ISSN [Online]: 2583-2654

# An Enhanced Sugarcane Yield Prediction Using Deep Learning approach

*Shalini G, Nisha N, Kruthika R, K S Sinchana, K S Sahana*
*Department of Computer Science & Engineering,Dr T Thimmaiah Institute of Technology KGF, Karnataka, India*

***Abstract:*** Sugarcane cultivation is of prime importance to the agricultural sector, contributing significantly to the global sugar and bio energy industries. Efficiently predicting sugarcane yield is crucial for optimizing resource management, enhancing agricultural productivity, and supporting sustainable practices. However, old methods of yield prediction results poor in capturing the complex dynamics inherent in sugarcane growth. This work addresses the inherent challenges in sugarcane yield prediction by leveraging a deep learning approach, adept at understanding complex patterns. The work intends for nuanced and precise sugarcane yield prediction, overcoming the drawbacks of conventional methods. The proposed framework ensembles neural networks for enhanced accuracy. Artificial intelligence analyzes vast datasets, including climate and soil conditions. The deep learning model adapts to nonlinear relationships, enabling precise predictions. This work, integrating deep learning and a web flask user interface contributes to precision agriculture and sustainable sugarcane cultivation.

***Keywords: DeepLearning, Precision Agriculture, LSTM, FFNN, Sugarcane yield prediction, Autoencoders***

## I. INTRODUCTION

Sugarcane is a kharif crop that has a pivotal role in global agriculture cultivation, contributing significantly in the manufacturing of sugar and bio ethanol. Accurate yield prediction is paramount for optimizing agricultural practices and ensuring economic viability. However, usage of statistical tools will not go well in capturing the nuanced relationships between various agronomic factors, weather patterns, and soil attributes. This project leverages cutting-edge algorithms to revolutionize sugarcane yield prediction. By amalgamating diverse datasets encompassing soil profiles, historical weather data, and crop management practices, our work is to construct a comprehensive predictive model. Through the usage of deep learning frameworks, we endeavor to unravel intricate patterns and dependencies, ultimately providing precise and reliable yield forecasts.

The proposed framework extends beyond yield prediction, aspiring to empower farmers and stakeholders with a user-friendly tool. A Flask-based web application is developed, to integrate the state-of-the art deep learning model for real-time predictions. This platform allows users to input relevant data, such as soil characteristics, weather conditions, and management practices, and receive instantaneous yield estimates. Moreover, the application provides visualizations for enhanced user interpretation of both input data and predicted outcomes. By deploying this application on a cloud-based server, we ensure accessibility and scalability, enabling widespread use within the sugarcane farming community.

The potential outcomes of the work is substantial, not only for sugarcane farming but also for the broader agricultural sector. Moreover, this research contributes to the adoption of precision agriculture, aligning with the global pursuit of sustainable farming practices.

## II. RELATEDWORK

There are several research works that have proved to be efficient in precision agriculture to address challenges faced by farmers in predicting the yield using conventional methods.

**M Naveen Kumar et al. [1]** ,. Here , the focus is to build a system to achieve high accuracy and resolve

ISSN [Online]: 2583-2654

the challenges of yielding in agriculture. They used the LS-SVM and LSSVM SAtechnique in predicting the yield of crop based on physical and chemical features of plant and soil. This systemhelps the farmer to opt the right decision about sowing particular crop in available land. Simulated Annealing is inclined to the procedure of annealing in metallurgy. In this natural procedure a metal is heated and it is slowly cooled using controlled conditions to magnify the size of the crystals in the materialand reduce their defects. Prediction of sugarcane yield with the help of LS-SVM was identified to be a significant tool. The predicted output from the present study proved that the LS-SVM model can be applied to real-life prediction and modeling problems and it appears to perform reasonably well. **Ramesh A. Medar et al. [2]** in the year 2019 has designed a model utilizing Supervised Machine Learning concept to improvise the Sugar cane Crop Yield prediction . In this work all the data consists of different features over the past few year 's daily records. All the information was collected from the various web sources over a period of years and implemented the data into this work. They use LTTS, weather and soil attributes, NDVI and SML algorithms to implement the sugarcane yield forecasting. The authors predicts sugarcane yield taking into consideration , sowing start period under different conditions in India. The study focuses on leveraging historical data related to meteorological conditions, soil characteristics, and previous crop yields which trains the machine learning model. Many supervised learning algorithms, such as regressionor ensemble methods, are likely explored to predict sugarcane yields accurately. **Dhivya Elavarasan et al. [3]** have used Reinforcement learning instead of supervised and unsupervised techniques to limit the mapping issues of crop yield data, as these data are very non-linear in nature. In this ,they have constructed a Deep Recurrent Q-Network model which is a Recurrent Neural Network deep learning algorithm over the Q-Learning reinforcement learning algorithm to forecast the crop yield. The layers are sequentially stacked , Recurrent Neural network is fed by the data parameters. The Q- learning network will construct a crop yield prediction environment from the input parameters. A linear layer maps the Recurrent Neural Network output values to the Q-values. The implemented model gave a pretty good accuracy of 93.7 and paved a way to the future work that can deploy LSTM over RNN which can handle the vanishing gradient problem. **Hrishikesh Kalbhor et al. [4]** In present world the very crucial most thing for living in the Indian economy is Agriculture. Above 70% of the world's population is likely to be dependent on agriculture. Variety of crops are cultivated in India, with wheat being one of the most important food grains cultivatedand exported by this country. It can thus be seen that wheat is a big part of the Indian agricultural system and the economy of India. Therefore, it is very important to maintain the steady production ofthe above-stated crop. To handle the segmentation of the system we use the crop predictive model. Planning for agriculture counts for a major role in agro-based countries' economic development and food security. In agricultural planning, the choice of crops is a significant question. **Sammy V. Militantte et al. [5],** in yield predictions even disease identification has a primary role in improving the yield of the crop hence in this research the authors have adopted deep learning for multi classification where 7 diseases of sugarcane-leaves is trained to the system using the classic CNN(Convolution neural network) and obtained an accuracy of 95%. The trained model has reached its goal by effectively detecting and classifying sugarcane images into healthy and diseased category based on patterns of leaves. hence, this study will be of great use in helping farmers to use computer vision and machine learning in detecting and classifying sugarcane diseases. For, Future work, different models may be implemented to identify the performance of the model on thetraining set. Different Learning Rates and Optimizers may also utilize for experimenting with the proposed models.**Thomas Van Klompernburg et al. [6] ,** This work has shown how the related works used different of attributes, depending on the kind of the experiment and the handiness of data. Each paperinvestigated yield prediction with machine learning but differed from the attribute being considered. The studies also differed in scale, geographic location , and also crop being taken into account. The selection of features is dependent on the availability of thedataset and the aim of the work. To find the best performing model, models with more and fewerfeatures should be tested. The results showcased that we cannot infer very specified conclusion from it and come to a conclusion on the best model , but they clearly show that few machine learning models are used more frequently than the others. The most used models are the random forest,neural networks, linear regression, and gradient boosting tree. **Debnath Bhattacharyya et al. [7] ,**This paper has deployed Hybrid CNN-SVM classifier to process semi structured data in sugarcane yield forecasting production. They have worked on collecting soil moisture data and its effect in classifying the target variable sugarcane yield This work is done in two phases, In the first phase, ensembelled classifiers are used to predict soil moisture . Next round, to boost the models performance, they have used Probabilistic models. The suggested approach aims to forsee soil moisture values measured that affects crop growth and cultivation. The output model is 89.53% accurate than other traditionalneural network classifiers. The put forward models results will aid farmers and agricultural agencies in improvising the production. **Kushal B J et al.[8]** Research in the realm of the agriculture sector is expanding. More than half of India's population depends on agriculture for livelihood, and it is a major contributor to the country's economic growth. Soil quality is changing drastically, affecting the agricultural crop yield. Data on soiltemperature, soil humidity, rain, soil moisture, and pH are needed to train the machine-learning models.This work has been carried out using the following machine learning models: Decision Tree classifier, K-Neighbor classifier, and Random Forest classifier models. The predicted accuracy of the Random Forest classifier is 93.11 percent, which is higher than the predicted accuracy of the Decision Tree classifier (90.96 percent) and the accuracy of the K-Neighbors classifier (87.63 percent). As farmers are less consciousof the nutrient quality of farmland's soil to cultivate the crop, real-time data from IoT devices like pH sensors, rainfall sensors, humidity sensors, humidity & temperature sensors, soil moisture sensors, and Node MCU to determine which crop would be the best fit for a certain plot of land. **Pezhman Taherei Ghazvinei et al. [9]** The research paper focuses on predicting sugarcane growth by leveraging meteorological parameters, employing two distinct machine learning approaches: Extreme Learning Machine (ELM) and Artificial Neural Network (ANN). The core objective is to develop accurate and reliable models for forecasting sugarcane growth based on environmental factors. The study acknowledges the significance of meteorological conditions in influencing cropperformance and aims to harness advanced computational techniques to enhance prediction accuracy. The comparative analysis of these two machine learning approaches provides insights into theirrespective strengths and limitations. Additionally, the study may discuss the dataset used,encompassing relevant meteorological variables, and the training/testing methodology employed tovalidate the predictive models. **Jaturong Som-ard et al. [10]** This paper explores the various applications of remote sensing technology in the context of sugarcane cultivation. Remote sensing involves the use of satellite or airborne sensors to gather information about the Earth's surface, and in this case, its application is tailored to benefit sugarcane farming. The primary aim of the research is to assess and analyze the potential of remote sensing tools in monitoring and managing various aspects of sugarcane cultivation.The paper likely delves into the specific remote sensing techniques employed, such as satellite imageryor aerial photography, to gather data on sugarcane fields. It may discuss the utilization of these technologies for tasks such as monitoring crop health, estimating yields, detecting diseases, and assessing environmental conditions. The research might highlight the advantages of remote sensing in providing real-time and large-scale information, contributing to precision agriculture practices in sugarcane cultivation. **Laura Martinez- Ferror et al. [11]** The paper focuses on crop yield estimation, specifically employing Gaussian processes as a modeling tool, and emphasizes the importance of interpretability in the context of agricultural predictions. Gaussian processes are probabilistic models that offer a flexible framework for capturing complex relationships in data. The primary objective of the paper islikely to assess the effectiveness of Gaussian processes in predicting crop yields while maintaining a high level of interpretability. Furthermore, the paper may present comparative analyses with other modeling techniques, showcasing the strengths and advantages of Gaussian processes in terms of bothprediction accuracy and interpretability. The findings could have implications for precision agriculture,where farmers and decision makers benefit from models that not only make accurate predictions but also offer insights into the factors influencing those predictions.**T R Rajesh et al. [12]** The paper introduces an advanced machine learning-based approach for enhancing disease prediction in sugarcane crops. By leveraging comprehensive datasets, including environmental and historical factors, the study employs sophisticated machine learning strategies to provide more accurate and timely predictions. The goal is to equip farmers with an accurate tool for early disease detection, contributing to improved crop management and reduced losses in sugarcane cultivation. It aims to improve disease prediction in sugarcane crops by employing advanced machine learning techniques. The central agenda is to develop a more effective and accurate model for early detection of diseases in sugarcane plants, ultimately aiding farmers in timely intervention and crop management.

## III. DATASET DESCRIPTION

The Sugarcane dataset collected has a wide range of years, from 1997 to 2020, and includes data for various states of India, It provides detailed information on the area under cultivation, production, rainfall, fertilizer and pesticide usage, and yield for sugarcane across different seasons in these states.This dataset can be useful for analysing sugarcane production trends, studying the impact of various factors(like rainfall, fertilizer, and pesticide usage) on yield, and comparing production levels across states and seasons. It Consists of 605 data points with 8 input features and 1 target feature.The classes: State, Season, Area, Production, Annual rainfall, Fertilizer, Pesticide and target feature Yield. For labelling the named values into numeric values label encoding is performed.

## IV METHODOLOGY

The Paper proposes to detect yield of sugarcane using the state-of-art deep learning approaches and techniques for fine tuning the model. The proposed system includes neural networks for enhanced accuracy. This advanced modeling technique harnesses the powerfulness of deep learning to accurately forecast sugarcane crop yields based ona comprehensive set of input features, including factors like crop year, season, state, area, production,annual rainfall, fertilizer, and pesticide usage. The features are fed as input to data preprocessing. This involves encoding categorical variables, feature scaling, normalization of target variable. Next Identify relevant features from the dataset that can be used to make accurate predictions. Train the model using the training dataset using deep learning algorithms. Then Test the model and evaluate its performance on a separate validation dataset to ensure it generalizes well to new, unseen data. By employing deep learning techniques, the system can effectively capture complex, non-linear relationships within the data, enabling it to adapt to varying agricultural conditions and provide moreaccurate yield predictions.

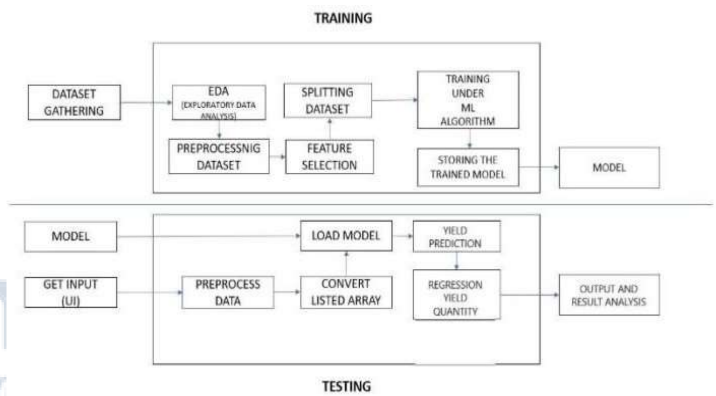


**Fig.1 SystemArchitecture**

The system architecture for our sugarcane yield prediction project involves the arrangement and interaction of various components to ensure the effective functioning of the entire system. At its core, the architecture includes three main stages: data preprocessing, model training, and yield prediction. In the data preprocessing stage, the system collects and refines input data, addressing issues like null values and conversion of named variables .

## V RESULTS &ANALYSIS

In the evaluation phase of our sugarcane yield prediction project, we meticulously assess the performance of the Feedforward Neural Network (FFNN) and Long Short-Term Memory (LSTM) models. As we have done work on regression task , the evaluation metrics to assess are Mean Squared Error

ISSN [Online]: 2583-2654

and Mean Absolute Error , we gauge their effectiveness in predicting sugarcane yields on the testing set, which comprises previously unseen data. Leveraging cross-validation techniques ensures the robustness of our evaluations, providing a comprehensive understanding of each model's generalization capabilities.

For experimentation purposes, the parameters, batch size were set to 32 to obtain the final results for different deep learning techniques.

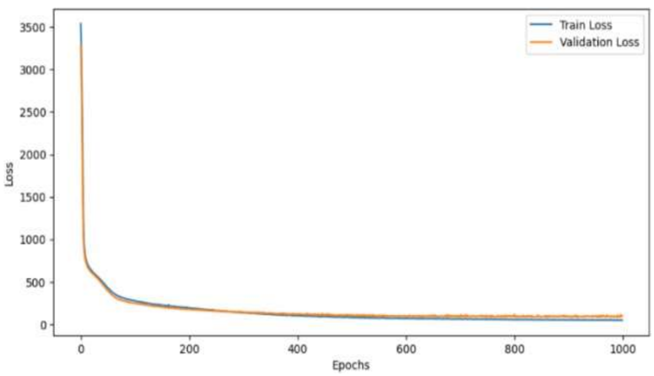


**Fig.2 Train and validation loss graph of FFNN model**

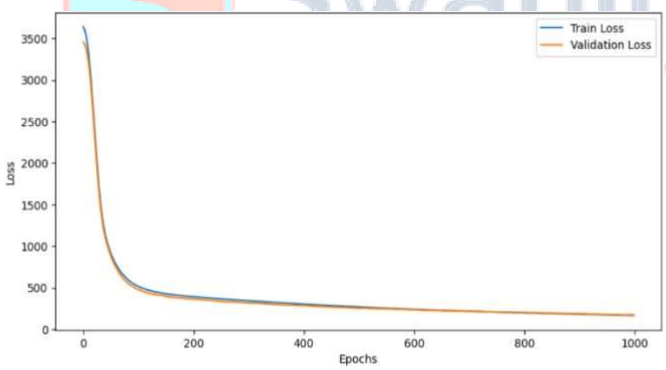


**Fig.3 Train and validation loss graph of LSTM model**

The Fig 2, graph depicts FFNN training progress with train and valid loss plotted over epochs. Initially high, both losses decrease steadily as training advances, indicating effective learning. By 200-300 epochs, both losses stabilize, suggesting optimal model performance. This convergence signifies successful generalization to unseen data without overfitting.

In Fig3, graph illustrates training and validation loss curve for LSTM model. The training loss initially high, rapidly declines, then gradually decreases as the model fine-tunes. Validation loss follows a similar pattern, indicating good generalization. Optimal training results in converging low losses signifying effective learning and generalization. Loss curve shapes and convergence rates depend on model architecture hyperparameters, and problem complexity.

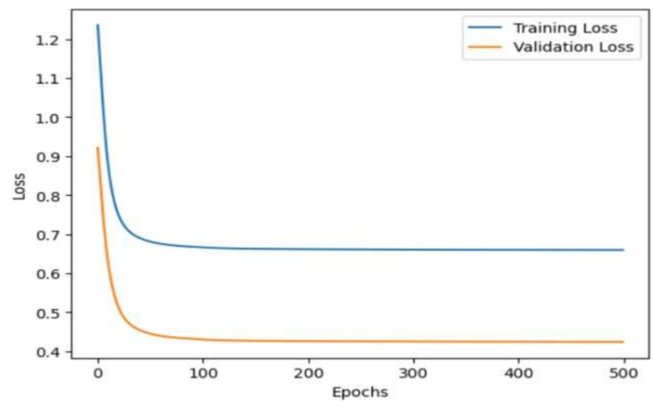


**Fig.4 Train and validation loss graph of Autoencoders model**

In Fig 4, graph illustrates loss curves of an autoencoder model, initially high but declining over epochs. Stabilization after around 200 epochs indicates effective data encoding and minimal reconstruction error, influenced by model architecture and data complexity. In this work, FFNN achieved the top score with accuracy of 86%, followed closely by LSTM of 84%, and Auto encoders with 65%. Given the superior performance of FFNN and LSTM, they were preferred for integrating into the Flask web application. FFNN, with its multi-layer architecture, excels in capturing complex patterns in the data. resulting in high prediction accuracy. Meanwhile, LSTM, known for its ability to retain and utilize sequential information, offers robust prediction capabilities, particularly in time-series data like crop yield. By leveraging both models within the Flask web application, users can access accurate sugarcane yield forecasts, empowering decision-making in agricultural management.


**Fig 5 :Front End**


**Fig 6: The user login page for registered users**


**Fig.7 User Registration Page**

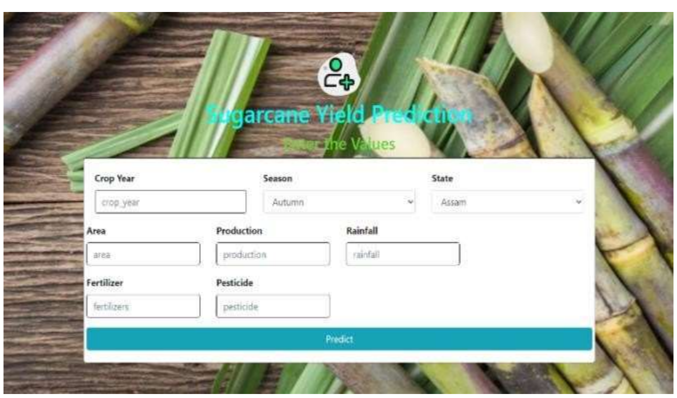


**Fig.8 Page to enter the values**

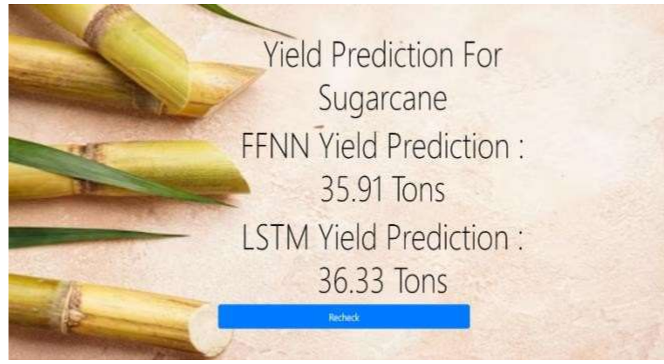


**Fig .9: Displays the predictions**

## VI. CONCLUSION AND FUTURE WORK

In this research, we have applied the Long Short-Term Memory (LSTM) networks, Feed forward Neural Networks(FFNN),and Auto encoders to design a User interface for an enhanced sugarcane yield prediction system. By exploiting the strengths of each approach, our ensemble model aims to deliver superior predictive performance. The LSTM model excels in capturing temporal dependencies and sequential patterns, which is very important for knowing the involvement of weather fluctuations and farming practices on sugarcane yields over time. Meanwhile, the FFNN excels in modeling intricate relationships among agricultural factors contributing to yield variations. Integration of Auto encoders facilitates noise reduction and data simplification, focusing on the most influential features, though we trained this model we did not use in web flask interface as the accuracy was struck at 65% for our surprise. Our model's robustness and generalizability are ensured through training on a diverse dataset spanning multiple seasons and regions, while a user friendly Flask-based web application enables stakeholders to access real-time yield forecasts. This work underscores the superiority of deep learning in precision agriculture, paving the way for smarter decision-making and resource optimization in sugarcane farming. Future endeavors could explore integrating remote sensing data, expanding to other crops, enhancing model interpretability, ensuring deployment scalability, integrating real-time data streams, and providing personalized farming recommendations, ultimately refining and advancing sustainable agricultural practices.

***REFERENCES***

*[1].M Naveen Kumar and Dr. M. Balakrishnan, "Prediction of Sugarcane Yield using LSSVM and LSSVM with Simulated Annealing based Algorithm" ICAR , IJRAR Research Gate ,Vol-06,2019*

*[2] Ramesh A. Medar "Sugarcane Crop Yield Forecasting Model Using Supervised Machine Learning" Intelligent Systems and applications, Vol- 8, pp 11-20,2019*

*[3] DHIVYA ELAVARASAN AND P. M. DURAIRAJ VINCENT, Crop Yield Prediction Using Deep ReinforcementLearning Model for Sustainable Agrarian Applications, IEEE Acess 2020, Digital Object Identifier 10.1109/ACCESS.2020.2992480*

*[4]Hrishikesh Kalbhor et al, "Crop Yield prediction Using Machine Learning "International Research Journal of Engineering and Technology, Vol-09 Issue 04, 2022*

*[5] Sammy V. Militante et al,"Sugarcane Disease Recognition Using Deep Learning" IEEE Eurasia Conference on IOT, Communication and Engineering 2019.*

*[6]Thomas Van Klompernburg et al,"Crop Yield Prediction Using Machine Learning" Information Technology Group 2020.*

*[7]Debnath Bhattacharyya 'Hybrid CNN-SVM Classifier Approaches to process semi structured data in Sugarcane yield Forecasting" Proceedings of the Third International Conference on Computing Methodologies and Communication (ICCMC), pp 775-778,2023.*

*[8]Kushal B J, "Real Time Crop Prediction based on Soil Analysis using IOT and ML" Proceedings of International Conference on Edge Computing and Applications(ICECAA) 2022.*

*[9]Pezhman Taherei Ghazvinei ,"Sugarcane growth prediction based on meteorological parameters using extreme learning machine and artificial neural network" , Engineering Applications Of Computational Fluid Mechanics, Vol. 12, No. 1, pp 738–749 2018*

*[10]Jaturong Som-ard,"Remote Sensing Applications in Sugarcane Cultivation" Institute of Geomatics, University of Natural Resources and Life Sciences, Vienna (BOKU), Peter-Jordan-Straße 82, 1190 Vienna, Austria , Vol 13, No. 20, 2021*

*[11]Laura Martinez- Ferror, "crop yield estimation and interpretability with guassian processes", IEEE GeoScience and Remote Sensing Letters, Vol 18, No. 12, pp 2043-2047, 2022.*

*[12]T R RAJESH et al, "Enhanced Approach for Disease Prediction in Sugarcane Crop with the support of advanced machine learning strategies", Vignan's Foundation for Science, Technology and Research , Vol 24, No. 4,ISSN: 1583-6258, pp 16805-16814 ,2022.*