International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Agricultural Crop Recommendation, Crop Disease Detection and Price Prediction Using Machine Learning

[*]Tumma Susmitha[1], Jinkala Swathy[2], Duvva Laxmiprasanna[3]
1Computer Science & Engineering
Vasavi College of Engineering
Hyderabad, India
[*] susmitha.vce@gmail.com

[2]Computer Science & Engineering
Vasavi College of Engineering
Hyderabad, India
swathijinka231@gmail.com

[3]Computer Science & Engineering
Vasavi College of Engineering
Hyderabad, India
duvva.prasanna5@gmail.com

Abstract— India's foundation is its agriculture. With over 60% of the workforce employed and producing over 18% of the nation's GDP, it is a vital sector of the Indian economy. Although there are many ways in which we can use technology to increase product production, a farmer can only profit if he is able to sell his crops. Three laws have been passed by the Indian government to encourage the export of agricultural products across the nation. But today, we witness farmers all over the nation fighting against these regulations to protect their rights. Farmers worry that big merchants will exploit them as puppets and undercut the price at which they sell their goods. After doing a thorough analysis of the situation, we developed the concept of creating an agricultural produce application that facilitates direct communication between farmers and retailers, allows for product reviews and crop yielding rate prediction, and predicts the price of agricultural produce based on quantity produced and previous years' sales rates. Unpredictable rains, unexpected temperature decreases, and heat waves have all been brought on by the shifting climate, and the ecosystem has suffered significant harm. Thankfully, machine learning has produced useful methods for tackling international problems, such as agriculture. These climate change-related agricultural issues can be resolved by using various machine learning methods. The purpose of this piece is to Create a method to identify crop diseases and suggest crops. For both objectives, publicly accessible datasets were utilized. Regarding the crop recommendation system, feature extraction was done, and a variety of machine learning methods were used to train the dataset, including Support Vector Machine (SVM), Random Forest, Decision Tree, Logistic Regression, and Multilayer Perceptron. 99.30% accuracy was attained via the random forest algorithm.CNN architectures such as ResNet50, and EfficientNetV2 were trained and compared for the plant disease identification system. EfficientNetV2 outperformed the rest, with a high accuracy of 96.08%.

Keywords-*: GPS Navigation, Decision Tree, SVM, Multilayer Perceptron(MLP), Random Forest(RF), CNN, ResNet50, EfficientNetV2*

I. INTRODUCTION

Agriculture produced in our nation needs a strong market. Farmers find it challenging to get customers to buy their goods. India's farmers have limited options for where to sell their goods at markets. All states, with the exception of three, mandate that farm produce be marketed and sold through state-owned mandis, or retail marketplaces, where middlemen put pressure on growers to raise their profit margins. Crop Cost Forecasting, Language Interpreter, Sorting by the farmer's or customer's geographic proximity, tailoring the app to that farmer's crop and profit, etc., utilizing machine learning, deep learning algorithms, such as the Decision Tree Regression Algorithm for Price Prediction, and other methods like GPS navigation, KNN, Haversine, nearest neighbor search, load balancing, This application would be a great crop selling tool for farmers that make significant profits by eliminating middlemen and mediators entirely, provided there is market analysis and a few number of APIs for geographic proximity, among other factors. An online shopping software that satisfies all needs

for farmers to sell their goods, learn about the costs and revenues of their products, and communicate with clients directly while remaining safe in their own homes. With the help of this program, farmers may communicate with clients in their own tongue and utilize it even if they don't know much about technology, smart phones, etc. Intelligent machines have the capacity to take on challenging tasks that are hard for humans to do. It can be used in a variety of contexts, including as business, sports, and agriculture. It is capable of doing tasks including identification, prediction, and classification. This article's main goal is to create a website that tackles two pressing issues: crop disease identification and crop suggestion. This strategy will address the demands of the farming sector and the needs of the farming sector [1].These issues were addressed by training models on publicly accessible datasets, and comparing the results produced by the various models. The last five years have seen significant changes in the climate, which has had a significant impact on agriculture. Choosing crops that are unsuitable for their intended application is often the result of a lack of knowledge about scientific agriculture methods [2]. Farmers are sometimes forced to make decisions based on a limited amount of experience, which can increase their risk of making mistakes. The ineffective use of vital information, such as the pH and composition of the soil, and the early identification of plant diseases, causes the agricultural sector to incur significant losses as a direct result [3].

II. Related Work

In [1] , The framework provides a platform at the government level, including an app for Android and a website app, so farmers have multiple options to sell their crop products at different stages of the marketing chain (market, merchant, or end user). With less time and effort, farmers may utilize the platform to learn about local marketplaces, stock levels, and demand for particular crops. The web-based application will contain data on, among other things, end-user lists, complaints lists, merchant lists, farmer lists, and market specifics. As a result, government management will be enhanced.

The problems of crop recommendation and plant disease detection have been attempted to be addressed. Based on the soil, G. Chauhan and A. Chaudhary proposed a crop type[9]. They made predictions using Random Forest and Decision Trees, and the outcomes showed that a random forest classifier could be used to predict crops based on patterns in the terrain. In order to make intelligent decisions in the field of agriculture, Jose M. Cadenas et al. introduced a decision support system based on time series datasets and an inference engine[5].

In [2], This study focuses on machine learning methods for support vector regression algorithm-based crop price forecasting. Regression is a data mining technique where the crop price is determined by learning. Regression tasks will take into account activities involving classification and tasks with particular class labels. One of the inputs to the algorithm is our training dataset, which we use to identify patterns in order to compute the crop price. The algorithm receives input values from the user in the form of yield, rainfall, minimum support price, and wholesale price index. The other parameters in the algorithm are probability, the number of dataset parameters, and the new record input.

In [3], To translate a text from English into Telugu using a rule-based translation system, one must be aware of the structures of both languages. The translation process is influenced by the syntax and structure of the two languages. Managing prepositions is among the most challenging parts. Prepositions in Telugu will be understood as postpositions. The objective of this paper is to translate the given text from English to Telugu, select the appropriate postposition, and formulate the prepositional phrase.

In [4], employs both qualitative and quantitative research methods to predict the product price. The main premise of qualitative cost evaluation methodologies is to detect similarities between a new product and previous items. The similarities discovered make it easier to incorporate historical data into the current product, which lessens the need to create a cost estimate from scratch.

III. Methodology

Linear Regression

Even while it can also handle regression problems, SVM is a supervised machine learning model that is most commonly employed for dataset classification. Every data item is a data point in the n-dimensional space used by the SVM algorithm. where n denotes the number of dimensions and n-dimensional space is the number of features. Every feature value denotes a specific coordinate. The next step in the classification process is to identify the hyper-plane that best illustrates how the two classes differ from one another.

Decision Tree

In both regression and classification issues, decision trees are useful instruments. It is depicted as a tree structure, with the leaves standing in for ultimate choices. Entropy, a measure of uncertainty or disorder in a dataset, is the foundation of decision trees. On the other hand, information gain quantifies the amount of uncertainty that a particular characteristic reduces, and it is an important consideration when deciding which attribute to use as the decision node or root node.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Random Forest Classifier

An ensemble model called a random forest is created by combining several decision trees. Each and every decision tree is trained using a portion of the input data, and the results from many trees are then combined to generate an output. Random forest is a highly effective classifier because it can handle high-dimensional data and minimize overfitting.

Multi-Layer Perceptron

For classifying tasks related to classification, the multilayer perceptron (MLP) neural network is a suitable choice. As an alternative to previous classification methods like Support Vectors and the Naive Bayes Classifier, the MLP Classifier uses a neural network to perform the classification. The Naive Bayes Classifier and the Support Vector Classifier are two further classification methods.

ResNet50

Max Pool, Average Pool, and an additional convolution layer are among the 48 convolution layers in the ResNet50 model. In this construction, these layers are stacked one on top of the other. The version of ResNet currently in use has been given the name "ResNet50". This particular ResNet model, which has been the subject of much research, was built using 3.8 x 109 different floating point operations. This method works well for computer vision applications like object detection, object location, and image classification. apps that utilize computer vision are sometimes referred to as computer vision apps.

EfficientNetV2

A kind of The CNN network known as EfficientNetV2 increases the depth and width in every dimension and employs a compound coefficient for resolution. When the size of the input image increases, this model requires more layers in order to extend the receptive field. The purpose of the inverted bottleneck residual blocks is to maximize effectiveness and efficiency. They nevertheless achieve equivalent accuracy but with fewer parameters than typical residual blocks. Efficient Net can achieve state-of-the-art accuracy on a range of picture classification applications because to the compound scaling algorithm. It is an effective method for growing neural networks without compromising precision or effectiveness.

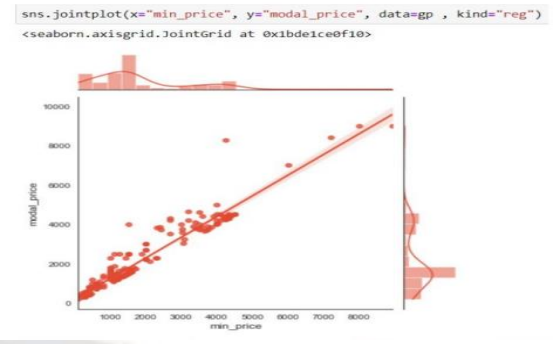

Fig. 1. Minimum vs Modal Comparison

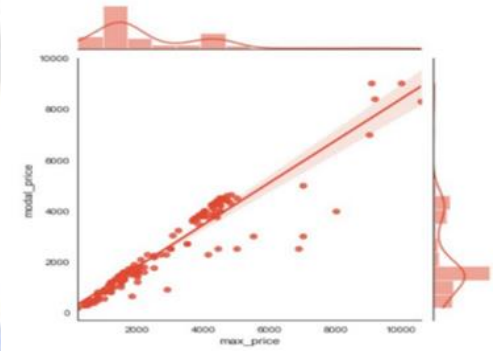

Fig. 2. Maximum vs Modal Comparison

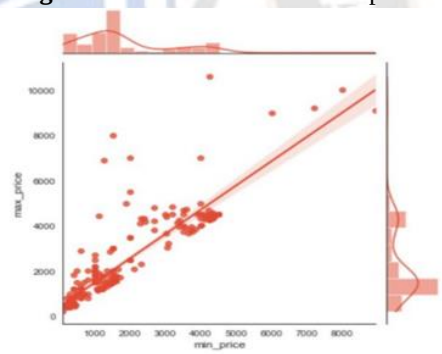

Fig. 3. Minimum vs Maximum Comparison

Crop Recommendation System

When choosing the crop for the season, farmers can make great use of crop prediction. Therefore, Crop is estimated using values appropriate for the local conditions. The Random Forest Classifier is utilized for these. As a result of its superior accuracy.

The amount of nitrogen, phosphate, and potassium in the soil can be ascertained by measuring its NPK value. These nutrients are critically necessary for the development of plants.

- An essential factor in the growth of plants is the surrounding temperature. Different temperatures are needed by different plants.
- The pH value of the soil can be used to determine how acidic it is.

- Different plant species prefer different pH levels of water.
- Every variety of plant may require a different amount of water.

Crop Price Prediction

A handy tool for dynamically obtaining product pricing in an application is Crop Price Prediction. To do this, a model can be trained to determine the crop's true price based on market rates. Min-Max Models can be utilized for pricing calculations. Here, the various algorithms are taken into account and contrasted with the most effective algorithm for price prediction.

Crop Disease Identification

There are 70,000 photos in the dataset that is made accessible for this assignment. Given the difficulty of feature recognition, image classification poses a significant challenge to traditional machine learning algorithms. Because convolution neural networks (CNNs) can identify features during training and offer more accuracy than conventional machine learning techniques, they are utilized for challenging image categorization applications. For classification, we suggested three CNN models: VGG16, ResNet50, and EfficientNetV2.

IV. Dataset Description

The data is sourced from the Kaggle website, which acts as the input source for the analysis. The data extraction process involves retrieving the raw data from the source and storing it in a dataset.

CONCLUSION AND FUTURE WORK

Thus, we draw the conclusion that our program supports farmers' economic growth and gives them access to individualized product statistics data. The farmer benefits from the application's reduction of manual labor, ease of product sales, and market trend data. Farmers benefit greatly from the application's usage of machine learning algorithms to forecast prices and crops. This study will help find a solution to the massive harvesting problem that the agriculture industry is now facing. Our analysis comparing five distinct algorithms demonstrates the efficiency of the random forest model. Random Forest's total accuracy is 99.3%. A comparison of ResNet50, VGG16, and EfficientNetV2 was offered for classification. EfficientNetV2, with an accuracy of 96.08%, outperformed VGG16 and ResNet50. To increase classification accuracy, ViT and CoAtNet models can take the place of future CNN. To enhance multiclass disease categorization, the plant leaves need to be gathered from various geographical locations.

References

[1] Kaneda, Y., Shibata, S., & Mineno, H. (2017). Multi-modal sliding window-based support vector regression for predicting plant water stress. Knowledge-Based Systems, 134, 135-148.

[2] Mrs. ManishaBhende,Ms. Mohini.S.Avatade,Mrs.Suvarna,Mrs. PoojaMishra,Ms. Pooja Prasad and Mr. ShubhamShewalkar, “Digital Market : E-Commerce Application For Farmers”,2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).

[3] Kulkarni, P., Karwande, A., Kolhe, T., Kamble, S., Joshi, A. and Wyawahare, M., 2021. Plant Disease Detection Using Image Processing and Machine Learning.ar XivpreprintarXiv:2106.10698

[4] Rohith R, Vishnu R, Kishore A and DeebanChakkarawarthi,”Crop Price Prediction and Forecasting System using Supervised Machine Learning Algorithms” 2020 International Journal of Advanced Research in Computer and Communication Engineering(IJARCCE).

[5] Cadenas, J. M., Garrido, M. C., & Martínez-España, R. (2023). A Methodology Based on Machine Learning and Soft Computing to Design More Sustainable Agriculture Systems. Sensors, 23(6), 3038.

[6] LakkanaRuekkasaem and MontaleeSasananan, “Forecasting agricultural products prices using time series methods for crop planning”, 2018 International Journal of Mechanical Engineering and Technology (IJMET).

[7] Susmitha, T., Prashanthi, T., Mishra, R.K. (2023). Quality-Produced Agricultural Crop Price Prediction Using Machine Learning. In: Seetha, M., Peddoju, S.K., Pendyala, V., Chakravarthy, V.V.S.S.S. (eds) Intelligent Computing and Communication. ICICC 2022. Advances in Intelligent Systems and Computing, vol 1447. Springer, Singapore. https://doi.org/10.1007/978-981-99-1588-0_18.

[8] Keerthi Lingam, E Rama Lakshmi and L Ravi Theja, “ Rule Based Machine Translation from English to Telugu with emphasis on prepositions”, First International Conference on Networks & Soft Computing IEEE.

[9] HassanSM,MajiAK,Jasi´nskiM,LeonowiczZ,Jasi´nskaE. Identification of Plant-Leaf Diseases Using CNN and TransferLearning Approach. Electronics. 2021; 10(12):1388

[10] G. Chauhan and A. Chaudhary, ”Crop Recommendation System using Machine Learning Algorithms,” 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART), 2021, pp. 109-112, doi: 10.1109/SMART52563.2021.9676210.

[11] BijayaAdhikari, ParikshitSondhi, Wenke Zhang, Mohit Sharma and B. AdityaPrakash,”Mining E-Commerce Query Relations using Customer Interaction Networks”,Department of Computer Science, Virginia Tech ,WalmartLabs.

[12] Prashanthi, T., Susmitha, T., Mishra, R.K. (2023). Crop Disease Detection Using NLP and Deep Learning. In: Kumar, A., Ghinea, G., Merugu, S. (eds) Proceedings of the 2nd International Conference on Cognitive and Intelligent Computing. ICCIC 2022. Cognitive Science and Technology. Springer, Singapore. https://doi.org/10.1007/978-981-99-2742-5_21

[13] Jian S Dai,StavroulaBalabani and LakmalSeneviratne,”Product Cost Estimation: Technique Classification and Methodology Review”Journal of Manufacturing Science and Engineering.

[14] UrveshRathod,Prof. ArunaPavate and VaibhavPatil,”Product Rank Based Search Engine for E-Commerce” 2018 3rd International Conference for Convergence in Technology (I2CT) IEEE.

[15] Sharada Prasanna Mohanty, David Hughes, Marcel Salathe, 2016, Using Deep Learning for Image-Based Plant Disease Detection.arXiv preprint arXiv:2106.10698.