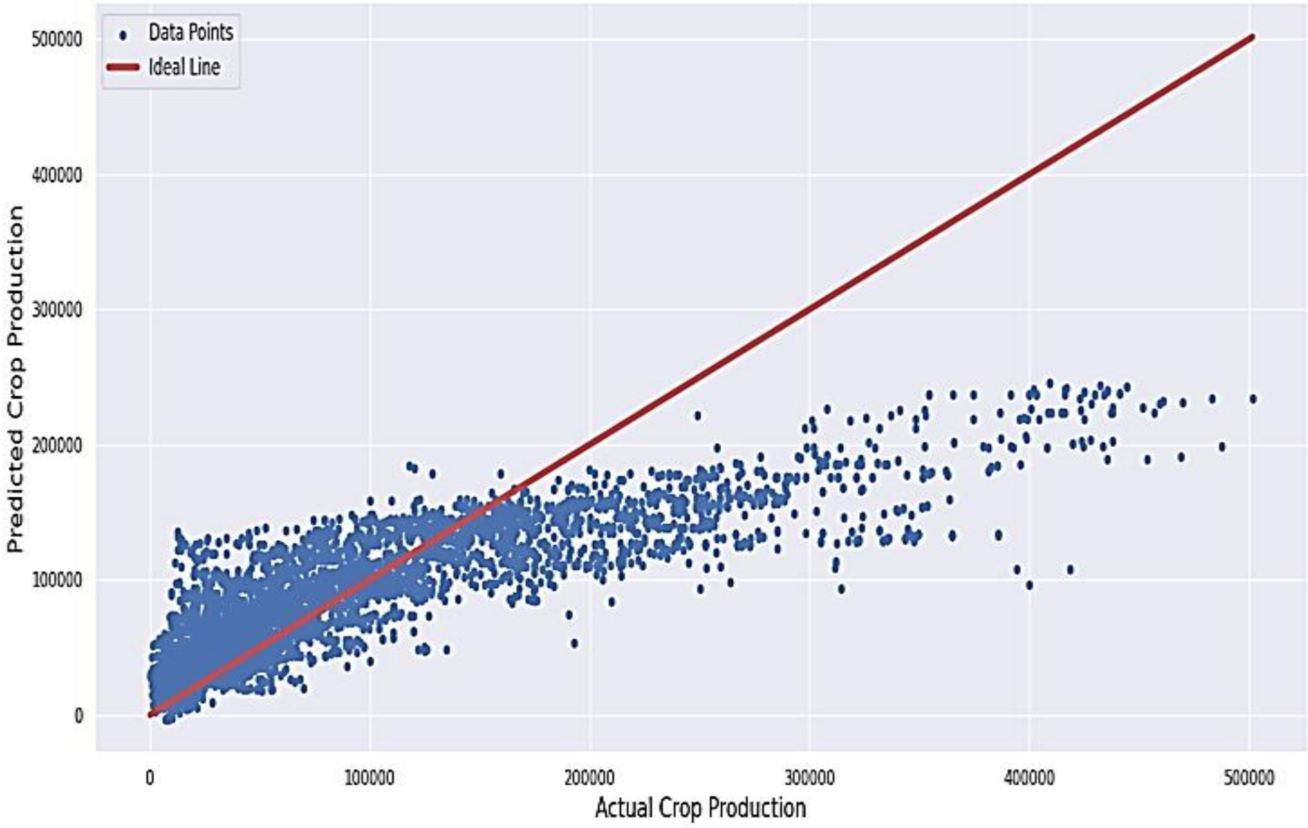

Figure 11. Elastic Net: Predicted crop yield vs. Actual crop yield

The predictions are then reverted to their original values using the MinMaxScaler and becomes the predicted crop yield for the testing dataset.

The model enables a precise and comprehensive examination of the complex relationship between crop production and climate change by integrating deep learning methodologies, including LSTM and Transformer Encoders. The predictions generated by the model offer significant insights into how the increased temperatures and decreased precipitation impacts on the crop yield, underscoring the importance of addressing regional variations in crop responses to shifting climatic conditions. The research findings highlight the importance of tailoring adaptive strategies to local conditions to allay the adverse effects of climate change on agricultural production. This novel approach allows for the development of proactive measures to safeguard global food security in the times of changing climatic conditions. The experimental results of CLSTMT, Support Vector Regression [42], SGDRegressor [43], Ridge Regression [42], Lasso Regression [42] and Elastic Net [42] are presented in Figure 6 to Figure 11.

It creates a clear and informative scatter plot to visualize the performance of a model's predictions in comparison to the actual crop production values. The correlation in Figure 6 shows the predicted value of crop yield using CLSTMT model is closer to the actual crop yield. It provides much accurate results than other models. Figure 7 shows the comparison of predicted crop yield of SVR is little closer to actual crop. Figure 8 represents the comparison of predicted crop yield of SGDRegressor is little less close to actual crop. Figure 9 shows the comparison of predicted crop yield of RidgeRegression is has some deviation to actual crop.

Figure 10 demonstrates the comparison of predicted crop yield of Lasso Regression much deviates from actual crop compared. Figure 11 shows the comparison of predicted crop yield of ElasticNet more deviates from actual crop compared.

5. Discussion

The correlation of actual and predicted crop yield of CLSTMT, Support Vector Regression [42], SGDRegressor [43], Ridge Regression [42], Lasso Regression [42] and Elastic Net [42] are compared using scatter plot and shown in Figure 12. It demonstrates that the CLSTMT predicted crop has high correlation with actual crop yield compared to others. The predicted crop yield produced by all the methodologies are compared against the actual crop yield in Figure 13. It shows the crop yield predicted by the CLSTMT (Yellow) is closer to the actual crop yield (Blue) compared to others. Figure 14 shows the comparison of the actual crop yield and predicted crop yield of first 100 samples for the clear understanding of the CLSTMT performance. In this, the CLSTMT graph is closest to actual crop yield graph in almost all data points compared to support vector regression, SGDRegressor, RidgeRegression, lasso regression and ElasticNet. The comparison of model performance is given in Table 2. The crop yield prediction performance is measured using MAPE and R^2. It shows the crop yield prediction performance of CLSTMT model is high. Followed by support vector regression, SGDRegressor, Ridge Regression and Lasso Regression. Finally the Elastic Net provides the least performance. The SGDRegressor, Ridge Regression and Lasso Regression provide the similar prediction performance.

Figure 12. Scatter plot of actual and predicted crop yield

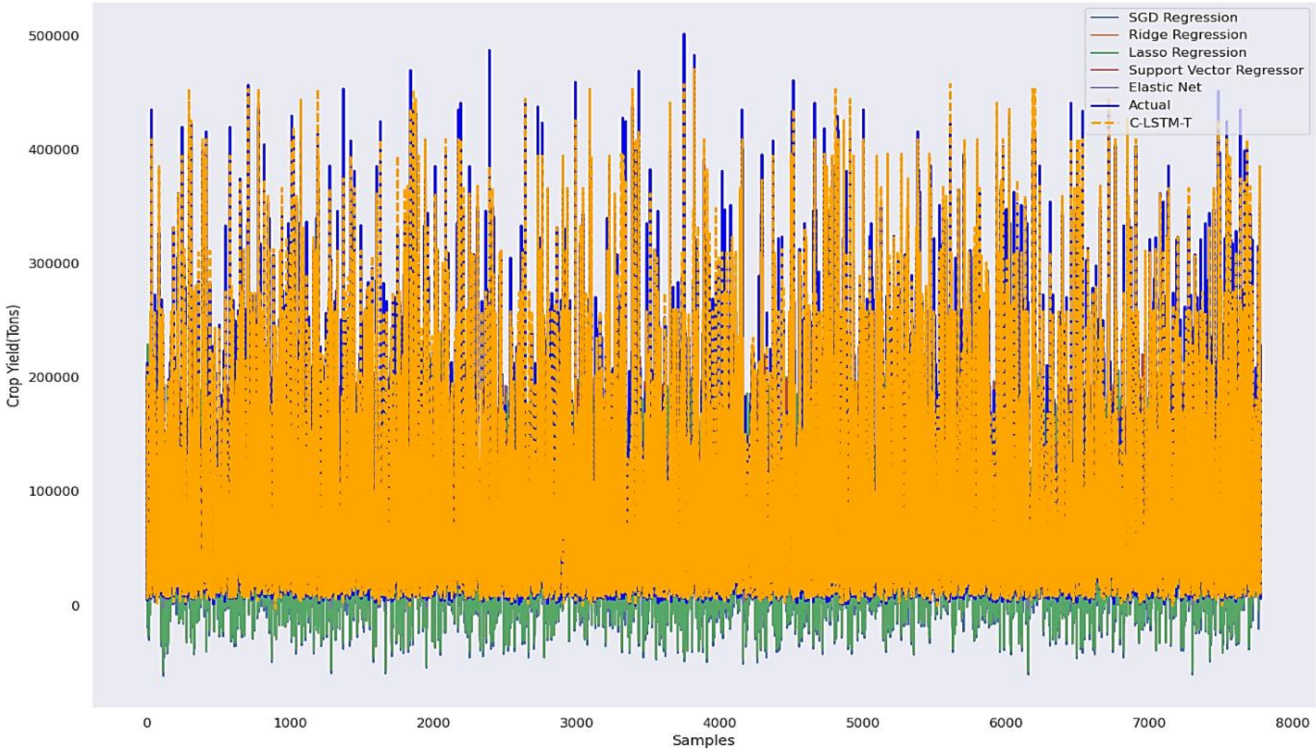

Figure 13. Comparison of actual and predicted crop yield

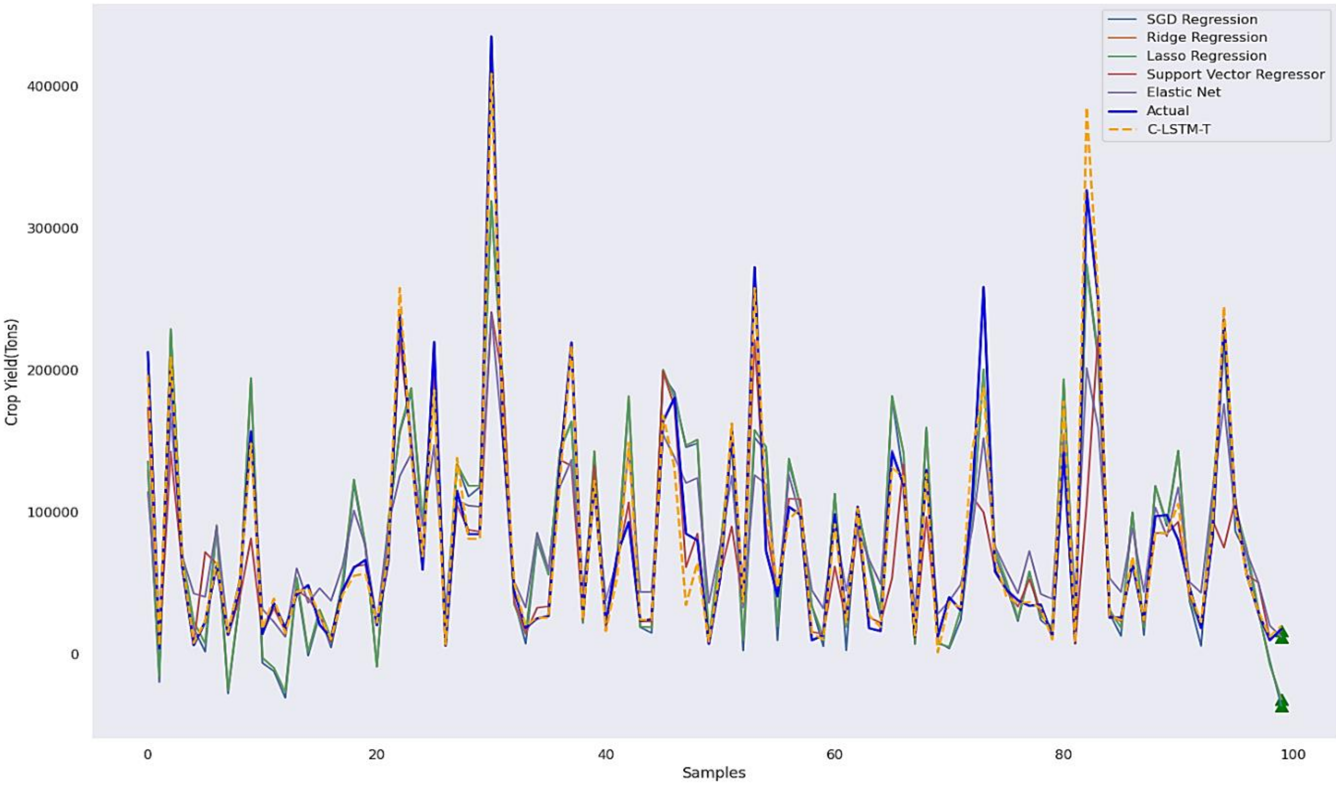

Figure 14. Comparison of actual and predicted crop yield for first 100 samples

As a result the deep learning based CLSTMT outperformed other competing regression methods by producing a least MAPE of 0.195 and high R-squared of 0.951. It shows the CLSTMT model outperforms other models in predicting crop yields using climatic data. The CLSTMT model proves its superior performance with R^2 of 0.951, indicating that it has the capability of approximately 95.1% of the variability in crop yields. This highlights its remarkable ability to capture complex internal patterns and connections in the dataset. Additionally, the model demonstrates a high level of accuracy with a MAPE value of 0.195 indicating a minimal average percentage deviation of 0.195% between observed and predicted yields.

Table 2. Comparison of Model Performance

Methodology	R^2	MAPE
CLSTMT	0.951	0.195
Support Vector Regression	0.788	0.253
SGD Regression	0.745	0.923
Ridge Regression	0.745	0.923
Lasso Regression	0.745	0.923
Elastic Net	0.668	1.044

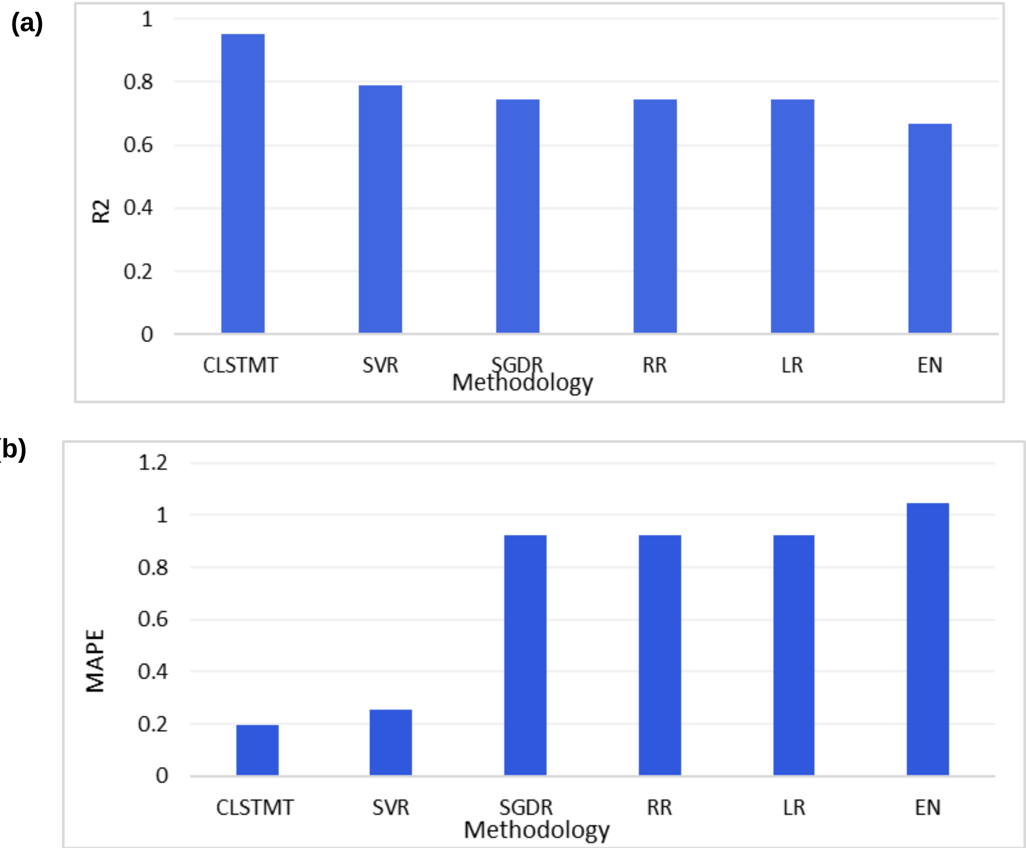

Figure 15. Comparison of actual and predicted crop yield for first 100 samples

The R^2 of Support Vector Regression, SGDRegressor, Ridge Regression, Lasso Regression and Elastic Net are 0.788, 0.745, 0.745, 0.745 and 0.668 respectively. The MAPE of Support Vector Regression, SGDRegressor, Ridge Regression, Lasso Regression and Elastic Net are 0.253, 0.923, 0.923, 0.923, 1.044 respectively. The CLSTMT model consistently outperforms other regression models in crop yield prediction under changing climatic conditions, establishing its superiority in terms of robustness and effectiveness.

The integration of clustering, LSTM and Transformer enables it to effectively comprehend the complex internal relationship between crop productivity and climate variables. This capability facilitates proactive measures and ensures food security in climate change. The performance of each model is compared in terms of R^2 and MAPE and it is shown graphically in Figure 15.

6. Conclusion

The application of predictive modeling techniques plays a pivotal role in assessing the detrimental impact of the climate change on crop yields, thereby addressing the agricultural issues on a global scale and safeguarding food security. The present study introduces an innovative methodology that integrates conventional regression models with sophisticated deep learning techniques, specifically LSTM and Transformer, for the purpose of predicting crop yields based on historical climate data. The proposed model effectively captures climate-induced patterns by integrating Transformer with LSTM layers and FNN components, resulting in accurate and comprehensive predictions of crop yield. The comparative analysis demonstrates the superiority of the presented CLSTMT model over other regression models. These insights can be utilized by policymakers and farmers to develop tailored adaptive strategies, thereby addressing regional disparities in crop responses to climate change and ensuring the protection of global food security.

With the proactive implementation of these strategies, it is possible to guarantee a future that is both sustainable and food-secure for all individuals. In this research, the dataset utilized is limited to a certain number of features. Future research endeavors could potentially delve into the integration of more features and the examination of regional disparities, thereby augmenting the model's accuracy and accommodating the dynamic nature of climate conditions. Numerous prospects for future research and enhancement exist within this domain. Consideration of more contemporary data will help to capture the most up-to-date patterns in climate change and its ramifications on crop yields. Performing a more detailed examination at the regional or local scale has the potential to yield more profound understandings regarding the distinct difficulties encountered by various geographical regions. Subsequent investigations may delve into strategies for integrating these uncertainties into the prediction models, thereby enhancing the resilience and dependability of crop yield projections across various climate scenarios.

References

[1] F.M. Talaat, Crop yield prediction algorithm (CYPA) in precision agriculture based on IoT techniques and climate changes. Neural Computing and Applications, 35(23), (2023) 17281-17292. https://doi.org/10.1007/s00521-023-08619-5

[2] M. Abdel-salam, N. Kumar, S. Mahajan, A proposed framework for crop yield prediction using hybrid feature selection approach and optimized machine learning. Neural Computing and Applications, 36, (2024) 20723–20750. https://doi.org/10.1007/s00521-024-10226-x

[3] E.S.M. El-Kenawy, A.A. Alhussan, N. Khodadadi, S. Mirjalili, M.M. Eid, Predicting Potato Crop Yield with Machine Learning and Deep Learning for Sustainable Agriculture. Potato Research, (2024) 1-34. https://doi.org/10.1007/s11540-024-09753-w

[4] K. Meghraoui, I. Sebari, J. Pilz, K. Ait El Kadi, S. Bensiali, Applied Deep Learning-Based Crop Yield Prediction: A Systematic Analysis of Current Developments and Potential Challenges. Technologies, 12(4), (2024) 43. https://doi.org/10.3390/technologies12040043

[5] S.N. Khan, D. Li, M. Maimaitijiang, Using gross primary production data and deep transfer learning for crop yield prediction in the US Corn Belt. International Journal of Applied Earth Observation and Geoinformation, 131, (2024) 103965. https://doi.org/10.1016/j.jag.2024.103965

[6] M. Habib-ur-Rahman, A. Ahmad, A. Raza, M. U. Hasnain, H.F. Alharby, Y.M. Alzahrani, A.A. Bamagoos, K. R. Hakeem, S. Ahmad, W. Nasim, S. Ali, Impact of climate change on agricultural production; Issues, challenges, and opportunities in Asia, Frontiers in Plant Science, 13, (2022) 925548. http://dx.doi.org/10.3389/fpls.2022.925548

[7] T. Hu, X. Zhang, S. Khanal, R. Wilson, G. Leng, E. M. Toman, X. Wang, Y. Li, K. Zhao, Climate change impacts on crop yields: A review of empirical findings, statistical crop models, and machine learning methods. Environmental Modelling & Software, 179, (2024) 106119. https://doi.org/10.1016/j.envsoft.2024.106119

[8] M. Albahar, A Survey on Deep Learning and Its Impact on Agriculture: Challenges and Opportunities. Agriculture, 13(3), (2023) 540. http://dx.doi.org/10.3390/agriculture13030540

[9] E.M. Al-Ali, Y. Hajji, Y. Said, M. Hleili, A.M. Alanzi, A. H. Laatar, M. Atri, Solar Energy Production Forecasting Based on a Hybrid CNN-LSTM-Transformer Model. Mathematics, 11(3), (2023) 676. http://dx.doi.org/10.3390/math11030676

[10] F. Andayani, L.B. Theng, M.T. Tsun, C. Chua, Hybrid LSTM-transformer model for emotion recognition from speech audio files. IEEE Access, 10, (2022) 36018-36027. http://dx.doi.org/10.1109/ACCESS.2022.3163856

[11] P. Datta, B. Behera, Climate change and Indian agriculture: A systematic review of farmers' perception, adaptation, and transformation. Environmental Challenges, 8, (2022) 100543. http://dx.doi.org/10.1016/j.envc.2022.100543

[12] S.S. Subbiah, P.S. Kumar, Deep learning based load forecasting with decomposition and feature selection techniques. Journal of Scientific & Industrial Research, 81(5), (2022) 505-517. http://op.niscpr.res.in/index.php/JSIR/article/view/56794

[13] Y. Liu, S. Wang, J. Chen, B. Chen, X. Wang, D. Hao, L. Sun, Rice yield prediction and model interpretation based on satellite and climatic indicators using a transformer method. Remote Sensing, 14(19), (2022) 5045. https://doi.org/10.3390/rs14195045

[14] L. Bi, O. Wally, G. Hu, A.U. Tenuta, Y.R. Kandel, D.S. Mueller, A transformer-based approach for early prediction of soybean yield using time-series images. Frontiers in Plant Science, 14, (2023) 1173036. https://doi.org/10.3389/fpls.2023.1173036

[15] P.S. Nishant, P. S. Venkat, B.L. Avinash, B. Jabber, (2020) Crop yield prediction based on Indian agriculture using machine learning. In 2020 International Conference for Emerging Technology (INCET), IEEE, India. http://dx.doi.org/10.1109/INCET49848.2020.9154036

[16] M. Chandraprabha, R.K. Dhanaraj, (2020) Machine learning based Pedantic Analysis of Predictive Algorithms in Crop Yield Management. In 2020 4th International conference on electronics, communication and aerospace technology (ICECA), IEEE, India. http://dx.doi.org/10.1109/ICECA49313.2020.9297544

[17] M.T. Shakoor, K. Rahman, S.N. Rayta, A. Chakrabarty, (2017) Agricultural production output prediction using supervised machine learning techniques, In 2017 1st international conference on next generation computing applications (NextComp), IEEE, Mauritius. http://dx.doi.org/10.1109/NEXTCOMP.2017.8016196

[18] S.S. Reddy, N. Sethi, R. Rajender, (2020) Evaluation of deep belief network to predict hospital readmission of diabetic patients, In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), IEEE, India. http://dx.doi.org/10.1109/ICIRCA48905.2020.9182800

[19] S. Agarwal, S. Tarar, A hybrid approach for crop yield prediction using machine learning and deep learning algorithms, In Journal of Physics: Conference Series, IOP Publishing, 1714(1), (2021) 012012. http://dx.doi.org/10.1088/1742-6596/1714/1/012012

[20] D. Elavarasan, P.D. Vincent, Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. IEEE access, 8, (2020) 86886-86901. http://dx.doi.org/10.1109/ACCESS.2020.2992480

[21] K. Palanivel, C. Surianarayanan, An approach for prediction of crop yield using machine learning and big data techniques. International Journal of Computer Engineering and Technology, 10(3), (2019) 110-118. http://dx.doi.org/10.34218/IJCET.10.3.2019.013

[22] S.K.S. Durai, M.D. Shamili, Smart farming using machine learning and deep learning techniques. Decision Analytics Journal, 3, (2022) 100041. http://dx.doi.org/10.1016/j.dajour.2022.100041

[23] L.S. Cedric, W.Y.H. Adoni, R. Aworka, J.T. Zoueu, F.K. Mutombo, M. Krichen, C.L.M. Kimpolo, Crops yield prediction based on machine learning models: Case of West African countries, Smart Agricultural Technology, 2, (2022) 100049. http://dx.doi.org/10.1016/j.atech.2022.100049

[24] K.L.M. Ang, J.K.P. Seng, Big data and machine learning with hyperspectral information in agriculture. IEEE Access, 9, (2021) 36699-36718. http://dx.doi.org/10.1109/ACCESS.2021.3051196

[25] M. Kalimuthu, P. Vaishnavi, M. Kishore, (2020) Crop prediction using machine learning. In 2020 third international conference on smart systems and inventive technology (ICCSIT), IEEE, India. http://dx.doi.org/10.1109/ICSSIT48917.2020.9214190

[26] B. Deforce, B. Baesens, J. Diels, E. Serral Asensio, Forecasting sensor-data in smart agriculture with temporal fusion transformers. Transactions on Computational Science & Computational Intelligence, (2022).

[27] S. Nagini, T.R. Kanth, B.V. Kiranmayee, (2016) Agriculture yield prediction using predictive analytic techniques. In 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), IEEE, India. http://dx.doi.org/10.1109/IC3I.2016.7918789

[28] T. Junankar, J.K. Sondhi, A.M. Nair, (2023) Wheat Yield Prediction using Temporal Fusion Transformers. In 2023 2nd International Conference for Innovation in Technology (INOCON), IEEE, India. http://dx.doi.org/10.1109/INOCON57975.2023.10101144

[29] Y. Ang, H.Z.M. Shafri, Y.P. Lee, S.A. Bakar, H. Abidin, M.U.U. Mohd Junaidi, S.J. Hashim, N.N. Che'Ya, M.R. Hassan, H.S. Lim, R. Abdullah, Oil palm yield prediction across blocks from multi-source data using machine learning and deep learning. Earth Science Informatics, 15(4), (2022) 2349-2367. http://dx.doi.org/10.1007/s12145-022-00882-9

[30] S. Goel, S. Markanday, S. Mohanty, (2022) Classification of Agriculture Crops Using Transfer Learning. In 2022 OITS International Conference on Information Technology (OCIT), IEEE, India. http://dx.doi.org/10.1109/OCIT56763.2022.00058

[31] S. Gurrapu, F.A. Batarseh, P. Wang, M.N.K. Sikder, N. Gorentala, M. Gopinath, (2021) Deepag: Deep learning approach for measuring the effects of outlier events on agricultural production and policy. In 2021 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, USA. http://dx.doi.org/10.1109/SSCI50451.2021.9659921

[32] C. El Hachimi, S. Belaqziz, S. Khabba, A. Chehbouni, (2021) Towards precision agriculture in Morocco: A machine learning approach for recommending crops and forecasting weather. In 2021 International Conference on Digital Age &

Technological Advances for Sustainable Development (ICDATA), IEEE, Morocco. http://dx.doi.org/10.1109/ICDATA52997.2021.00026

[33] K. Parasuraman, U. Anandan, A. Anbarasan, (2021) IoT based smart agriculture automation in artificial intelligence, In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), IEEE, India. http://dx.doi.org/10.1109/ICICV50876.2021.9388578

[34] S.S. Subbiah, J. Chinnappan, A review of bio-inspired computational intelligence algorithms in electricity load forecasting. Smart Buildings Digitalization, (2022) 169-192. http://dx.doi.org/10.1201/9781003201069-11

[35] S.S Subbiah, J. Chinnappan, Short-term load forecasting using random forest with entropy-based feature selection, In Artificial Intelligence and Technologies: Select Proceedings of ICRTAC-AIT 2020, Springer Singapore, (2021) 73-80. http://dx.doi.org/10.1007/978-981-16-6448-9_8

[36] H. Kukadiya, D. Meva, N. Arora, S. Srivastava, Effective Groundnut Crop Management by Early Prediction of Leaf Diseases through Convolutional Neural Networks. International Research Journal of Multidisciplinary Technovation, 6(1), (2024) 17-31. http://dx.doi.org/10.54392/irjmt2412

[37] V.G. Kiruthika, V. Arutchudar, P. Senthil Kumar, Highest humidity prediction using data mining techniques. International Journal of Applied Engineering Research, 9(16), (2014) 3259-3264.

[38] G. Swaroop, P. Senthil Kumar, T. Muthamil Selvan, An efficient model for share market prediction using data mining techniques. International Journal of Applied Engineering Research, 9(17), (2014) 3807-3812.

[39] V. Sellam, E. Poovammal, Prediction of crop yield using regression analysis. Indian Journal of Science and Technology, 9(38), (2016) 1-5. http://dx.doi.org/10.17485/ijst/2016/v9i38/91714

[40] S.S. Sankari, P.S. Kumar, A Review of Deep Transfer Learning Strategy for Energy Forecasting. Nature Environment and Pollution Technology, 22(4), (2023) 1781-1793. http://dx.doi.org/10.46488/NEPT.2023.v22i04.007

[41] M. Qiao, X. He, X. Cheng, P. Li, H. Luo, L. Zhang and Z. Tian. Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks. International Journal of Applied Earth Observation and Geoinformation, 102, (2021) p.102436. https://doi.org/10.1016/j.jag.2021.102436.

[42] P.P Jorvekar, S.K. Wagh, J.R. Prasad. Predictive modeling of crop yields: a comparative analysis of regression techniques for agricultural yield prediction. Agricultural Engineering International: CIGR Journal, 26(2), (2024) 102436.

[43] I. Gupta, S. Ayalasomayajula, Y. Shashidhara, A. Kataria, S. Shashidhara, K. Kataria, S. Raj, M. Kurtz. A. Undurti, Innovations in Agricultural Forecasting: A Multivariate Regression Study on Global Crop Yield Prediction, International Journal of Advanced Research in Computer and Communication Engineering, 13(9), (2024) 171-181. http://dx.doi.org/10.17148/IJARCCE.2024.13922 .

Authors Contribution Statement

Yash Pravesh S: Conceptualization, Methodology, Data collection, Formal analysis, Writing -Original Draft Nakshatra Garg: Formal analysis, Writing - Original Draft, Ravik Arora: Methodology, Software validation, Sudhanshu Singh: Editing, Validation, Siva Sankari S: Supervision, Review & Editing.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Has this article screened for similarity?

Yes

About the License