ESTIMATING THE EFFICIENCY OF MACHINE LEARNING IN FORECASTING HARVESTING TIME OF RICE

[1]Dr.R.Saravanakumar, [2]SrilathaToomula, [3]N.Zareena, [4]Abhay Chaturvedi, [5]Dr. S. Thowseaf, [6]Dr. R. Manikandan

[1]Associate Professor ,Institute of ECE,Saveetha School of Engineering , Saveetha Institute of Medical and Technical Sciences
Email ID :saravanakumarr.sse@saveetha.com
[2]Assistant Professor,Department of Computer Science,RBVRR Women's College, (Autonomous), Hyderabad,Telangana State
Email ID:toomula.srilatha@gmail.com
[3]Assistant Professor, CSE Department,RVR & JC College Of Engineering,Guntur-522019
Email ID:zareenacse@gmail.com
[4]Associate Professor, Department of Electronics & Communication Engineering, GLA University, Mathura-281406 (U.P.) INDIA
Email ID: abhaychat@gmail.com
[5]Assistant Professor, The QuaideMilleth College For Men, Medavakkam, Chennai, India
thowseaf786@gmail.com
[6]Head, Department of CS, The Quaide Milleth College for Men, Medavakkam, Chennai, India
manisankar27@gmail.com

Abstract: The study aims to measure the effectiveness of a system that used a machine-learning algorithm to predict the harvesting times of the rice crop. The study estimates the forecasting accuracy of the system using the opinions of experts and agriculture farmers. The system was designed to analyse the Moisture Content of the Grain, Ripe Grains Per Panicle, Number of Days After Sowing, Number of Days After Heading, Temperature at the time Heading and Surface Moisture of the soil. When the system is feed with before mentioned input, it generates output in the Likert scale of (Highly Not Recommended – Highly Recommended) using the mean score derived. For the study, 150 agriculture experts and 150 farmers practising rice cultivation for more than 5 years were considered as the samples. Further 24 fields not less than a hectare were considered for the study purpose. The opinion of agricultural experts and farmers were compared against the opinion generated by the system on the Likert scale. From the result obtained through analysis, it can be well perceived that there is no significant difference in an opinion posted by an expert, farmer with system generated output. Also, the system is generating output very lose to experienced farmed. Further, the standard deviation estimated is very least, which indicates the system is efficient enough in producing accurate opinion close enough with experienced farmer and experts. From the correlation value, it can be interpreted that there is a 79.1% relationship between opinions posted by an expert with system generated output. Further, there is a 72.8% relationship between opinion posted by experienced farmers with system generated output. Thereby, it is understood from the result that considering individual responses without considering the mean score, the system generated output is close to expert opinion. Furthermore, the estimated coefficient value indicates that the system generated output can be predicted using the opinion posted by Experts and Experienced Farmers. The regression equation is given; System Generated Opinion = 0.463 + (0.573×Expert) + (0.307×Farmers). From the regression plot, it can be additionally interpreted that most of the opinions posted by the Experts and Experienced farmers coincided with the opinion generated by the System using a machine learning algorithm.

Keywords: Agriculture, Information Communication Technology, Machine Learning

Introduction

Rice, like wheat and oats, is a grain or cereal. A grain is a plant's entire seed, which is cultivated, harvested, and processed for human consumption. Rice is the seed obtained from the Oryza sativa (Asian rice) or Oryza glaberrima plant's long, grass-like stem (African rice). It is the most widely consumed cereal grain by a large portion of the world's human population, particularly in Asia. Except for Antarctica, rice is grown on every continent on the planet. After sugarcane and maize, it is the world's third-largest producer. All types of Asian rice, according to genetic evidence, arose from a single domestication that occurred 8,200–13,500 years ago in the Pearl River delta.In 2011, genetic evidence revealed that all types of Asian rice descended from a single domestication that took place 8,200–13,500 years ago in Ancient China's Pearl River valley. Rice was introduced to South and Southeast Asia from East Asia. Rice was brought to Europe via Western Asia and the Americas via European colonisation. Cultivated rice (the grass species Oryza sativa) is known to come in over 40,000 different varieties. However, the exact number is unknown (irri.org, 2021).

The method of extracting the mature rice crop from the field is known as harvesting. Cutting, stacking, handling, threshing, washing, and hauling are all part of paddy harvesting. Grain yield is maximised while grain loss and quality degradation are minimised with proper harvesting techniques. Harvesting may be achieved by hand with sickles and knives, or by machine with threshers or combine harvesters.Crop failure can be avoided by harvesting at the right time. Rats, birds, lodging, insects, and shattering can all cause grain losses. Harvesting promptly guarantees good grain quality and a high market value. Too early harvesting will result in a higher percentage of unfilled or immature crops, lowering yield and increasing grain breakage during milling (International Rice Research Institute, 2021). Harvesting rice too late can result in unnecessary losses and breakage. Rice seed germination ability is also affected by harvest time. Being such an important staple crop throughout the world, this research was conducted to create an Information Communication Technology-based (ICT) based device that would accurately forecast the timing ofthe harvest of the rice using a machine learning algorithm.

Review of Literature

Agriculture is extremely important to the global economy. With the continued growth of the human population, pressure on the agricultural system will increase(Liakos, K. G., Busato, P., et. al, 2018). Agri-technology and precision farming, also known as digital agriculture, are emerging as new scientific fields that employ data-intensive methods to improve agricultural productivity while reducing environmental effects(Fashoto, S., Mbunge, E., et. al, 2021). A variety of sensors provide data produced in modern agricultural operations, allowing for a better understanding of the operating environment (an interaction of complex crop, soil, and weather conditions) as well as the process itself (machinery data), resulting in more precise and faster decisions(Sharma, R., Kamble, S. S., et. al, 2020).Machine learning (ML) has risen to prominence alongside big data technologies and high-performance computing to open up new avenues for unravelling, quantifying, and understanding data-intensive processes in agricultural operations. ML is characterised as a scientific field that allows machines to learn without being strictly programmed, among other things(Wei, M. C. F., Maldaner, L. F., et. al, 2020).

The yield prediction, disease detection, weed detection, crop quality, and species recognition are the important applications of machine learning in the crop section(Morota, G., Ventura, R. V., et. al, 2018)(Banerjee, A., Mitra, A., & Biswas, A. 2021). Animal protection and livestock production were the two sub-categories of ML applications in the livestock section(Yahata, S., Onishi, T., Yamaguchi, et. al, 2017).The system segments and detects occluded crops with full ripped even when these are inconspicuous(Mekonnen, Y., Namuduri, S., et. al, 2019). The main aim of the system was to reduce labour requirements in manual harvesting and handling operations.

The objective of the Study

The research looked into the effectiveness of a system that used a machine-learning algorithm to predict the harvesting times of the rice crop. The studyestimates the forecasting accuracy of the system using the opinions of experts and agriculture farmers.

Methodology

The system was designed to analyse the Moisture Content of the Grain, Ripe Grains Per Panicle, Number of Days After Sowing, Number of Days After Heading,Temperature at the time Heading and Surface Moisture of the soil. These are considered as the variables to forecast the harvesting time because; the ideal moisture content has to be 20-25%, further at-least 80-85% of the grain in a panicle should be straw, also it is recommended to harvest between 110th – 135th day after sowing, Also, 25-35 days are standard requirements after heading, furthermore the surface moisture dries off at the time of harvest. When the above-mentioned variables are feed into the system as input, the system generates output in the Likert scale of (Highly Not Recommended – Highly Recommended) using the mean score derived. For the study, 150 agriculture experts and 150 farmers practising rice cultivation for more than 5 years were considered as the samples. Further 24 fields not less than a hectare were considered for the study purpose. The opinion of agricultural experts and farmers were compared against the opinion generated by the system on the Likert scale.

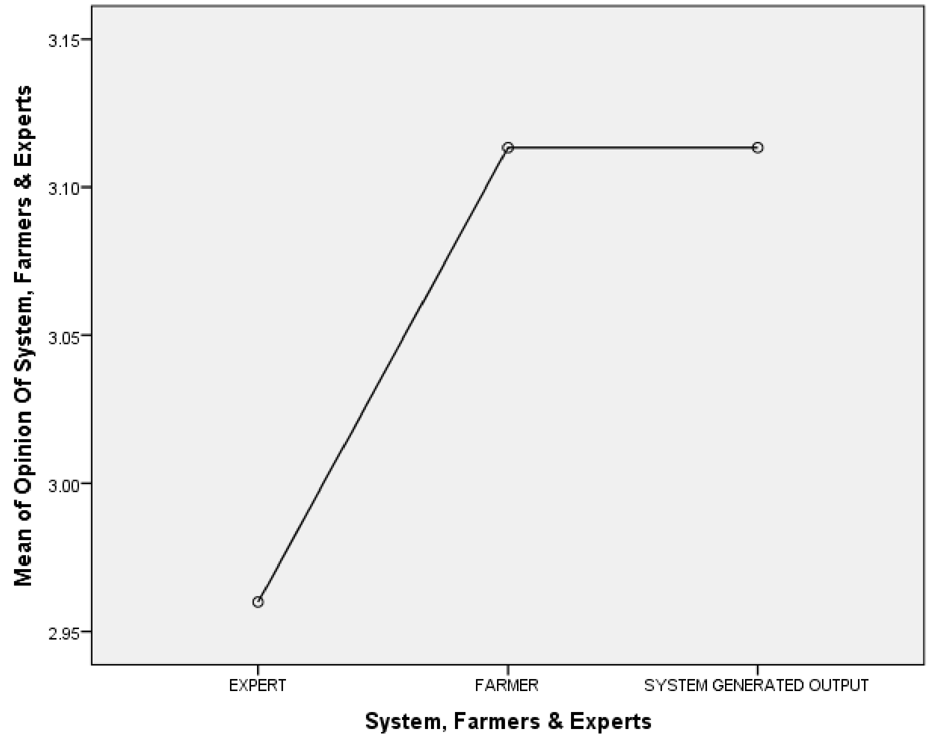
Analysis and Interpretation

Herein ANOVA test was performed to identify whether there is a significant difference in an opinion posted by an expert, farmer with system generated output.

Table No. 1: ANOVA Test

ANOVA					
Opinion Of System, Farmers & Experts					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.351	2	1.176	.584	.558
Within Groups	899.907	447	2.013		
Total	902.258	449			

Descriptive				
Opinion Of System, Farmers & Experts				
	N	Mean	Std. Deviation	Std. Error
EXPERT	150	2.9600	1.43251	.11696
FARMER	150	3.1133	1.37346	.11214
SYSTEM GENERATED OUTPUT	150	3.1060	1.44954	.11835
Total	450	3.0622	1.41756	.06682



Source: (Primary data)

The estimated significance value is 0.558, which is greater than 0.05. Meaning the null hypothesis is accepted i.e. there is no significant difference in an opinion posted by an expert, farmer with system generated output. The mean score estimated using the opinion value is; 2.96 for Experts, 3.1133 for farmers and 3.1060 for the system. This indicates that the system is generating output very lose to experienced farmed. Also, the standard deviation estimated is 0.06682, which indicates the system is efficient enough in producing accurate opinion close enough with experienced farmer and experts.

Having found the opinion given by expert and farmer is equivalent to system-generated output. The correlation analysis using the Pearson technique was adopted to identify the level of relationship in their opinion.

Table No. 2: Pearson Correlation Analysis

Correlations				
		Expert	Farmer	System Generated Opinion
Expert	Pearson Correlation	1	.773**	.791**
	Sig. (2-tailed)		.000	.000
	N	150	150	150
Farmer	Pearson Correlation	.773**	1	.728**
	Sig. (2-tailed)	.000		.000
	N	150	150	150
System Generated Opinion	Pearson Correlation	.791**	.728**	1
	Sig. (2-tailed)	.000	.000	
	N	150	150	150
**. Correlation is significant at the 0.01 level (2-tailed).				

Source: (Primary data)

The estimated significance value of all the variable considered for the study was found to be less than 0.05. This indicates the Null hypothesis is rejected, meaning there is a significant relationship between their opinion. From the correlation value, it can be interpreted that there is a 77.3% (estimated Correlation value is 0.773) relationship between opinions posted by an expert with experienced farmers. Similarly, there is a 79.1%(estimated Correlation value is 0.791) relationship between opinions posted by an expert with system generated output. Further, there is a 72.8% (estimated Correlation value is 0.728) relationship between opinion posted by experienced farmers with system generated output. Thereby, it is understood from the result that considering individual responses without considering the mean score, the system generated output is close to expert opinion.

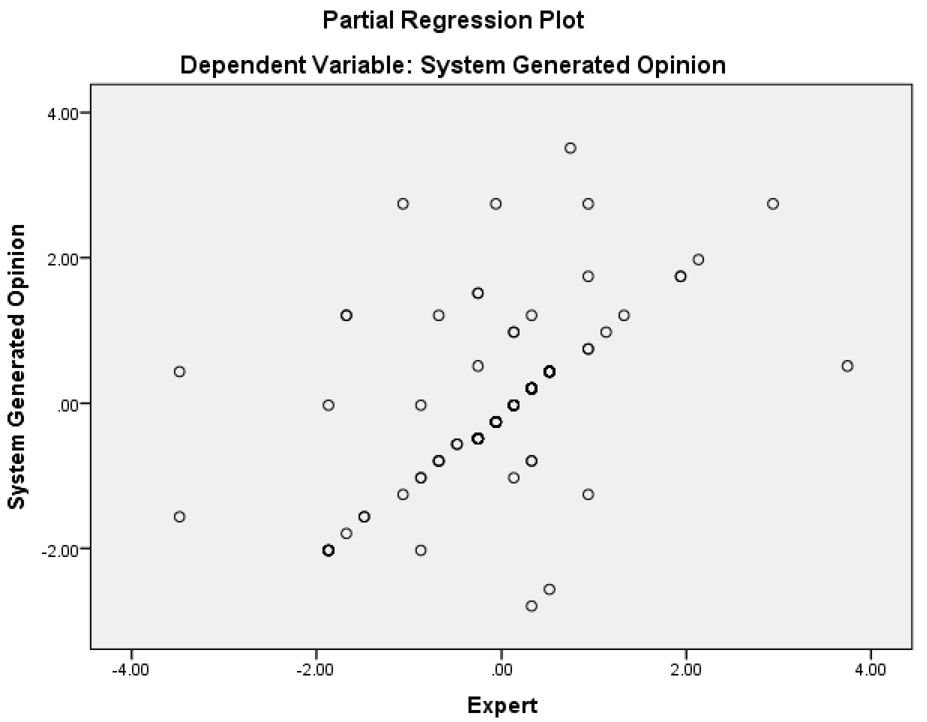
Having found there is a significant relationship between the opinion posted by an expert, farmer with system generated output. The regression analysis was carried to identify whether it is possible to predict the output of the system generated output using opinion posted by the experts and experienced farmers.

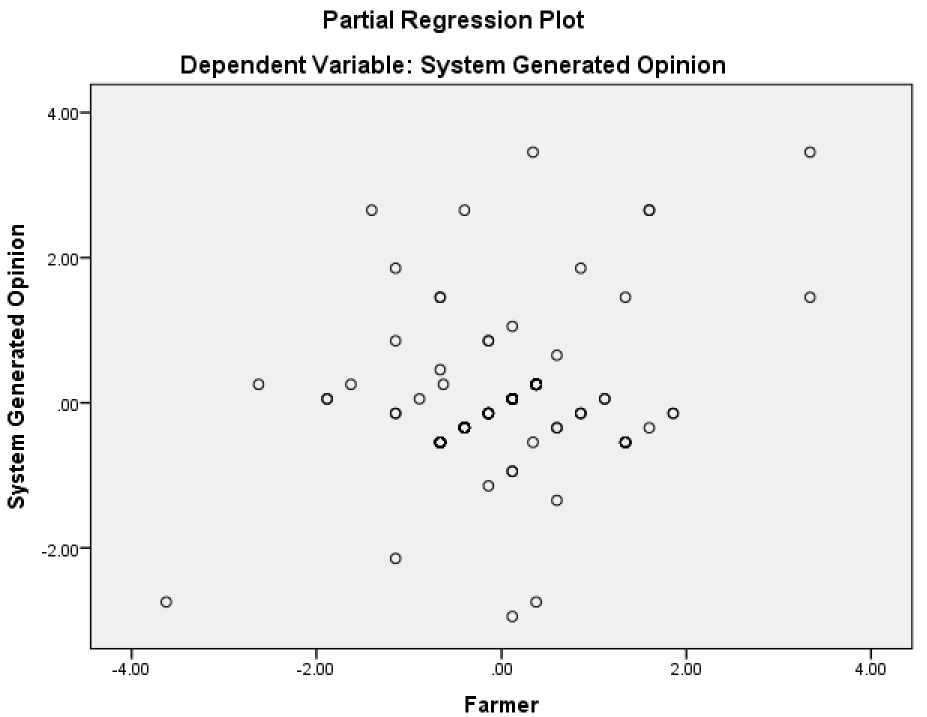
Table No. 3: Linear Regression Analysis

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.812[a]	.659	.655	.85170
a. Predictors: (Constant), Farmer, Expert				
b. Dependent Variable: System Generated Opinion				

ANOVA						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	206.441	2	103.220	142.296	.000[b]
	Residual	106.633	147	.725		
	Total	313.073	149			
a. Dependent Variable: System Generated Opinion						
b. Predictors: (Constant), Farmer, Expert						

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.463	.176		2.626	.010
	Expert	.573	.077	.566	7.457	.000
	Farmer	.307	.080	.291	3.830	.000
a. Dependent Variable: System Generated Opinion						





Source: (Primary data)

The estimated R-value is 0.812, meaning there is an 81.2% relationship among the considered variables. Further, the R-Square value is calculated to be 0.659 (Estimated Value > 0.6 – Acceptable Value) meaning, the regression equation derived will be 65.9% accurate in forecasting the system-generated output using the opinion posted by Experts and Experienced Farmers. Also, the ANOVA significance value is less than 0.05, meaning the model is fit. Furthermore, the estimated coefficient value indicates that the system generated output can be predicted using the opinion posted by Experts and Experienced Farmers. The regression equation is given;

System Generated Opinion = 0.463 + (0.573×Expert) + (0.307×Farmers)

From the regression plot, it can be additionally interpreted that most of the opinions posted by the Experts and Experienced farmers coincided with the opinion generated by the System using a machine learning algorithm.

Findings and Conclusion

From the result obtained through analysis, it can be well perceived that there is no significant difference in an opinion posted by an expert, farmer with system generated output. Also, the system is generating output very lose to experienced farmed. Further, the standard deviation estimated is very least, which indicates the system is efficient enough in producing accurate opinion close enough with experienced farmer and experts. From the correlation value, it can be interpreted that there is a 79.1%(estimated Correlation value is 0.791) relationship between opinions posted by an expert with system generated output. Further, there is a 72.8% (estimated Correlation value is 0.728) relationship between opinion posted by experienced farmers with system generated output. Thereby, it is understood from the result that considering individual responses without considering the mean score, the system generated output is close to expert opinion. Furthermore, the estimated coefficient value indicates that the system generated output can be predicted using the opinion posted by Experts and Experienced Farmers. The regression equation is given; System Generated Opinion = 0.463 + (0.573×Expert) + (0.307×Farmers). From the regression plot, it can be additionally interpreted that most of the opinions posted by the Experts and Experienced farmers coincided with the opinion generated by the System using a machine learning algorithm.

References

1. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., &Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, *18*(8), 2674.
2. Sharma, R., Kamble, S. S., Gunasekaran, A., Kumar, V., & Kumar, A. (2020). A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Computers & Operations Research*, *119*, 104926.
3. Morota, G., Ventura, R. V., Silva, F. F., Koyama, M., & Fernando, S. C. (2018). Machine learning and data mining advance predictive big data analysis in precision animal agriculture. *Journal of Animal Science*.
4. Mekonnen, Y., Namuduri, S., Burton, L., Sarwat, A., & Bhansali, S. (2019). Machine learning techniques in wireless sensor network based precision agriculture. *Journal of the Electrochemical Society*, *167*(3), 037522.
5. Banerjee, A., Mitra, A., & Biswas, A. (2021). An Integrated Application of IoT- Based WSN in the Field of Indian Agriculture System Using Hybrid Optimization Technique and Machine Learning. *Agricultural Informatics: Automation Using the IoT and Machine Learning*, 171-187.
6. Fashoto, S., Mbunge, E., Ogunleye, G., & Van den Burg, J. (2021). Implementation of machine learning for predicting maize crop yields using multiple linear regression and backward elimination. *MALAYSIAN JOURNAL OF COMPUTING*, *6*(1), 679-697.
7. Sharma, R., Kamble, S. S., Gunasekaran, A., Kumar, V., & Kumar, A. (2020). A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Computers & Operations Research*, *119*, 104926.
8. Wei, M. C. F., Maldaner, L. F., Ottoni, P. M. N., & Molin, J. P. (2020). Carrot Yield Mapping: A Precision Agriculture Approach Based on Machine Learning. *AI*, *1*(2), 229-241.
9. Kale, S. S., & Patil, P. S. (2019). Data mining technology with fuzzy logic, neural networks and machine learning for agriculture. In *Data management, analytics and innovation* (pp. 79-87). Springer, Singapore.

10. Yahata, S., Onishi, T., Yamaguchi, K., Ozawa, S., Kitazono, J., Ohkawa, T., ... & Tsuji, H. (2017, May). A hybrid machine learning approach to automatic plant phenotyping for smart agriculture. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 1787-1793). IEEE.
11. Dr.A.Senthil Kumar, Dr.G.Suresh, Dr.S.Lekashri, Mr.L.Ganesh Babu, Dr. R.Manikandan. (2021). Smart Agriculture System With E – Carbage Using Iot. International Journal of Modern Agriculture, 10(01), 928 - 931. Retrieved from http://www.modern-journals.com/index.php/ijma/article/view/690
12. Dr.G.Suresh, Dr.A.Senthil Kumar, Dr.S.Lekashri, Dr.R.Manikandan. (2021). Efficient Crop Yield Recommendation System Using Machine Learning For Digital Farming. International Journal of Modern Agriculture, 10(01), 906 - 914. Retrieved from http://www.modern-journals.com/index.php/ijma/article/view/688