INTERNATIONAL JOURNAL OF PROFESSIONAL BUSINESS REVIEW
ISSN: 2525-3654
OPEN ACCESS

CROP INSURANCE PREMIUM RECOMMENDATION SYSTEM USING ARTIFICIAL INTELLIGENCE TECHNIQUES

Prakash M C[A], P. Saravanan[B]

ARTICLE INFO

Article history:

Received 31 January 2023

Accepted 06 April 2023

Keywords:

Crop Insurance Premium;
Right Farming Practices;
Agriculture;
Ada Boost Regressor;
Gradient Boosting Regressor;
Extra Trees Regressor;
Machine Learning (ML);
Artificial Intelligence (AI).

PREREGISTERED
OPEN DATA
OPEN MATERIALS

ABSTRACT

Purpose: The objective of this study is to build a crop insurance premium recommender model which will be fair to both crop insurance policy holders and crop insurance service providers.

Theoretical Framework: The Nonparametric Bayesian Model (modified) is the name of the proposed model suggested by Maulidi et al. (2021) and it consists of six variables which are regional risk, cultivation time period, land area, claim frequency, discount eligibility (local variable) and premium. Discount eligibility variable is introduced to encourage right farming practices among farmers.

Design/methodology/approach: Descriptive research method is used in this study as it is used to accurately represent the characteristics of a group of items. The population for this study is 943 respondents. The entire dataset is used for in-depth and accurate analysis. Five Artificial Intelligence models (Machine Learning models) are proposed for crop insurance premium prediction and they are Ada Boost Regressor, Gradient Boosting Regressor, Extra Trees Regressor, Support Vector Regressor and K-Neighbors Regressor. Among them Gradient Boosting Regression model has given the highest accuracy. Thus, Gradient Boosting Regression model is the most suitable model to be recommended for crop insurance premium prediction.

Findings and Suggestions: Regional risk, land area, claim frequency and cultivation time period is the order of independent variables from highest to least in terms of regression coefficient. This relative importance helps Non-Banking Financial Companies (NBFCs) to suggest farmers that they should concentrate most on the regional risk or chances of crop failure in a particular region in which they are doing agriculture and least on the cultivation time period of a crop or the season in which a crop is cultivated. Two suggestions for future researchers are to extend this research work to other parts of Tamil Nadu and to apply hybrid machine learning techniques to the proposed model.

Practical Implication: Unlike the existing formula-based traditional method used for calculating crop insurance premium, artificial intelligence models (machine learning models) can automatically learn the changes that take place with respect to the nature of variables in the proposed model and improve its accuracy based on new data. Hence, the crop insurance premium suggested by the most accurate model among the artificial intelligence models used in this study will be fair to both NBFCs and farmers. Here, fair means moderate. On the other hand, the crop insurance premium suggested by the existing formula-based method may not be fair in the long term as they cannot automatically learn the changes that take place with respect to the nature of variables in the proposed model and improve.

[A] *Research Scholar, SRM College of Management, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu-603203, Tamil Nadu, India. E-mail: pm8443@srmist.edu.in*
Orcid: https://orcid.org/0000-0003-1534-9216
[B] *Associate Professor, SRM College of Management, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu-603203, Tamil Nadu, India. E-mail: saravanp2@srmist.edu.in*
Orcid: https://orcid.org/0000-0003-1400-3571

Originality/value: In this research article, the relative importance of independent variables in the proposed model is determined and it helps NBFCs to suggest farmers that they should concentrate most on the region they are doing agriculture and least on the cultivation time period of a crop. Additionally, a machine learning model which can automatically learn and improve itself is used and hence the crop insurance premium predicted by it will be fair. Finally, the entire population containing 943 respondents details is analysed.

Doi: https://doi.org/10.26668/businessreview/2023.v8i4.1270

SISTEMA DE RECOMENDAÇÃO DE PRÊMIO DE SEGURO DE SAFRA USANDO TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL

RESUMO

Objetivo: O objetivo deste estudo é construir um modelo de recomendação de prêmio de seguro agrícola que seja justo tanto para os titulares de apólices de seguro agrícola quanto para os prestadores de serviços de seguro agrícola.

Referencial Teórico: O Modelo Bayesiano Não Paramétrico (modificado) é o nome do modelo proposto sugerido por Maulidi et al. (2021) e consiste em seis variáveis que são risco regional, período de cultivo, área de terra, frequência de reivindicação, elegibilidade de desconto (variável local) e prêmio. A variável de elegibilidade de desconto é introduzida para incentivar práticas agrícolas corretas entre os agricultores.

Desenho/metodologia/abordagem: O método de pesquisa descritivo é usado neste estudo, pois é usado para representar com precisão as características de um grupo de itens. A população para este estudo é de 943 respondentes. Todo o conjunto de dados é usado para uma análise aprofundada e precisa. Cinco modelos de Inteligência Artificial (modelos de Aprendizado de Máquina) são propostos para predição de prêmio de seguro agrícola e são eles Ada Boost Regressor, Gradient Boosting Regressor, Extra Trees Regressor, Support Vector Regressor e K-Neighbors Regressor. Entre eles, o modelo Gradient Boosting Regression forneceu a maior precisão. Assim, o modelo Gradient Boosting Regression é o modelo mais adequado para ser recomendado para a previsão de prêmio de seguro agrícola.

Resultados e Sugestões: Risco regional, área de terra, frequência de reivindicação e período de tempo de cultivo é a ordem das variáveis independentes do maior para o menor em termos de coeficiente de regressão. Essa importância relativa ajuda as Empresas Financeiras Não Bancárias (NBFCs) a sugerir aos agricultores que eles devem se concentrar mais no risco regional ou nas chances de quebra de safra em uma determinada região em que estão praticando agricultura e menos no período de cultivo de uma safra ou a estação em que uma cultura é cultivada. Duas sugestões para futuros pesquisadores são estender este trabalho de pesquisa para outras partes de Tamil Nadu e aplicar técnicas híbridas de aprendizado de máquina ao modelo proposto.

Implicação prática: Ao contrário do método tradicional baseado em fórmula existente usado para calcular o prêmio do seguro agrícola, os modelos de inteligência artificial (modelos de aprendizado de máquina) podem aprender automaticamente as mudanças que ocorrem com relação à natureza das variáveis no modelo proposto e melhorar sua precisão com base em novos dados. Portanto, o prêmio de seguro agrícola sugerido pelo modelo mais preciso entre os modelos de inteligência artificial usados neste estudo será justo tanto para os NBFCs quanto para os agricultores. Aqui, justo significa moderado. Por outro lado, o prêmio de seguro agrícola sugerido pelo método baseado em fórmula existente pode não ser justo a longo prazo, pois eles não podem aprender automaticamente as mudanças que ocorrem com relação à natureza das variáveis no modelo proposto e melhorar.

Originalidade/valor: Neste artigo de pesquisa, a importância relativa das variáveis independentes no modelo proposto é determinada e ajuda os NBFCs a sugerir aos agricultores que eles devem se concentrar mais na região em que estão fazendo agricultura e menos no período de cultivo de uma cultura . Além disso, um modelo de aprendizado de máquina que pode aprender e melhorar automaticamente é usado e, portanto, o prêmio de seguro agrícola previsto por ele será justo. Finalmente, toda a população contendo 943 detalhes dos respondentes é analisada.

Palavras-chave: Prêmio de Seguro de Cultura, Práticas Agrícolas Corretas, Agricultura, Ada Boost Regressor, Gradient Boosting Regressor, Extra Trees Regressor, Machine Learning (ML), Inteligência Artificial (IA).

SISTEMA DE RECOMENDACIÓN DE PRIMAS DE SEGUROS DE CULTIVOS UTILIZANDO TÉCNICAS DE INTELIGENCIA ARTIFICIAL

RESUMEN
Propósito: El objetivo de este estudio es construir un modelo de recomendación de primas de seguros de cosechas que sea justo tanto para los titulares de pólizas de seguros de cosechas como para los proveedores de servicios de seguros de cosechas.
Marco Teórico: El Modelo Bayesiano No Paramétrico (modificado) es el nombre del modelo propuesto sugerido por Maulidi et al. (2021) y consta de seis variables que son riesgo regional, período de tiempo de cultivo, área de tierra, frecuencia de reclamos, elegibilidad de descuento (variable local) y prima. Se introduce la variable de elegibilidad de descuento para fomentar prácticas agrícolas correctas entre los agricultores.
Diseño/metodología/enfoque: En este estudio se utiliza un método de investigación descriptivo, ya que se utiliza para representar con precisión las características de un grupo de elementos. La población para este estudio es de 943 encuestados. Todo el conjunto de datos se utiliza para un análisis profundo y preciso. Se proponen cinco modelos de inteligencia artificial (modelos de aprendizaje automático) para la predicción de primas de seguros de cultivos y son Ada Boost Regressor, Gradient Boosting Regressor, Extra Trees Regressor, Support Vector Regressor y K-Neighbors Regressor. Entre ellos, el modelo de regresión de aumento de gradiente ha dado la mayor precisión. Por lo tanto, el modelo de regresión de aumento de gradiente es el modelo más adecuado para recomendar para la predicción de primas de seguros de cosechas.
Hallazgos y Sugerencias: El riesgo regional, el área de la tierra, la frecuencia de reclamos y el período de tiempo de cultivo es el orden de las variables independientes de mayor a menor en términos de coeficiente de regresión. Esta importancia relativa ayuda a las empresas financieras no bancarias (NBFC, por sus siglas en inglés) a sugerir a los agricultores que deberían concentrarse más en el riesgo regional o las posibilidades de mala cosecha en una región particular en la que están haciendo agricultura y menos en el período de tiempo de cultivo de un cultivo o la temporada en que se cultiva un cultivo. Dos sugerencias para futuros investigadores son extender este trabajo de investigación a otras partes de Tamil Nadu y aplicar técnicas híbridas de aprendizaje automático al modelo propuesto.
Implicación práctica: a diferencia del método tradicional basado en fórmulas existente que se utiliza para calcular la prima del seguro de cosechas, los modelos de inteligencia artificial (modelos de aprendizaje automático) pueden aprender automáticamente los cambios que tienen lugar con respecto a la naturaleza de las variables en el modelo propuesto y mejorar su precisión en función de sobre nuevos datos. Por lo tanto, la prima del seguro de cosecha sugerida por el modelo más preciso entre los modelos de inteligencia artificial utilizados en este estudio será justa tanto para las NBFC como para los agricultores. Aquí, justo significa moderado. Por otro lado, la prima del seguro de cosecha sugerida por el método basado en fórmulas existente puede no ser justa a largo plazo ya que no pueden aprender automáticamente los cambios que ocurren con respecto a la naturaleza de las variables en el modelo propuesto y mejorar.
Originalidad/valor: en este artículo de investigación, se determina la importancia relativa de las variables independientes en el modelo propuesto y ayuda a los NBFC a sugerir a los agricultores que deben concentrarse más en la región en la que están haciendo agricultura y menos en el período de cultivo de un cultivo. . Además, se utiliza un modelo de aprendizaje automático que puede aprender y mejorar automáticamente y, por lo tanto, la prima del seguro de cosecha prevista por él será justa. Finalmente, se analiza toda la población que contiene 943 datos de encuestados.

Palabras clave: Prima de Seguro de Cultivos, Prácticas Agrícolas Correctas, Agricultura, Regresor de Refuerzo de Ada, Regresor de Aumento de Gradiente, Regresor de Árboles Adicionales, Aprendizaje Automático (ML), Inteligencia Artificial (AI).

INTRODUCTION

India is a country whose economy is significantly influenced by agriculture. Agriculture provides direct and indirect job opportunities to millions of people in India. Though it is the biggest employment opportunity provider, there are many challenges faced by farmers that need to be addressed as quickly as possible.

Crop insurance is a type of insurance policy that provides coverage to farmers for losses in crop yields or revenue due to natural disasters, such as droughts, floods, pest attacks or diseases. These losses can have a significant financial impact on farmers, making crop insurance an important risk management tool in the agricultural industry.

However, the cost of crop insurance can be a significant burden for farmers, especially for small and medium-sized farmers who may not have the resources to absorb large losses. Therefore, there is a need for accurate and reliable methods to determine the appropriate level of insurance coverage and premium for each farmer.

The development of a Crop Insurance Premium Recommender using machine learning techniques could help farmers make informed decisions about the level of insurance coverage they need and the premium they should pay. This could help to improve the accessibility and affordability of crop insurance, enabling more farmers to protect their livelihoods against natural disasters.

Manual determination of crop insurance premiums can be time-consuming and resource-intensive. Machine learning can help automate the process, making it more efficient and cost-effective. This can also help insurance companies scale their operations and provide coverage to a larger number of farmers.

The objective of this research is to develop a crop insurance premium recommendation model that ensures fairness for both crop insurance policyholders and providers of crop insurance services.

Artificial Intelligence techniques are used in agriculture for various applications such as crop yield prediction, intelligent spraying, predictive insights, agriculture robots, crop and soil monitoring, disease diagnosis etc., but this research article is focused on recommending fair crop insurance premium. Here, fair crop insurance premium means premium amount which is moderate. Mutaqin et al. (2015) proposed a method for crop insurance premium count estimation. Providing insurance for protection against losses due to crop failure was discussed by Maulidi et al. (2021). Significance of charging fair premium to crop insurance policy holders and ensuring crop insurance service providers achieve financial balance was pointed out by Maulidi et al. (2021).

LITERATURE REVIEW

Cotton yield forecasting was done using Artificial Intelligence techniques and a novel idea for measuring cotton yield was given by Xu et al. (2021). Crop loss detection at field parcel scale was done using Random Forest model and it was used to confirm whether crop loss claims made by farmers were genuine (Hiremath et al. (2021)). According to Schmidt et al. (2022), Machine Learning techniques were successful in predicting the crop yield-weather conditions relationship with better accuracy compared to traditional approaches. Kittichotsatsawat et al. (2022) found out that the forecasted yield of cherry coffee crop increased each year and identified that Multiple Linear Regression and Artificial Neural Networks can be used for forecasting arabica coffee yield to exactly meet the market demand. Islam et al. (2021) proposed yield forecasting models to forecast rice yield and proved that yield in a major part of the study area (more than 70%) can be precisely forecasted using such yield forecasting models. Arumugam et al. (2021) proposed using Gradient Boosted Regression for rice yield prediction and it perhaps met the needs of insurance companies and government agencies with respect to rice yield prediction. Crop insurance model for covering crop failure losses was proposed by Maulidi et al. (2021) and a method based on rounding off for data discretization was also provided by them. Agarwal et al. (2020) highlight the benefits of using machine learning techniques to predict crop insurance premium and the benefits mentioned were better accessibility, better affordability, and better sustainability of the crop insurance program. Gupta et al. (2019) tell that their research work will be useful for policymakers and insurance companies in designing better insurance schemes for farmers. Singh and Agarwal (2018) suggest the use of satellite images too as a data source for better prediction of crop insurance premium.

MATERIAL AND METHODOLOGY

There are many research articles which have discussed about various aspects such as crop yield forecasting, crop loss detection etc. (Islam et al. (2021), Kittichotsatsawat et al. (2022), Xu et al. (2021)) etc., but there are only few research articles which have discussed about crop insurance premium prediction that too with the limitation of giving less accurate results (Maulidi et al. (2021)). The Nonparametric Bayesian Model (existing model) used in (Maulidi et al. (2021)) does not have the variable “claim frequency” and predictions were made using traditional formula-based algorithm.

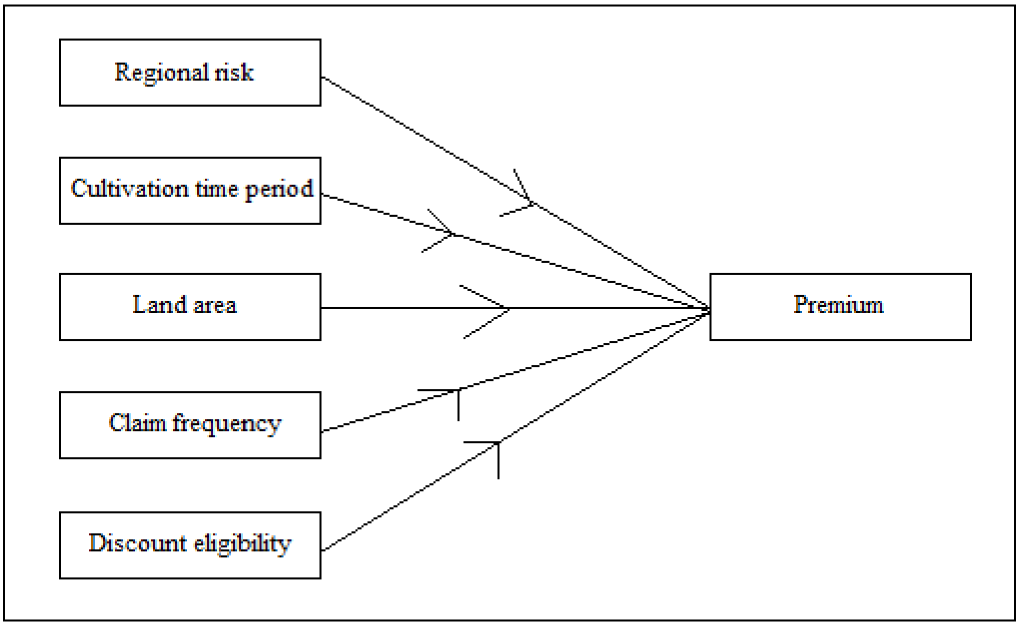
To build a crop insurance premium recommender model which will be fair to both crop insurance policy holders and crop insurance service providers.

Crop insurance policy holders will benefit from fair premium that is charged to them and crop insurance service providers will benefit from achievement of financial balance or fair profit (Maulidi et al. (2021)).

Descriptive research method is used in this study as it is used to accurately represent the characteristics of a group of items (Kothari (2004)).

The material used in this study is a dataset consisting of secondary data (nature of data) that was given by Non-Banking Financial Companies (NBFCs) in Coimbatore. The entire dataset (population) is chosen for this study. It contains 943 respondents details and six variables.

Figure 1. Proposed model – Nonparametric Bayesian Model (modified)



(Source: Maulidi et al. (2021))

The characteristics of the crops are explained by the below variables:

Regional risk: The probability of crop failure in each region.

Cultivation time period: The time period during which a given crop is cultivated.

Land area: Area of land under cultivation in hectares.

Claim frequency: The number of times a crop insurance policy holder has claimed crop insurance.

Discount eligibility: This variable denotes whether a crop insurance policy holder is eligible for 5% discount or not based on the criterion that those who have not availed crop insurance for 3 years continuously are eligible for the above-mentioned discount. This is an initiative to encourage right farming practices.

Premium: The crop insurance premium that is charged to crop insurance policy holders.

Among the variables mentioned above, premium is the dependent variable and all the other variables – regional risk, cultivation time period, land area, claim frequency and discount eligibility are independent variables. The existing model was proposed by Maulidi et al. (2021) and it consists of one dependent variable "premium" and three independent variables – regional risk, cultivation time period and land area claim. The Nonparametric Bayesian Model (modified) is the name of the proposed model suggested by Maulidi et al. (2021) and it consists of six variables which are mentioned and explained above. It differs from the existing model by excluding the variable named "land area claim" and by including three variables – land area, claim frequency and discount eligibility.

Tatsat et al. (2021) describe Support Vector Machine technique as follows – It is a supervised machine learning technique that can be used for regression as well as classification. It uses a hyperplane and two support vectors to meet its objective. Its objective is to maximise the distance between the hyperplane and each support vector when used for classification. In case of regression, its objective is to fit as many datapoints as possible within the area between the two support vectors.

K-Nearest Neighbors technique works by determining the K most similar instances using distance measures like Euclidean distance and assigns datapoints to class labels based on majority vote (Tatsat et al. (2021)). This technique too can be used for regression and classification.

Extremely randomized trees (Extra trees) technique is a slightly different form of random forest technique. It works like random forest technique in terms of building trees and splitting nodes, but differs by picking observations without replacement and choosing a random split to separate the parent node into two random child nodes (Tatsat et al. (2021)).

Adaptive Boosting (AdaBoost) as the name suggests is a boosting technique which works by trying predictors in a sequential manner and each subsequent model tries to fix the errors of its predecessor. Each iteration involves altering the sample distribution by changing the weights attached to each instance (Tatsat et al. (2021)).

Gradient boosting technique too is a boosting technique which is like AdaBoost in terms of trying predictors sequentially. It adds the previous underfitted predictions to the ensemble in a sequential manner and ensures that previous errors are corrected. It differs from AdaBoost by attempting to fit the new predictor to the outstanding errors made by its predecessor (Tatsat et al. (2021)).

RESULTS AND DISCUSSION

The Multiple Linear Regression equation for the proposed model is given below.

$$y = m_1x_1 + m_2x_2 + m_3x_3 + m_4x_4 + m_5x_5 + c$$

y - dependent variable

c - constant

x_1, x_2, x_3, x_4 and x_5 – independent variables

m_1, m_2, m_3, m_4 and m_5 - regression coefficients of x_1, x_2, x_3, x_4 and x_5

H_{01}: There is no significant relationship between Regional risk and Premium.

H_1: There is a significant relationship between Regional risk and Premium.

H_{02}: There is no significant relationship between Cultivation time period and Premium.

H_2: There is a significant relationship between Cultivation time period and Premium.

H_{03}: There is no significant relationship between Land area and Premium.

H_3: There is a significant relationship between Land area and Premium.

H_{04}: There is no significant relationship between Claim frequency and Premium.

H_4: There is a significant relationship between Claim frequency and Premium.

H_{05}: There is no significant relationship between Discount eligibility and Premium.

H_5: There is a significant relationship between Discount eligibility and Premium.

The research objective is to build a crop insurance premium recommender model which will be fair to both crop insurance policy holders and crop insurance service providers, and Ordinary Least Squares (OLS) method in Multiple Linear Regression was used to meet this objective by testing the hypotheses. The above stated hypotheses are accepted or rejected based on the statistical significance of independent variables used in the proposed model.

Figure 2. p-values of independent variables in the proposed model

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-78.3571	222.717	-0.352	0.725	-515.439	358.725
Regional_risk	4446.8065	1202.011	3.699	0.000	2087.862	6805.751
Cultivation_time_period	629.0301	48.530	12.962	0.000	533.790	724.270
Land_area	1627.6058	49.144	33.119	0.000	1531.160	1724.052
Claim_frequency	1014.6344	44.726	22.685	0.000	926.859	1102.410
Discount_eligibility	-123.4410	123.023	-1.003	0.316	-364.873	117.991

(Source: Prepared by the authors (2023))

We can infer from Figure 2 that all independent variables except Discount eligibility are statistically significant. Hence, H_{01}, H_{02}, H_{03}, H_{04} and H_5 are rejected, and H_1, H_2, H_3 H_4 and H_{05} are accepted. In other words, there is a significant relationship between Regional risk and Premium, Cultivation time period and Premium, Land area and Premium, and Claim frequency and Premium. Since Discount eligibility does not have a significant relationship with Premium it can be eliminated from the proposed model. Hence, the proposed model consists of only five variables and the regression coefficient values of the four independent variables in it are 4446.8065 for regional risk, 629.0301 for cultivation time period, 1627.6058 for land area and 1014.6344 for claim frequency.

The multiple linear regression equation (formula) for the updated proposed model is

$$y = m_1x_1 + m_2x_2 + m_3x_3 + m_4x_4 + c$$

OR

premium = 4446.8065 (regional risk) + 629.0301 (cultivation time period) + 1627.6058 (land area) + 1014.6344 (claim frequency) - 78.3571

Note: The formula used by the Nonparametric Bayesian Model proposed by (Maulidi et al. (2021)) for premium estimation is, $E[X|(\bar{x}_1 = 1, \bar{x}_2 = 0)] = \sum_{k=1}^{n} P[\theta_k|(\bar{x}_1, \bar{x}_2)]\, E(X|\theta_k)$.

The accuracy scores of Ada Boost Regression model, Gradient Boosting Regression model, Extra Trees Regression model, Support Vector Regression model and K-Neighbors Regression model are mentioned below in Table 1.

Table 1. Accuracy scores of machine learning models

Model Name	R^2 value to measure model accuracy
Ada Boost Regression	67.60%
Gradient Boosting Regression	93.58%
Extra Trees Regression	77.94%
Support Vector Regression	61.51%
K-Neighbors Regression	68.71%

(Source: Prepared by the authors (2023))

The accuracy of Ada Boost Regression model with respect to the given dataset is 67.60%. When Gradient Boosting Regression model was applied to the given dataset it gave an accuracy of 93.58%. Extra Trees Regression model resulted in 77.94% accuracy followed by Support Vector Regression model and K-Neighbors Regression model whose accuracy scores are 61.51% and 68.71% respectively. From this we observe that for the given dataset Support Vector Regression model has given the lowest accuracy and Gradient Boosting Regression model has given the highest accuracy. Thus, Gradient Boosting Regression model is the most suitable model to be recommended for crop insurance premium prediction as it has given the highest accuracy among the proposed Machine Learning models.

CONCLUSION

The relative importance of independent variables in the proposed model is determined using regression coefficients. Regional risk, land area, claim frequency and cultivation time period is the order of variables from highest to least. This relative importance helps NBFCs to suggest farmers that they should concentrate most on the region they are doing agriculture and least on the cultivation time period of a crop. The model may not perform at its best when applied to different regions than the ones used for training. In order to overcome this limitation, addition or removal of independent variables from the proposed model may be necessary. Two suggestions for future researchers are to extend this research work to other parts of Tamil Nadu and to apply hybrid machine learning techniques to the proposed model.

This study is focused on building a crop insurance premium recommender model and that model is Nonparametric Bayesian Model (modified). The relative importance of independent variables in the proposed model is determined using OLS method and it helps NBFCs to give valuable suggestions to farmers which they may be unaware of. The Gradient

Boosting Regression model is the most accurate among the machine learning models used in this study and it can automatically learn and improve itself. This capability of it will enable prediction of fair crop insurance premium which the existing formula-based method may not be able to do in the long run due to its inability to automatically learn and improve itself.

REFERENCES

Agarwal, R., Singh, A., & Singh, M. (2020). Crop Insurance Premium Prediction Using Machine Learning Techniques. *International Journal of Professional Business Review*, *4*(1), 12-20.

Arumugam, P., Chemura, A., Schauberger, B., & Gornott, C. (2021). Remote sensing based yield estimation of rice (Oryza sativa L.) using gradient boosted regression in India. *Remote Sensing*, *13*(12), 2379.

Gupta, S., Kumar, S., & Kumar, A. (2019). A Comparative Study of Machine Learning Techniques for Crop Insurance Premium Prediction. *International Journal of Professional Business Review*, *3*(1), 21-28.

Hiremath, S., Wittke, S., Palosuo, T., Kaivosoja, J., Tao, F., Proll, M., ... & Mamitsuka, H. (2021). Crop loss identification at field parcel scale using satellite remote sensing and machine learning. *PloS one*, *16*(12), e0251952.

Islam, M. M., Matsushita, S., Noguchi, R., & Ahamed, T. (2021). Development of remote sensing-based yield prediction models at the maturity stage of boro rice using parametric and nonparametric approaches. *Remote Sensing Applications: Society and Environment*, *22*, 100494.

Kittichotsatsawat, Y., Tippayawong, N., & Tippayawong, K. Y. (2022). Prediction of arabica coffee production using artificial neural network and multiple linear regression techniques. *Scientific Reports*, *12*(1), 1-14.

Kothari, C.R. (2004) *Research methodology*: *Methods and techniques*. New Age International.

Maulidi, I., Syahrini, I., Oktavia, R., Ihsan, M., & Emha, R. (2021, February). An analysis on determination of land area claim prediction for rice farmers insurance business in Indonesia using nonparametric bayesian method. In *Journal of Physics: Conference Series* (Vol. 1796, No. 1, p. 012006). IOP Publishing.

Mutaqin, A. K., Kudus, A., & Karyana, Y. (2015). Metode Parametrik Untuk Menghitung Premi Asuransi Usahatani Padi (AUTP) di Indonesia. *Prosiding SNaPP*, 15-23.

Schmidt, L., Odening, M., Schlanstein, J., & Ritter, M. (2022). Exploring the weather-yield nexus with artificial neural networks. *Agricultural Systems*, *196*, 103345.

Singh, N., & Agarwal, A. (2018). Machine Learning Techniques for Crop Insurance Premium Prediction: A Review. *International Journal of Professional Business Review*, *2*(1), 30-38.

Tatsat, H., Puri, S., & Lookabaugh, B. (2021). *Machine Learning and Data Science Blueprints for Finance*. O'Reilly Media.

Xu, W., Chen, P., Zhan, Y., Chen, S., Zhang, L., & Lan, Y. (2021). Cotton yield estimation model based on machine learning using time series UAV remote sensing data. *International Journal of Applied Earth Observation and Geoinformation*, *104*, 102511.