INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

IJARIIT

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 8, Issue 3 - V8I3-1453)
Available online at: https://www.ijariit.com

Pests Prediction and Detection of Disease Spreading Frequency in Native crops using Machine Learning Technique

Impana V.
impanavgowda1@gmail.com
R. V. College of Engineering, Bengaluru, Karnataka

Vanishree K.
vanishreek@rvce.edu.in
R. V. College of Engineering, Bengaluru, Karnataka

Hemanth Kumar
hemanth.a@brillio.com
Brillio, Karnataka, Bengaluru

Sumit Atram
sumitsk1522@gmail.com
Rubixe, Karnataka, Bengaluru

ABSTRACT

Disease and pest management models can generate information about the use of pesticides only when needed, reduce costs and impact the environment. Machine learning algorithms can be used to develop models used in disease and pest warning systems to improve the effectiveness of chemical control of coffee tree pests. Algorithms tested for predictable events (a) Multi-Line Detachment (RLM); (b) K-nearest neighbor regressionrs (KNN); (c) Random forest regressors (RFT); (d) Artificial neural networks. Pearson correlation analysis shows three different periods of 1-10 days (1-10 days before incident assessment), 11-20 days, 21-30 days, with a single correlation between climate change and infection rates. Establish. Days, hot and humid, weather conditions above 80% show a strong correlation with disease. It has a negative relationship with precipitation, and as precipitation increases, insect resistance decreases. Machine learning algorithms are used to predict diseases and pests.

Keywords*: Pearson correlation, MultipleLinear regression, K Neighbors Regressor (KNN) Random ForestRegressor (RFT), and Artificial Neural Networks*

1. INTRODUCTION

The production and quality of small grains are greatly affected by diseases and pests, which depends on the weather. The most common strategy for controlling these diseases and pests is to install foliar fungicides and pesticides, keeping in mind how difficult the environment can be. There are many challenges facing modern Agriculture.

Today, the agricultural sector has evolved into a highly competitive and globally competitive industry. Therefore, farmers and other stakeholders need to consider regional and indigenous climate, environmental and political factors to ensure a sustainable and productive economy. Expected obvious global climate change, fluctuating precipitation patterns, warming, drought, or increased frequency and duration of bad weather events threaten shared production areas and pose new risks and uncertainties to global crops.

Recognizing signs of damage to a variety of herbivorous insects is important for effective, efficient and economical management. Insect-resistant pests cause direct or indirect crop damage, ensuring food, its growth, and future generations. Insects attack different parts of theplant: roots, stems, bark, leaves, shoots, flowers and fruits.

Based totally on location and signs of damage, bugs may be divided into distinctive. The procedure of diagnosing plant health issues with none special laboratory device is known as field diagnosis.

Farmers often collect the wrong part of the harvest. Or the sample is too good to take away. There may be a need to visit the fields to see new signs and learn more about pests. Fearing pests and diseases, farmers applypesticides and fertilizers throughout the farm in the same way. This can damage the soil and crops.. The aim of this project is to enable the farmer to spray a limited amount of pesticides and fertilizer on the target area where there are pests and diseases or possibly future attacks. This helps the farmers primarily to prevent any such attacks on his farm andto eliminate them if any by spraying them in moderation and without contaminating the soil and other parts of the crop.

2. RELATED WORK

Author Wu et al., Suggests a Vector Regression (SVR) Support method for predicting travel time [1] using traffic data provided by the Intelligent Transportation Web Service Project (ITWS). As SVR has the advantages of having a high standard of normal operation and ensures the minimum of any set of training data. Here the SVR is compared to other basic models. Some common basic methods such as Current Travel Prediction Method and History Definition Method are used to compare SVR predictive methods.

In one of the papers [2], the author uses a fixed time travel taxi prediction using the XGBoost algorithm. Data used for the New York Taxi Commission taxi and Limousine taxi route under the Freedom of Information Act (FOIL). Web Service Project (ITWS). As SVR has the advantages of having a high standard of normal operation and ensures the minimum of any set of training data.

The data set contains millions of records including geo-coordinate passenger downloads, geocoordinate time stamps, download and departure time stamps, medal id, driver id, trip length (seconds), trip distance (miles), fair value, tax rates and much more. features related to taxi services. Since 70% of the data is considered an internal trip, after pre-processing data and critical travel filtering, the XGBoost algorithm is used to estimate travel time. In time). the XGBoost regression predictor model series surpasses even outsiders.

Algorithms for researching pesticides and coffee arabica diseases [3] Lucas Eduardo, Jose Reinaldo, Glauco de Souza, Cicero Teixeira -Coffee and pest control models developed in this work provide better diagnostic support over the years. of fruits with high and low yields.

A good model for modeling crop illness uses machine learning algorithms [4] Nilima Ashtankar-This paper uses raw data from the WSN sector and algorithms for crop and pest illnesses. It focuses on the previous dataset used to predict the pattern. Mechanical Study of Plant Pathogens and Stress Levels from Naive BayesKernel-Leaf Pictures [5] GodliverTeacher, Ernest Mwebaze-In this article, a smartphone-based Cassaba recently mechanically studied to solve field diagnostic problems in leaves. [5] Godliver Teacher, Ernest Mwebaze-This article introduced a smartphone-based diagnostic program for cassava plants. It has recently been mechanically studied to solve the problem of field diagnostics in leaf-analysis. Predictive Predictive Models Based on Meteorological Observatory [6] William Dalmorra de Souza, Thainan Bystronski Remboski, Marilton Sanchotene de Aguiar, PauloRoberto Ferreira-Predict the use of pesticides to control price-increases.

3. METHODOLOGY

Data on disease resistance and insect infection are provided on site from wild trials without phytosanitary treatment in Table 1. These data were collected in high and low doses and monthly.Events are randomly measured and plants are randomly selected with a curved pattern. Incidence is determined by the third or fourth branch of the branch in the central third of the plant.

Table 1. Climate change using pest analysis and disease resistance standards

Weather acronyms	Definition
Tmin	Maximum-temperature average(℃)
Tmax	Maximum-temperature average(℃)
Rainfall	Total Raiinfall(mm)
NDR>=1mm	Number of days with rainfall>=1mm and <9 mm
NDR>=10mm	Number of days with rainfall>=10mm
RH	Average relative humidity(%)
NdRH90%	Number of days with relative humidity>=90%
NdRH90%	Number of days with relative humidity>=80%

Leaf rate (%) eq.1 used. It is not uncommon for each node formed in a plagiotropic branch to produce two leaves .

$$\textit{Leafiness(\%)} = [(\text{LEAF NUMBER}/2)/\textit{Number of nodes}]*100 \quad \text{--(1)}$$

Diseases and pests have been tested in both “high” and “low” production conditions, which are due to a two- year crop condition. The cut-off distance between thesestages is due to the difference observed in the field during the high and low production season.Pearson Correlations is performed with three different considerations: 1-10d (from 1 to 10 daysprior to incident testing); 11-20d (from 11 to 20 days prior to incident assessment); and 21-30d (from 21 to 30 days prior to incident assessment);

Machine Learning Methods

Various methods are used to predict diseases and pests incidence. The prevalence of diseases and pests varies and the climatic factors are independent models. In all cases 40% of the data is used for training and 60% is used for modeling.This type of division is done using the python library (sklearn.model_selection.train_test_split). The prediction methods are as follows: Multiple LinearRegression (RLM); Random-Forest Regressor (RFT), and Artificial Neural Networks- MultilayerPerceptron.

The ANN algorithm is simple, easy to implement, and very flexible. ANN identified the three nearest neighbors and the The metric used to calculate the space changed into the Euclidean distance. RFT randomly created a wooded area with a mixture of a hundred decision-making bushes to are expecting pests and sicknesses depending at the climate.

The artificial neural network used became a multi-layer perceptron (MLP) with three layers of neurons. Ten neurons were used in each of these layers (hidden_layer_sizes = 10, 10, 10). MLP educationwas performed using backpropagation.

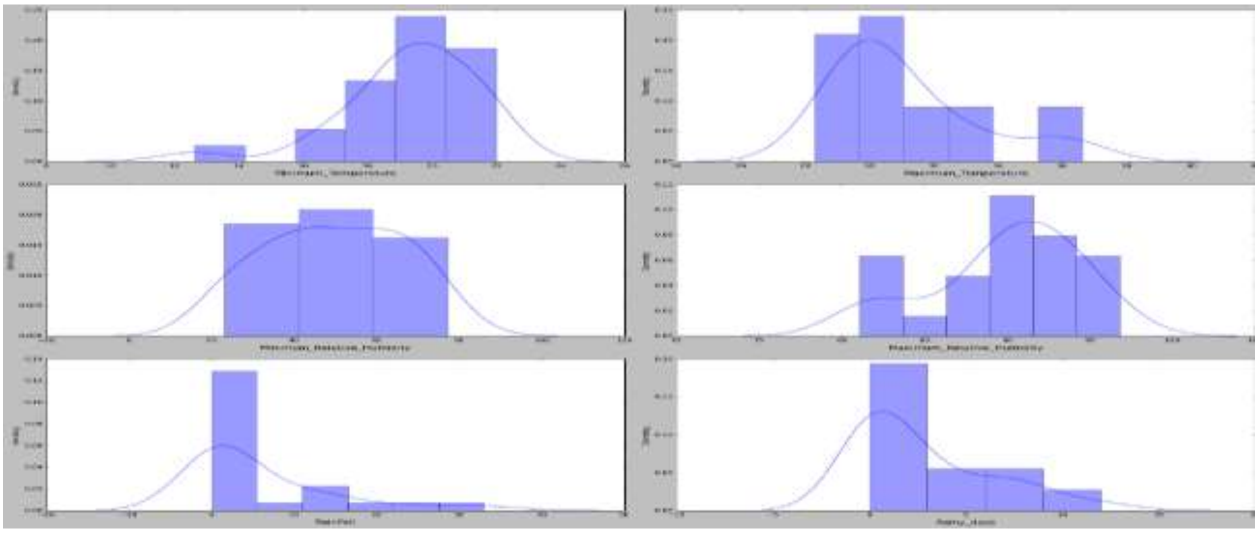
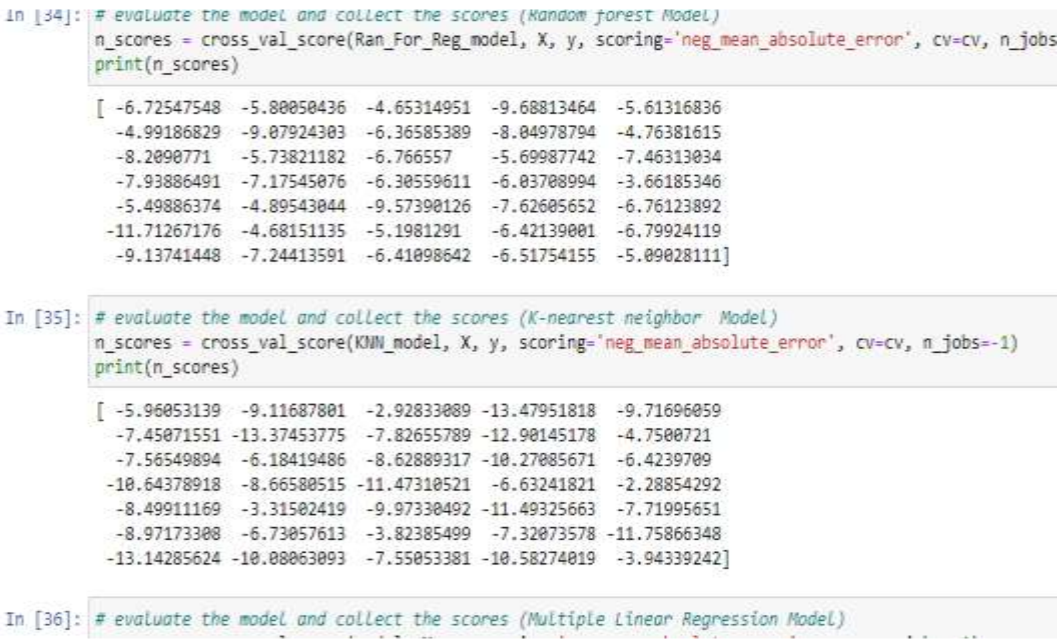


Fig1:Distribution of data based on weatherfactors(Temperature,Humidity,Rainfall)

4. MODEL EVALUATION

The actual target discipline facts and the effects of all models had been in comparison the use of several statistical signs together with accuracy, accuracy, and score stage. It way rootimply square mistakes (RMSE). Accuracy is the capacity of the model to multiply the measurements and was evaluated the use of a correction factor (R ^ 2 adj). All algorithms are limited to pest and disease predictions, RMSE indicators, actual yields, and the most accurate algorithm predictions. The Ridge Route (RLM) avoids the adverse conditions of a flexible matrix and controls the expansion and jitter normally associated with measuring small squares. Ridge avoids the problem of multicollinearity without leaving a variable regressor, so no information is lost.



```
In [34]: # evaluate the model and collect the scores (Random forest Model)
         n_scores = cross_val_score(Ran_For_Reg_model, X, y, scoring='neg_mean_absolute_error', cv=cv, n_jobs
         print(n_scores)

[ -6.72547548  -5.80050436  -4.65314951  -9.68813464  -5.61316836
  -4.99186829  -9.07924303  -6.36585389  -8.04978794  -4.76381615
  -8.2090771   -5.73821182  -6.766557    -5.69987742  -7.46313034
  -7.93886491  -7.17545076  -6.30559611  -6.03708994  -3.66185346
  -5.49886374  -4.89543044  -9.57390126  -7.62605652  -6.76123892
 -11.71267176  -4.68151135  -5.1981291   -6.42139001  -6.79924119
  -9.13741448  -7.24413591  -6.41098642  -6.51754155  -5.09028111]

In [35]: # evaluate the model and collect the scores (K-nearest neighbor  Model)
         n_scores = cross_val_score(KNN_model, X, y, scoring='neg_mean_absolute_error', cv=cv, n_jobs=-1)
         print(n_scores)

[ -5.96053139  -9.11687801  -2.92833089 -13.47951818  -9.71696059
  -7.45071551 -13.37453775  -7.82655789 -12.90145178  -4.7500721
  -7.56549894  -6.18419486  -8.62889317 -10.27085671  -6.4239709
 -10.64378918  -8.66580515 -11.47310521  -6.63241821  -2.28854292
  -8.49911169  -3.31502419  -9.97330492 -11.49325663  -7.71995651
  -8.97173308  -6.73057613  -3.82385499  -7.32073578 -11.75866348
 -13.14285624 -10.08063093  -7.55053381 -10.58274019  -3.94339242]

In [36]: # evaluate the model and collect the scores (Multiple Linear Regression Model)
```

Fig 2:Evaluation of models using K-fold cross validation

5. CONCLUSION

High temperatures, days with relative humidity above 80%, and moderate humidity showed a strong correlation with rust. In both cases, precipitation has a negative relationship with miners, and they lose weight-as-percipitation-changes.Machine gaining knowledge of algorithms can be used to make coffee and infer insects.The predictable sorts of pests advanced in this examine offer better pest control in high-yielding and low-yielding fruit years. The system gaining knowledge of algorithm makes no distinction between excessive yield and low yield guesses. The random forest version turned into very correct in predicting the outbreak of pests and illnesses based on climate situations.

6. REFERENCES

[1] X. J. Pei, ''Spread simulation of forest disease and pest based on multiagent system,'' M.S. thesis, Key Lab. Silviculture Conservation Ministry Educ., Beijing Forestry Univ., Bejing, China, 2010.

[2] Y. S. Song, H. J. Su, H. Y. Yu, T. Qu, G. B. Chang, and J. Zhao, ''Evaluation of economic losses caused by forest pest disasters between 2006 and 2010 in China,'' Forest Pest Disease, vol. 30, no. 6, pp. 1–4, 2011, doi: 10.3969/j.issn.1671-0886.2011.06.001.

[3] J. Patil and V. D. Mytri, ''A prediction model for population dynamics of cotton pest (Thrips tabaci Linde) using multilayer-perceptron neural network,'' Int. J. Comput. Appl., vol. 67, no. 4, pp. 19–26, Apr. 2013, doi: 10.5120/11384-6663.

[4] R. Jin and S. C. Li, ''Forecasting model for the occurrence degree of wheat aphids based on wavelet neural network,'' Acta Entomologica Sinica, vol. 58, no. 8, pp. 893–903, 2015, doi: 10.16380/j.kcxb.2015.08.010.
[5] Y. W. Yan, C. C. Feng, and P. H. Wan, ''Multiple regression and artificial neural network for the prediction of crop pest risks,'' in Proc. Int. Conf. Inf. Syst. Crisis Response Manage. Medit. Countries, 2015, pp. 73–84, doi: 10.1007/978-3-319-24399-3_7.
[6] M.-W. Shi, ''Based on time series and RBF network plant disease forecasting,'' Procedia Eng., vol. 15, pp. 2384–2387, Jan. 2011, doi: 10.1016/j. proeng.2011.08.447.
[7] C. S. Xiang and Z. Y. Zhou, ''Application of ARIMA and SVM hybrid model in pest forecast,'' Acta Entomologica Sinica, vol. 53, no. 9, pp. 1055–1060, 2010, doi: 10.16380/j.kcxb.2010.09.016.
[8] S. Sannakki, V. S. Rajpurohit, F. Sumira, and H. Venkatesh, ''A neural network approach for disease forecasting in grapes using weather parameters,'' in Proc. 4th Int. Conf. Comput., Jul. 2013, pp. 1–5, doi: 10.1109/ICCCNT.2013.6726613.
[9] X. F. Wang, C. L. Zhang, S. W. Zhang, and Y. H. Zhu, ''Forecasting of cotton diseases and pests based on adaptive discriminant deep belief network,'' Trans. Chin. Soc. Agricult. Eng., vol. 34, no. 14, pp. 157–164, 2018, doi: 10.11975/j.issn.1002-6819.2018.14.020.
[10] Z. K. Zhang, H. Qi, S. Wang, S. W. Cai, Y. Yao, and D. Li, ''Forest pest prediction method with apriori algorithm,'' J. Northeast Forestry Univ., vol. 45, no. 8, pp. 93–96, 2017.
[11] A. Anggraeni and C. Lin, ''Application of SAM and SVM techniques to burned area detection for landsat TM images in forests of South sumatra,'' in Proc. Int. Conf. Environ. Sci. Technol., 2011, pp. 160–164.
[12] M. C. Hansen, P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend, ''High-resolution global maps of 21st-century forest cover change,'' Science, vol. 342, no. 6160, pp. 850–853, Nov. 2013, doi: 10.1126/science.1244693.
[13] Y. N. Song, W. P. Liu, S. X. Zong, and Y. Q. Luo, ''Segmentation algorithm for unmanned aerial vehicle imagery based on superpixel and ultrametric contour Map,'' J. Comput.-Aided Design Comput. Graph., vol. 31, no. 8, pp. 1294–1300, 2019, doi: 10.3724/SP.J.1089.2019.17563.
[14] P. Rajan, B. Radhakrishnan, and L. P. Suresh, ''Detection and classification of pests from crop images using support vector machine,'' in Proc. Int. Conf. Emerg. Technol. Trends (ICETT), Oct. 2016, pp. 1–6, doi: 10.1109/ICETT.2016.7873750.
[15] E. Mwebaze and G. Owomugisha, ''Machine learning for plant disease incidence and severity measurements from leaf images,'' in Proc. 15th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA), Dec. 2016, pp. 158–163, doi: 10.1109/ICMLA.2016.0034