DOI:10.48047/INTJECSE/V15I4.89 ISSN: 1308-5581 Vol 15, Issue 04 2023

CROP YIELD PREDICTION USING MACHINE LEARNING

Mrs. RAJESHWARI
ASSISTANT PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SREYAS INSTITUTE OF ENGINEERING AND TECHNOLOGY
bele.rajic22@gmail.com

JAHNAVI V.S.
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SREYAS INSTITUTE OF ENGINEERING AND TECHNOLOGY
j252ahnavi@gmail.com

P.JAYASREE
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SREYAS INSTITUTE OF ENGINEERING AND TECHNOLOGY
sjaya2430@gmail.com

S.DEEPTHI REDDY
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SREYAS INSTITUTE OF ENGINEERING AND TECHNOLOGY
deepthireddu0212@gmail.com

S.SAI MAHESH
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SREYAS INSTITUTE OF ENGINEERING AND TECHNOLOGY
smaheshr15@gmail.com

ABSTRACT

Agriculture growth mainly depends on diverse soil parameters, namely Nitrogen, Phosphorus, Potassium,Crop rotation, Soil moisture, pH, surface temperature and weather aspects like temperature, rainfall, etc. Technology will prove to be beneficial to agriculture which will increase crop productivity resulting in better yields to the farmer. The proposed work provides a solution for Smart Agriculture by monitoring the agricultural field which can assist the farmers in increasing productivity to a great extent. This work presents a system, in a form of a website, which uses Machine Learning techniques in order to predict the most profitable crop in the current weather and soil conditions. This system can also help in predicting the yield of the crop using weather parameter, soil parameter and historic crop yield. Thus, the work develops a system by integrating data from various sources, data analytics, prediction analysis which can improve crop yield productivity and increase the profit margins of farmer helping the mover a longer run.

INTRODUCTION

Agriculture is the backbone of many economies, and its sustainability is critical for ensuring food security globally. With the advent of technology, machine learning (ML) has emerged as a powerful tool to enhance various aspects of agriculture, including crop yield prediction. Accurate yield prediction is essential for farmers, policymakers, and stakeholders to make informed decisions regarding crop management, resource allocation, and market planning. This article provides an overview of the application of machine learning techniques in predicting crop yields and explores the existing literature to understand the advancements and challenges in this domain. Machine learning involves the development of algorithms that enable computers to learn patterns and make predictions or decisions without explicit programming. In agriculture, ML techniques are applied to analyze vast datasets, including information on weather patterns, soil quality, crop types, and historical yield data. The goal is to create models that can predict crop yields

DOI:10.48047/INTJECSE/V15I4.89 ISSN: 1308-5581 Vol 15, Issue 04 2023

accurately, providing valuable insights for optimizing agricultural practices.

Accurate crop yield prediction is crucial for several reasons. Firstly, it aids in optimizing resource allocation, as farmers can adjust the use of fertilizers, water, and other inputs based on predicted yields. Secondly, it facilitates better risk management by helping farmers anticipate potential losses due to adverse weather conditions or pest infestations. Thirdly, it contributes to market planning by providing stakeholders with information about the expected supply, influencing pricing and distribution strategies.

LITERATURE SURVEY

Traditional Approaches:

Early efforts in crop yield prediction predominantly relied on traditional statistical methods. Linear regression, time series analysis, and other statistical techniques were employed to model the relationship between various factors and crop yields. While these approaches provided valuable insights, their accuracy was often limited by the complexity of agricultural systems and the dynamic nature of environmental factors.

2. Machine Learning Techniques:

The integration of machine learning techniques has marked a paradigm shift in crop yield prediction. Numerous studies have explored the effectiveness of algorithms such as Support Vector Machines (SVM), Random Forests, Neural Networks, and Ensemble methods. These algorithms are capable of handling complex, nonlinear relationships within datasets, making them well-suited for the multifaceted nature of agriculture.

3. Data Sources:

The accuracy of machine learning models depends heavily on the quality and quantity of data used for training. Researchers have incorporated diverse data sources, including satellite imagery, climate data, soil information, and historical yield records, to build comprehensive datasets. The integration of remote sensing technology has enabled the extraction of real-time information, enhancing the precision of predictions.

4. Feature Selection and Engineering:

Feature selection and engineering play a pivotal role in improving the performance of machine learning models for crop yield prediction. Researchers have explored the identification of relevant features such as climate variables, soil characteristics, and crop management practices. Additionally, feature engineering techniques aim to create new meaningful features that contribute to better model accuracy.

5. Challenges and Limitations:

Despite the advancements, challenges persist in the application of machine learning to crop yield prediction. Issues such as data scarcity, the need for robust model interpretability, and the dynamic nature of agricultural systems pose significant hurdles. Additionally, the deployment of machine learning models in real-world agricultural settings requires addressing the constraints of computational resources and the development of user-friendly interfaces for farmers.

6. Case Studies and Success Stories:

Several case studies and success stories highlight the practical implementation of machine learning in crop yield prediction. For instance, works leveraging satellite imagery and machine learning algorithms have successfully predicted crop yields with high accuracy in specific regions. These examples demonstrate the potential for ML to revolutionize agricultural practices and contribute to sustainable and efficient farming.

Application of machine learning in predicting crop yields represents a promising frontier in agriculture. The shift from traditional statistical methods to advanced ML techniques has significantly improved the accuracy and reliability of predictions. As researchers continue to address challenges and refine models, the integration of machine learning in agriculture holds immense potential for fostering sustainable practices, mitigating risks, and ensuring global food security. The literature survey underscores the diverse approaches and methodologies employed in this field, providing a comprehensive understanding of the current landscape and future directions for research and implementation.

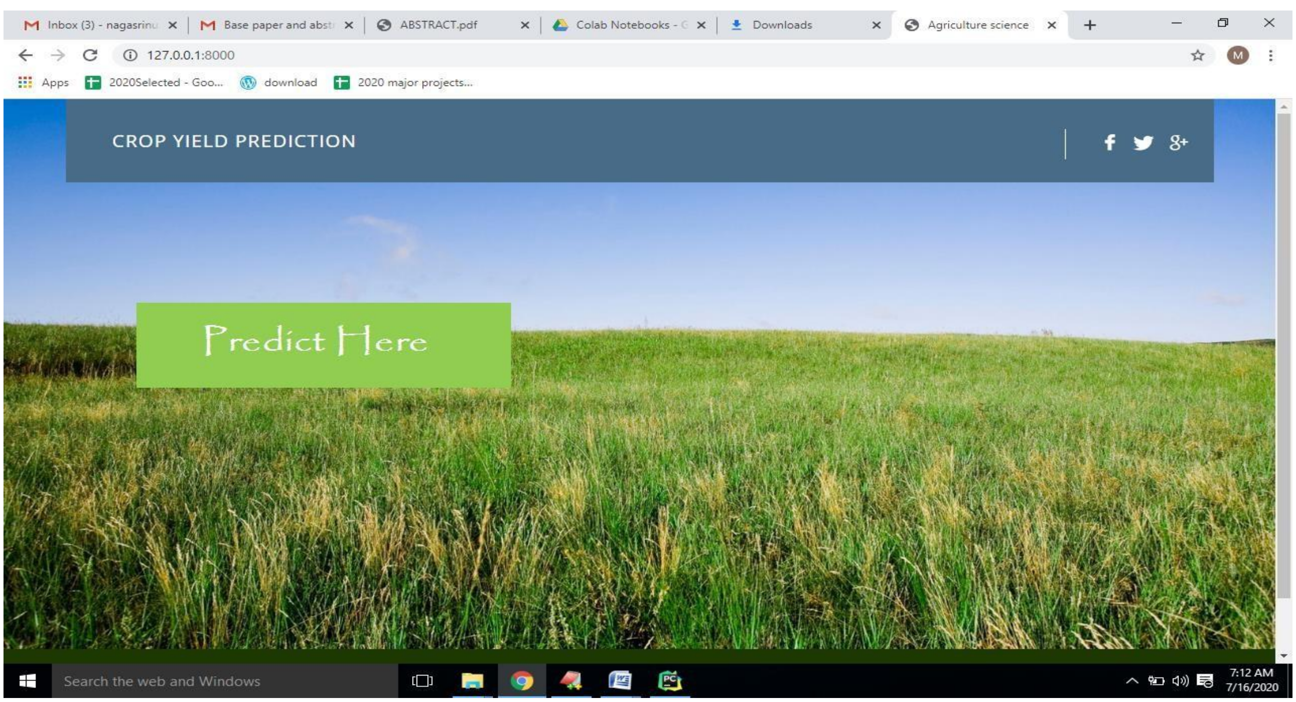
PROPOSED SYSTEM

The work mainly focuses on the basic types of crops and the irnature of yield invarious types of seasons. This work is included with a huge amount of datasets where in almostall types of crops and their wide range of behavior and

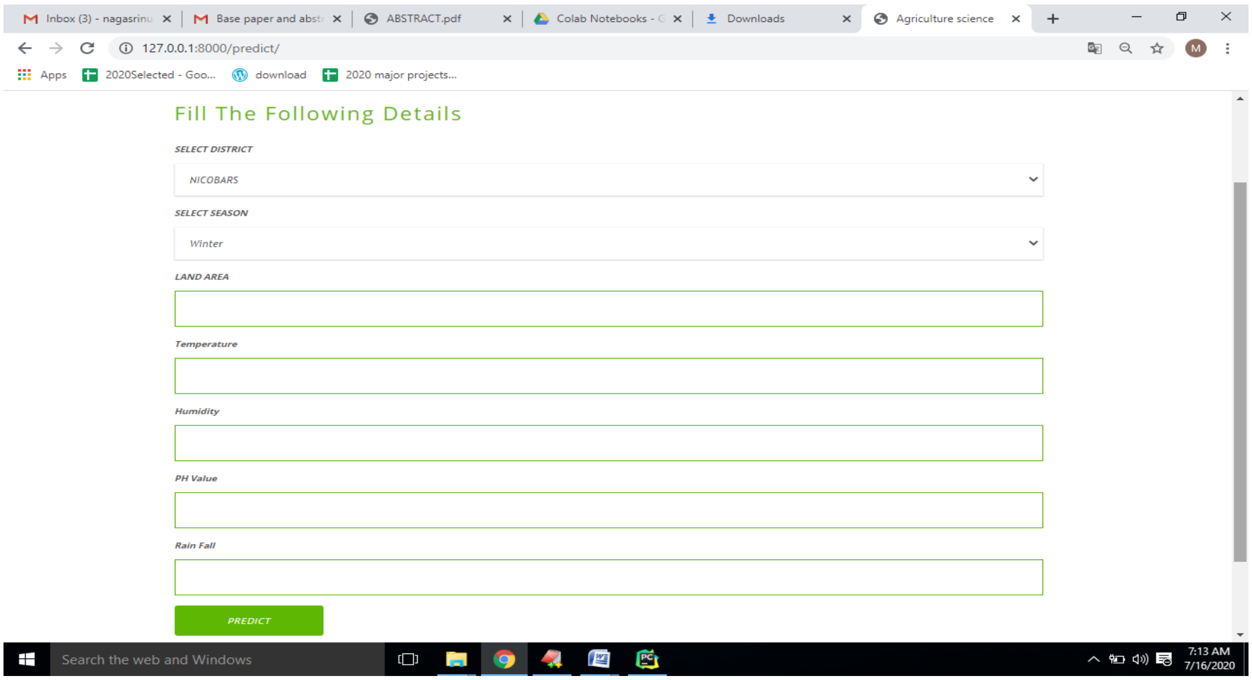
DOI:10.48047/INTJECSE/V15I4.89 ISSN: 1308-5581 Vol 15, Issue 04 2023

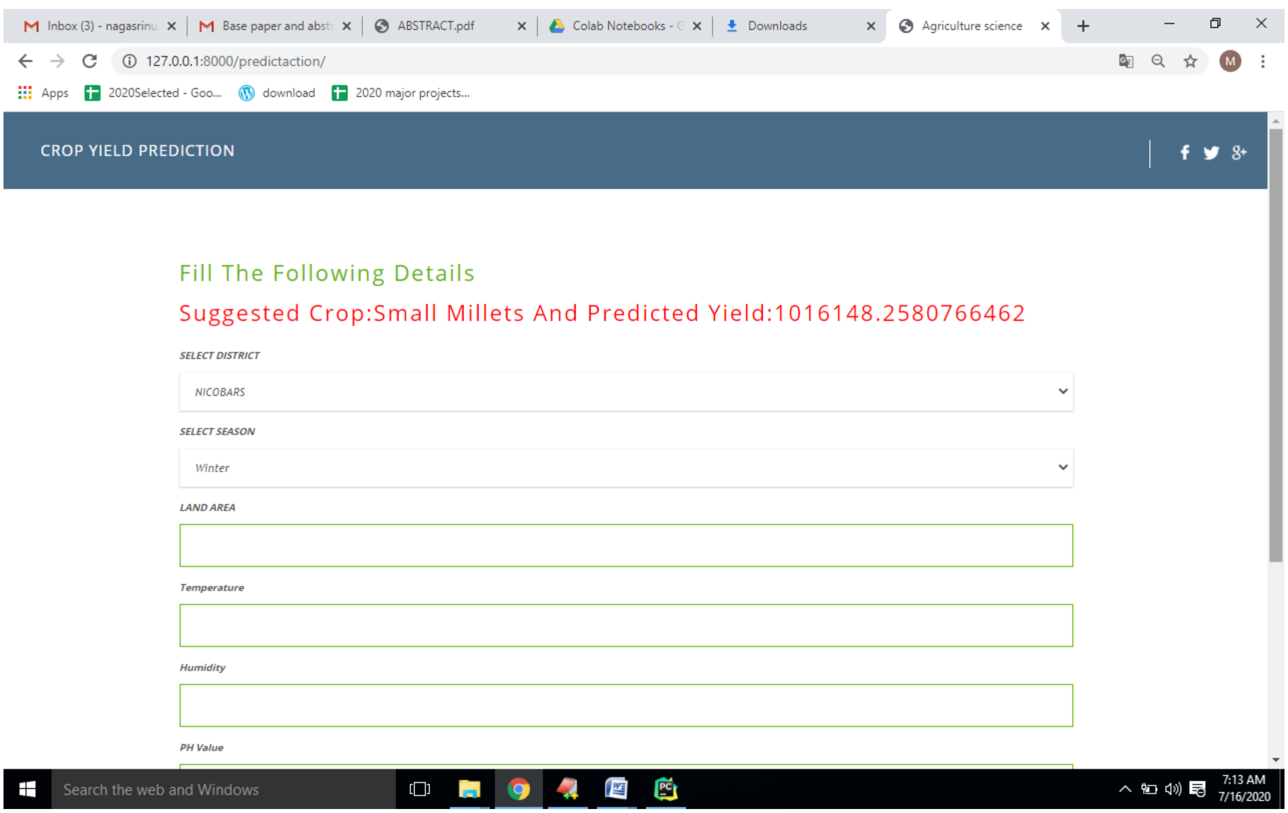
yield in various types of seasons is provided. These large amount of datasets helps in predicting the right crop with high amount of productivity. Through this, a person who does farming as their major occupation will get to know which type of plant serves them with best results. Agriculture is a business with risk which depends on climate, geography and economic factors. The main intent and objective of the work is to help the farmers choose the right crop in the right time to increase their earnings and returns on their farms. This work estimates which crop is tobeharvested in the right time andi n the right season. Basedon the requirements of the users (farmer), prediction is done and the necessaryinformationisprovided to him. Through this, they get benefitted with high returns for the amount of crop they planted and grown it with at most efforts.

Large amount of data sets are given containinginformation about types of crops, yield, soil types, seasons etc. Providing the user to select based on which field he/she wants to perform the analysis. Analyzing the data present with us based on the provided user requirement. Showing results with at most available possibilities which will enhance the chances of yield in the future. Data is composed from a different source and optimized for data sets. And the data is used to evaluate descriptively. Several abstract online outlets, like Kaggle, Google weather forestation and data government, provide the data for up to 10years in series. The data sets such as soil nature, climatic conditions and seed dataareused forth e crop prediction and better crop yields. Preprocessing the data is considered as a significant step machine learning phase.Preprocessing involves adding the missing values, the correct set of data, and extracting the functionality. Data set form is important to the process of analysis. The data collected in this step will induced in Google Colab platform in theform of python programming in order to get the desired output. Extraction of the features would reduce the data size involved to characterize a wide collection of data. The characteristics of soil, crop and weather collected from the pretreatment process establish the final training data collection. This approach selects the features based on the correlation matrix i.e. the features that has more correlation value is selected as an important predictive function for yield as shown in fig.1 In advance to this step there need to split the data into train dataset and test dataset. By applying the random forest the data is trained with available input and output data. In the test phase, th edata are tested if the accuracy of e model is satisfied. Then the new data is predicted by machine learning module.



International Journal of Early Childhood Special Education (INT-JECSE)
DOI:10.48047/INTJECSE/V15I4.89 ISSN: 1308-5581 Vol 15, Issue 04 2023





CONCLUSION

With this model, we have created a rudimentary model that is able to forecast to a certain extent. Even though the model is not perfect, we have one that can approximate to the past data pretty well. But for new data, we would require more parameter tuning. There exists great potential to improve sales forecasting accuracy in the Ecommerce domain. One good opportunity is to utilize the correlated and similar sales patterns available in a product portfolio. In this paper, we have introduced a novel demand forecasting framework based on LSTMs that exploits non-linear relationships that exist in the E-commerce business. We have used the proposed approach to forecast the sales demand by training a global model across the items available in a product assortment hierarchy. Our developments also present several systematic grouping strategies to our base model, which are in particular useful in situations where product sales are sparse. Our methodology has been evaluated on a real-world E-commerce database from Walmart.com. To demonstrate the robustness of our frame work, we have evaluated our methods on both category level and super-department level datasets. The results have shown that our methods have outperformed the state-of-the-art univariate forecasting techniques. Furthermore, the results indicate that E-commerce product hierarchies contain various cross-product demand patterns and correlations are available, and approaches to exploit this information are necessary to improve the sales forecasting accuracy in this domain.

REFERENCES

1. Hyndman,R.etal.,2008. Forecasting with Exponential Smoothing: The State Space Approach, Springer Science& Business Media.
2. Box,G.E.P.etal.,2015.Time Series Analysis: Forecasting and Control, John Wiley & Sons.
3. Box,G.E.P.&Cox,D.R.,1964. An Analysis of Transformations. Journal of the Royal Statistical Society. Series B, Statistical methodology, 26(2),pp.211252.
4. Yeo,J.etal.,2016.Browsing 2 purchase: Online Customer Model for Sales Forecasting in an E-Commerce Site. 25th Int. Conf. Comp.on World Wide Web.
5. Ramanathan,U.,2013.Supply chain collaboration for improved forecast accuracy of promotional sales. Int. Journal of Operations & Production Management.
6. Kulkarni,G.,Kannan,P.K.&Moe,W.,2012. Using online search data to forecast new product sales. Decision support systems, 52(3), pp.604611.
7. Zhao,K.&Wang,C.,2017. Sales Forecast in E-commerce using Convolutional Neural Network.ar Xiv [cs.LG].
8. Seeger,M.W.,Salinas,D.&Flunkert,V.,2016. Bayesian Intermittent Demand Forecasting for Large Inventories. In Proceedings of the 29th NIPS.
9. Snyder,R.,Ord,J.K.&Beaumont,A.,2012. Forecasting the intermittent demand for slow-moving inventories: A modelling approach.IJF, 28(2), pp.485496.
10. Zhang,G.,Patuwo,B.E.&Hu,M.Y.,1998. Forecasting with artificial neural networks: The state of the art. International journal of forecasting, 14(1),pp.3562.
11. Yan,W.,2012. Toward automatic time-series forecasting using neural networks. IEEE transactions on neural networks and learning systems,23(7), pp.10281039.
12. Zimmermann, H.-G., Tietz, C.& Grothmann, R.,2012. Forecasting with RNNs:12 Tricks. In Neural Networks. Lecture Notes in Computer Science. pp. 687707.
13. Trapero, J.R.,Kourentzes,N.&Fildes,R.,2015.Onthe identification of sales forecasting models in the presence of promotions. The Journal oftheORS.
14. Borovykh,A.,Bohte,S.&Oosterlee,C.W.,2017.Conditional Time Series Forecasting with Convolutional Neural Networks arXiv [cs.AI].
15. Flunkert,V.,Salinas,D.&Gasthaus,J.,2017. Deep AR: Probabilistic Forecasting with Autoregressive Recurrent Networks.ar Xiv [cs.AI].
16. Wen,R.etal.,2017. A Multi-Horizon Quantile Recurrent Forecaster.ar Xiv[stat.ML].