iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 9 **Issue:** IV **Month of publication:** April 2021

**DOI:** https://doi.org/10.22214/ijraset.2021.33922

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429*
*Volume 9 Issue IV Apr 2021- Available at www.ijraset.com*

# Solving Market Instability by Forecasting Vegetables Price using Machine Learning Algorithms

Mohanaprasath S[1], Moniswaran V[2], Prasanna S[3], Dr. C. Rajan[4]
[1, 2, 3]*Department of Information Technology, K.S. Rangasamy College of Technology*
[4]*Professor, M.E., Ph.D, Department of Information Technology, K.S. Rangasamy College of Technology, Tiruchengode - 637 215.*

***Abstract:*** ***India being a horticulture nation, its economy prevalently relies upon horticulture yield development and agro-industry items. In onion price prediction information mining plays a major role. Yield forecast is a significant issue in horticultural.Dataset with product name, particular date, price and location are the parameters mandatory for prediction of onion price in future. Initially, dataset with irrelevant data, noisy information are removed by preprocessing then feature extraction is implemented which extract required parameter from a large dataset and undergone for classification. Machine learning algorithm is used for efficient classification of dataset which leads to accurate prediction of onion price in future. One of the best algorithms is KNN classifier is used for accurate prediction. From the trained dataset, input is given it will process and predicts the price of onion in future. Hence our system achieves better result compared to other existing approaches that are present currently.***

***Keywords: price prediction, machine learning, Random forest, accuracy.***

## I. INTRODUCTION

Agricultural sector has a significant task to carry out in creating Indian economy, however the issues looked by ranchers includes not getting the normal benefit for the vegetables developed. Thus, with the assistance of irregular backwoods calculation and AI procedures vegetable value forecast should be possible. The value of the vegetable increments when it is solid. In India, there are a few different ways to expand the financial development of horticulture.

### *A. Weather Role In Price Prediction*

A seasonal change is one of the principle elements to choose the expansion and reduction in the yield of the harvest. They influence the economy of the two individuals and country. For tackling this issue making an application to foresee the pace of the vegetables by the climate projection utilizing AI was useful. In nowadays, occasional changes are more normal issue on account of Global Warming, this makes harm the development of plants. This difference in season influences the plants in their digestion, so they don't give sufficient collect of vegetables to individuals true to form. Along these lines, the pace of vegetables arrive at statures with the goal that it might influence individuals' typical life. Simultaneously sometimes when the vegetables rate diminishes it influences the rancher. To settle this, the meteorological forecast would assist with foreseeing the pace of the vegetable, how long it might exist and the amount they can broaden their cost. This may exhort an insurance for the Government and recommend to take further exercises.

### *B. Role Of Data Mining In Price Prediction*

Information mining in horticulture is a novel examination field. Ranchers are collecting vegetables and yields as well as gathering enormous measure of information. Information mining gives the technique to change these information into helpful data for dynamic. Vegetable value changes quick and flimsy which has extraordinary effect in our every day life. Vegetable cost has properties like high nonlinear and high clamor. In this way, it is difficult to foresee the vegetable cost. Information mining order procedures can be utilized to build up an imaginative model to anticipate the market cost of individual item. Value expectation is profoundly helpful in agribusiness for guaging the market cost for the individual products and furthermore helpful for ranchers to design their yield development exercises so they could get more cost on the lookout. Government can utilize the market gauge cost for arranging and execution of agribusiness advancement projects to balance out the market cost for the individual product. Government can likewise take choice if to permit to fare and import of individual products. Buyers can utilize this value forecast for their every day way of life arranging. This creative application isn't just valuable for ranchers and purchasers yet additionally helpful for agribusiness arranging, outlining polices and plans in farming and market arranging. By utilizing these prescient models, the public authority can do rural advancement wanting to settle the separate item costs. The recommended plot applies AI and conjecture calculations, for example, choice tree calculations to characterize data examples and afterward measure them as indicated by input conditions.

## II. LITERATURE SURVEY

Gangasagar HL et.al (2020), presents serious issue when the yield does not merit the cost and when ignorant of promoting cost in regards to trim specifically group. This model includes forecast of yield cost and prescribing best spot close to the rancher, to sell his produce for a decent benefit.

The way toward estimating crop cost is truly difficult, since it includes numerous boundaries. Value forecast should be possible by utilizing various calculations like SVR, Multi Linear Regression, Decision Tree Regressor, Random Forest Regressor, among which Random Forest Regressor was dissected to have a decent precision with negligible blunder.

M.Subhasree and Mrs.C.ArunPriya (2016), portrays foreseeing the vegetable cost is fundamental in horticulture area for compelling dynamic. This anticipating task is very troublesome. Neural organization is self-adjust and has astounding learning capacity and used to address assortment of errands that are complex. This model is utilized to anticipate the following day cost of vegetable utilizing the past cost of time arrangement information.

The three AI calculations are consolidated in this work in particular Radial premise work, back engendering neural organization and hereditary based neural organization are looked at. The models are surveyed and it is closed from the inferred precision that the exhibition of hereditary based neural organization is better compared to back spread neural organization and spiral premise work and improves the exactness level of vegetable value forecast.

P.S.Keerthana et.al (2020), talks about everyday life each individual buckle down for Food and Shelter. Food needs numerous fixings like Cereals, Pulses, Vegetables, Fruits and Millet. Yet, individuals endure a great deal when there is a limit expansion in pace of those fixings. The general population as well as the ranchers are dealing with a similar issue when there is an abatement in the rate.

Forecast of pace of the vegetable is troublesome in light of the fact that, the yield may diminish or increment whenever. The digestion of the harvest decides by storm. To deal with the circumstance there ought to be an application for forecast of the pace of vegetables. In this paper, the proposed framework is intended to clarify making an application which predicts the pace of Vegetables. After forecast of the rate the notice is sent by means of Mail to the more significant position authority. This gave a prudent plan to deal with the lessen in vegetable rate by the Government. We have utilized FRNN calculation for foreseeing the Weather.

G. M. Nasira and N. Hemageetha (2012), depicts the agrarian area needs more help for its improvement in non-industrial nations like India. Value expectation helps the ranchers and furthermore Government to settle on viable choice. In light of the intricacy of vegetable value forecast, utilizing the qualities of neural organizations, for example, self-adjust, study and high adaptation to internal failure, to develop the model of Backpropagation neural organization to foresee vegetable cost. A forecast model was set up by applying the neural organization. Accepting tomato for instance, the boundaries of the model are dissected through try. Toward the finish of the consequence of Backpropagation neural organization shows supreme blunder level of month to month and week by week vegetable value expectation and investigate the precision level of the value forecast.

Changshou Luo et.al (2011), presents the hypothesis and development strategies for four models are introduced for foreseeing the vegetable market value, which are BP neural organization model, the neural organization model dependent on hereditary calculation, RBF neural organization model and a coordinated forecast model dependent on the three models above. The four models are utilized to foresee the Lentinusedodes cost for Beijing Xinfadi discount market. A sum of 84 records gathered somewhere in the range of 2003 and 2009 were taken care of into the four models for preparing and testing. In outline, the foreseeing capacity of BP neural organization model is the most noticeably awful. The neural organization model dependent on hereditary calculation was by and large more precise than RBF neural organization model. The incorporated expectation model has the best outcomes.

Md. Mehedi Hasan et.al (2020), presents Price is the critical factor in monetary exercises. Unforeseen vacillation in cost is the indication of market flimsiness. These days Machine learning gives tremendous procedures to estimate cost of items to adapt up to showcase precariousness. In this paper, we investigate the use of AI way to deal with gauge the cost of onion. The estimate depends on the information gathered from Ministry of Agriculture, Bangladesh.

For making expectation we utilized AI calculations for example K-Nearest Neighbor (KNN), Naïve Bayes, Decision Tree, Neural Network (NN), Support Vector Machine (SVM). At that point we evaluated and contrasted our methods with discover which procedure gives the best exhibition in term of precision. We discover the entirety of our procedures give practically equivalent to execution. By previously mentioned strategies we try to group whether the cost of onion would be best (low), prudent (mid), costly (high).

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429*
*Volume 9 Issue IV Apr 2021- Available at www.ijraset.com*

## III. PROPOSED SYSTEM

### *A. Introduction*

By collecting large number of dataset regarding onion price details predicting onion price is proposed in our work. Initially large dataset has been gathered and preprocessed to remove irrelevant details and noisy data in the dataset. Once preprocessed data is undergone for the feature extraction process here most required field for our prediction is extracted and unwanted field gets removed. Feature extraction is an important process because it drastically affects the performance if more number of fields are included it will reflects in prediction with minimum performance. Here efficient feature extraction is implemented to attain better result. Once preprocessed and feature extracted dataset is utilized for prediction, RF classifier is used for prediction with improved performance which increases prediction accuracy. From our base work when data usage rate increases in RF average performance is not reduced hence by utilizing feature extraction our proposed method surely attains maximum performance.

### *B. Working Of Proposed Work*

Initially Dataset collected regarding various vegetable price is loaded in our system. These datasets are preprocessed here irrelevant data and noisy data are removed from it. This pruned data is next inputted to feature extraction. In this step mandatory fields alone extracted from a large data set which helps in accuracy achievement in results. Once feature extracted dataset in classified using Random Forest classifier it is a tree structure algorithm based on iterated condition verification various condition are checked and tree has been formed which results in accurate classification of price dataset. Dataset is given as input and it leads to prediction of output.

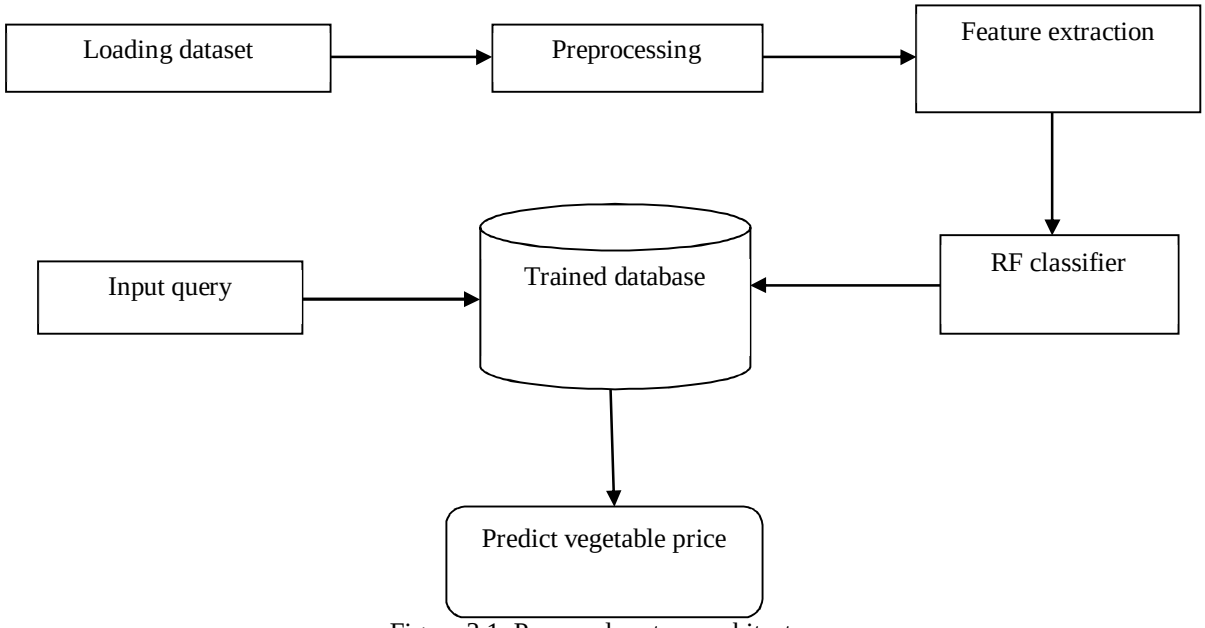



Figure 3.1: Proposed system architecture

### *C. Implementation Of Proposed Work*

1) *Data Gathering Module:* Interviewing with domain experts, identification of influencing factors and delivery of historical data from key stakeholders are the processes initially taken for data gathering. Data from various online website were searched and based on our required parameter a selective dataset with n number of data has been chosen for price prediction process. The process taken in future are cleaning and transforming the downloaded data into a machine-readable format, developing multiple price prediction algorithms. Testing the output of the algorithms and short listing best performing algorithms and then Deploy the algorithm for live testing over a period and further improvement using key learning's.
2) *Preprocessing Module:* Data preprocessing is the second step and it contains two steps. Original dataset can contain lots of missing values so initially all these should be removed. Missing values are denoted by a dot in the dataset and their presence can deteriorate the value of entire data and it can reduce the performance. So, to solve this problem we replace these values with large negative values which will be treated as outliers by the model. Generating the class labels is the second step. Since we are using a supervised learning method, for each entry in the dataset there should be a class label which is created during the preprocessing step.

3) *Feature Extraction Module:* Feature extraction is a piece of the dimensionality decrease measure, in which, an underlying arrangement of the crude information is partitioned and diminished to more sensible gatherings. So when you need to deal with it will be simpler. The main quality of these enormous informational collections is that they have countless factors. These factors require a great deal of figuring assets to deal with them. So Feature extraction assists with getting the best element from those huge informational indexes by select and consolidate factors into highlights, in this way, adequately lessening the measure of information. These highlights are not difficult to measure, yet ready to portray the genuine informational index with the precision and innovation. In our vegetable dataset there will be more number of unessential boundaries that are eliminated successfully and obligatory boundaries are removed in this module.
4) *Classification Module:* In our work Random forest algorithm is used to predict price of vegetable in accurate way. Random forests are an ensemble learning algorithm that can be used for classification, that is predicting a categorical response variable and they can also be used for regression which involves predicting a continuous response variable. Random forest regression and classification models fit an ensemble of decision tree models to a set of data. For each tree, the data are recursively split into more homogenous units, which are commonly referred to as nodes, in order to improve the predictability of the response variable. Split points are based on values of predictor variables. Thus, variables used to split the data are considered important explanatory variables. Random forests fit separate decision trees to a predefined number of bootstrapped data sets. The predicted value of a categorical response is the mode of the classes from all the individual fitted decision trees, and the predicted value of a continuous response is the mean fitted response from all the individual trees that resulted from each bootstrapped sample.

## IV. RESULT AND DISCUSSION

In this section, result of our proposed work is discussed and it comparison with other available algorithms are shown and discussed clearly. Initially dataset loading is shown in below figure 4.1.


Figure 4.1: loading datasets

The below figure shows prediction of various algorithm and its results are discussed. First result represent Random Forest result, followed by linear Regression and Support Vector machine. Among these three Random Forest achieves better results compared to other algorithms.

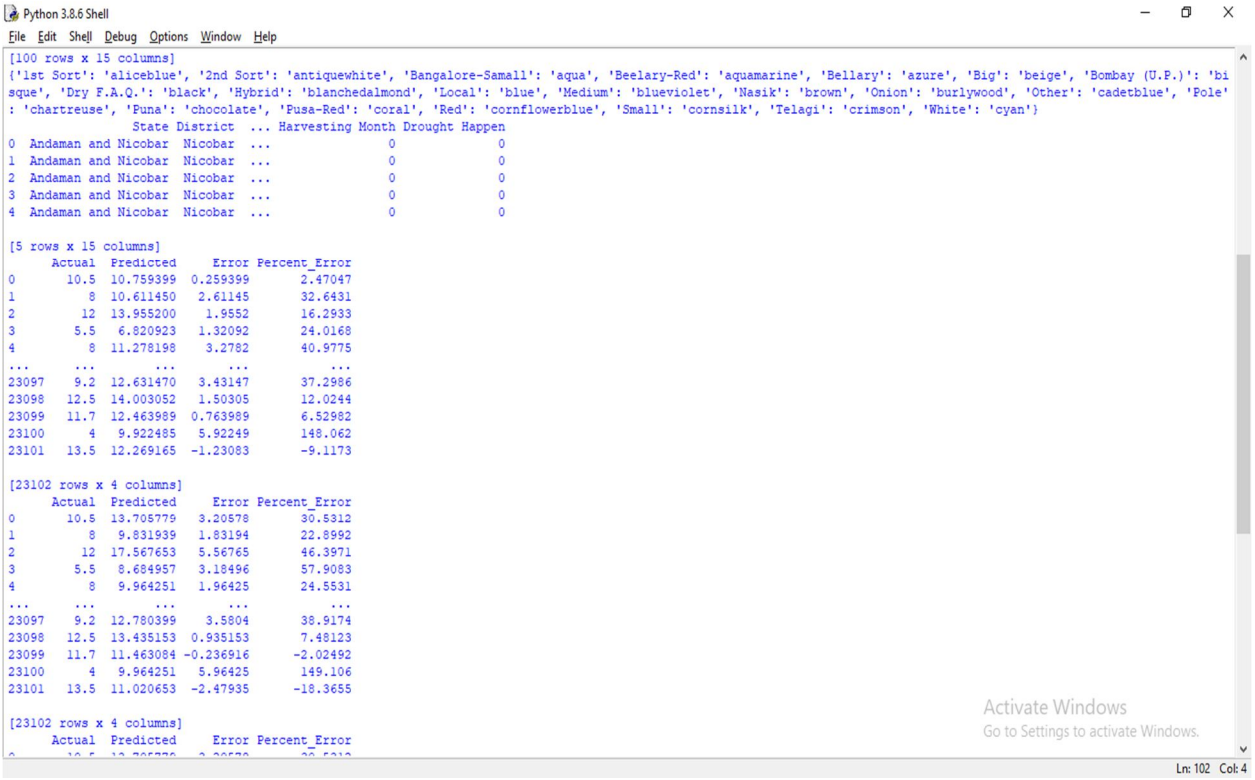



Figure 4.2: Random Forest and linear regression result accuracy

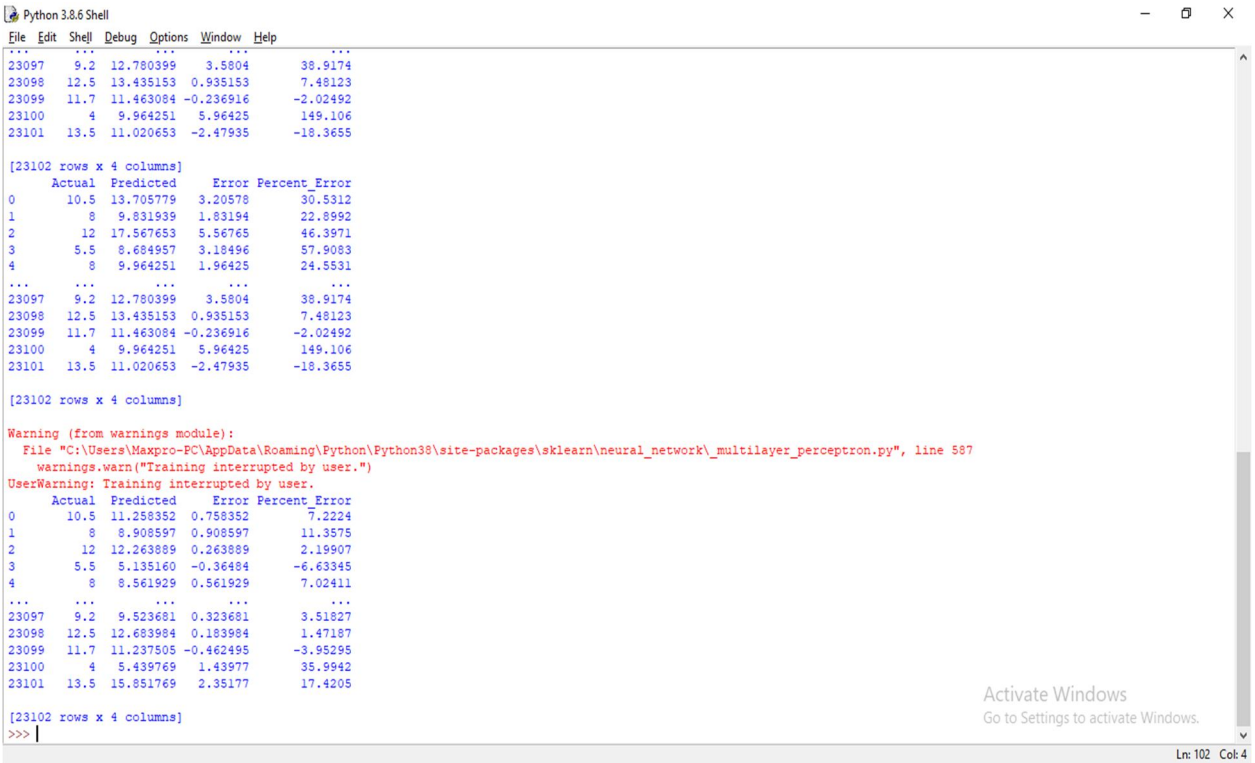



Figure 4.3: SVM accuracy results

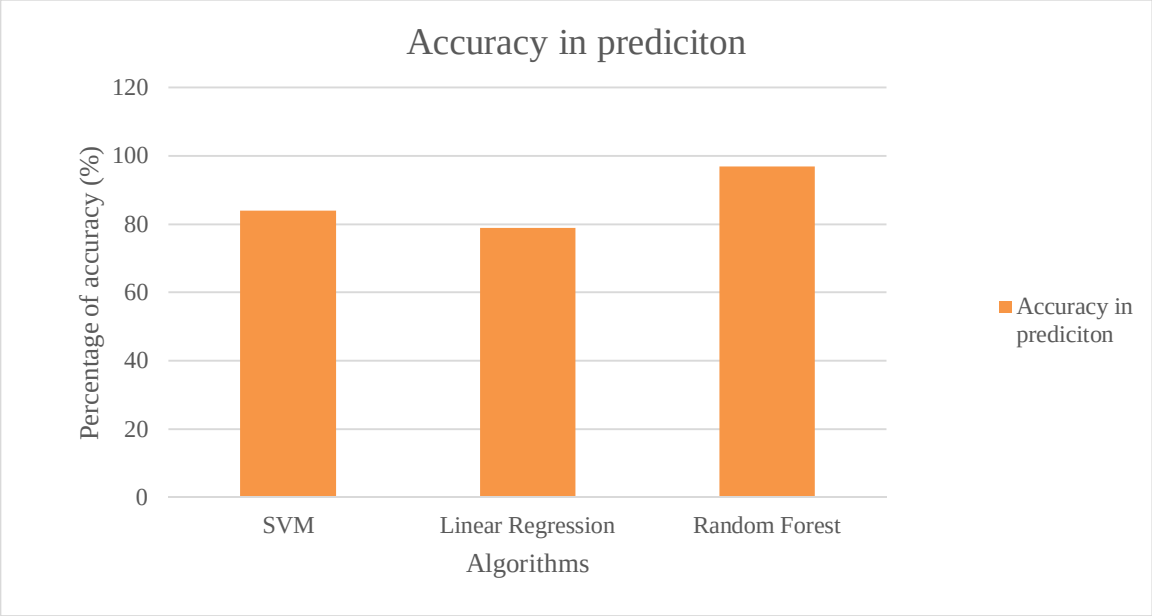



Figure 4.4: accuracy in prediction

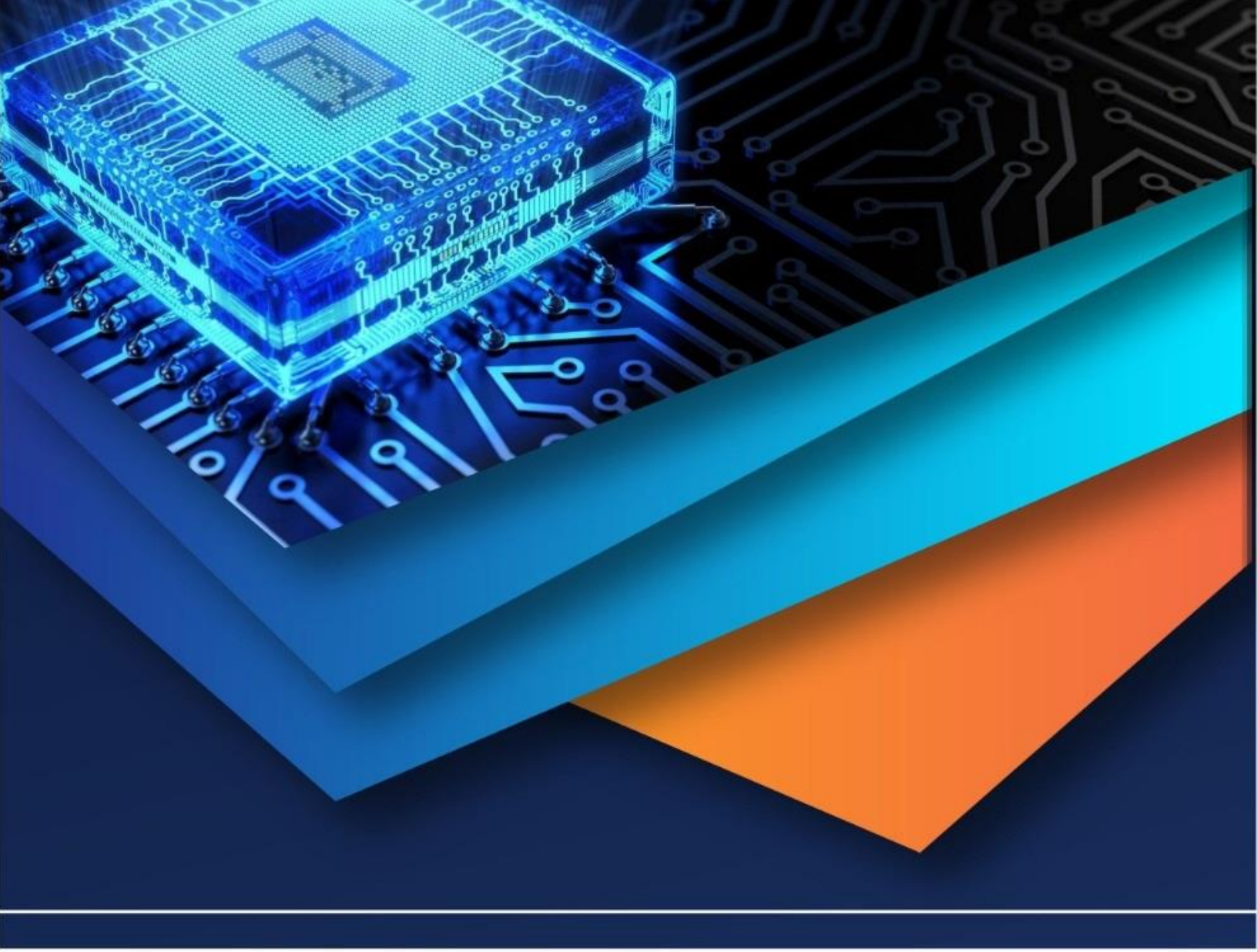
Finally the prediction of price is shown in graph is included below. Accuracy of our algorithm compared to other algorithms is discussed briefly. SVM achieves 84% of accuracy, Linear regression achieves 79% of accuracy and random forest our proposed method achieves 96% of accuracy in prediction.

## V. CONCLUSION

The main concept of the proposed application is to predict the vegetable rate. By this application we can able to control the increase and decrease in the rate of vegetables. This will be helpful for people, farmer and Government to take precautionary act for controlling the rate of vegetable. It also generates report of the cost predicted by the calculations made by the application, and send notifications. Finally we have a model by which we can expectedly forecast the future price of vegetable. Depending on this forecasting price we can calculate the demand and supply of vegetable, as we know demand-supply plays the main role in market equilibrium state. Now if we can calculate the future demand-supply of vegetable based on this prediction we can maintain equilibrium state in vegetable market which will help us to remove vegetable market instability. The main limitation of our work is unusual behavior of data and low number of records.

## REFERENCES

[1] Gangasagar HL, Jovin Dsouza, Bhagyashree B Yargal, Arun Kumar SV, AnuradhaBadage, “Crop Price Prediction Using Machine Learning Algorithms” International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)Volume 9, Issue 10, October 2020.
[2] M.Subhasree, Mrs.C.ArunPriya, “FORECASTING VEGETABLE PRICE USING TIME SERIES DATA” International Journal of Advanced Research (2016), Volume 3, Issue 535-541.
[3] P.S.Keerthana, B.Keerthika, R.Shalini, Mr.M.Saravanakarthikeyan, “Prediction of Vegetable Cost Based on Weather Condition Using Fixate Recurrent Neural Networks (FRNN)” International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 07 Issue: 03 | Mar 2020.
[4] G. M. Nasira and N. Hemageetha, “Vegetable Price Prediction Using DataMining Classification Technique” Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering , March 21-23, 2012.
[5] Changshou Luo, Qingfeng Wei, Liying Zhou,Junfeng Zhang, and Sufen Sun, “Prediction of Vegetable Price Based on Neural Network and Genetic Algorithm” IFIP International Federation for Information Processing 2011.
[6] Md. Mehedi Hasan, Muslima TuzZahara, Md. MahamudunnobiSykot, Rubaiya Hafiz and Mohd. Saifuzzaman, “Solving Onion Market Instability by Forecasting Onion Price Using Machine Learning Approach” 2020 International Conference on Computational Performance Evaluation (ComPE) North-Eastern Hill University, Shillong, Meghalaya, India. Jul 2-4, 2020.
[7] W. Huang, Y. Nakamori, S.-Y. J. C. Wang, and o. research, "Forecasting stock market movement direction with support vector machine," vol. 32, no. 10, pp. 2513-2522, 2005.
[8] M. Rafieisakhaei, B. Barazandeh, and M. Tarrahi, "Analysis of supply and demand dynamics to predict oil market trends: A case study of 2015 price data," in SPE/IAEE Hydrocarbon Economics and Evaluation Symposium, 2016: Society of Petroleum Engineers.
[9] N. Gandhi, L. J. Armstrong, O. Petkar, and A. K. Tripathy, "Rice crop yield prediction in India using support vector machines," in 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2016, pp. 1-5: IEEE
[10] J. M. Keller, M. R. Gray, J. A. J. I. t. o. s. Givens, man,, and cybernetics, "A fuzzy k-nearest neighbor algorithm," no. 4, pp. 580-585, 1985.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY