Binary Logistics Regression Model to Analyze Factors Influencing Technology Adoption Process Vegetable Farmers Case in Central Java Indonesia

Muhamad Rusliyadi[1], Yanuar Wahyu Yudistira Ardi[1](✉), and Kodrad Winarno[2]

[1] Polytechnic of Agricultural Development Yogyakarta-Magelang, Magelang, Indonesia
yanuarwya@gmail.com

[2] Agricultural Development Polytechnic of Yogyakarta - Magelang, Yogyakarta, Indonesia

Abstract. Agricultural extension activities as a form of educational facilities for the community, especially farmers have an important role in making changes with innovations delivered to vegetable farmers. The adoption process of innovation can affect the occurrence of behavior changes that can be observed directly or indirectly. This study aims to determine the application of binary logistics regression in analyzing the factors that influence the adoption process of innovative technology in vegetable farmers. The results of the research after some analysis of stages such as the formation of the initial model test, simultaneous tests, partial test, and model match test obtained. The results of factors that affect the adoption process of technology is a factor in the number of livestock and vegetable with a significance value of 0.006 ($p < 0.05$), an intensity factor of extension with a significance value of 0.039 ($p < 0.05$), a factor in the distance of residential of women farmers groups to the source of innovation with a significance value of 0.020 ($p < 0.05$), and the distance factor of the residential of farmers to the source of capital with a significance value of 0.010 ($p < 0.05$). The research showed that the probability factor influencing of technology adoption process base on the distance farmer from the capital and sources innovation, the intensity of extensions and also number of livestock and vegetable.

Keywords: Agricultural extension · Logistic regression · Adoption process · Innovation · Vegetable farmers · Binary logistic

1 Introduction

The role of agriculture in the national economic development is very important and strategic. This is mainly due to the fact that the agricultural sector still provides jobs for the majority of the population in rural areas and provides food for the population [1]. The Agriculture movement depends on the participation in agricultural extension activities are a form of activity for the community which is a means of education outside of school (non-formal), with the aim of improving community welfare, especially for farmers and their families [2]. Through agricultural extension activities there are various new

A. G. Abdullah et al. (Eds.): SEAVEG 2021, ABSR 23, pp. 460–470, 2023.
https://doi.org/10.2991/978-94-6463-028-2_48

innovations needed by the community and the stakeholder [3]. This innovation must be conveyed to the community, in this case farmers/breeders who do business traditionally, so that they can change their behavior. Where it is expected that all new innovations that are outsourced can be well absorbed through a directed adoption process [4].

The process of adopting innovation can be seen from the occurrence of changes in behavior (knowledge, attitudes, and skills) of the target that can be observed directly or indirectly. The adjustment speed or opportunity for innovation adoption can come from the characteristics of respondents such as age, education, family dependents, experience, land/business control and can come from business accessibility such as the distance of settlements to roads, sources of income. Innovation, and capital [5]. So that in implementing the technology adoption process, statistical analysis is needed to support the increasing adoption of innovation, technology, which in this paper aims to determine the implementation of binary logistic regression in analyzing the factors that influence the innovation technology adoption process and to find out the best model of binary logistic regression analysis.

Logistic regression is a statistical analysis method used to predict the relationship between the response variable (dependent variable) which has two or more categories with one or more explanatory variables (independent variable) on a category or interval scale [6]. The description of the relationship between response variables that have qualitative or categorical properties with explanatory variables that have two or more categories cannot be solved by ordinary linear regression models using the Ordinary Least Square (OLS) method [7]. If the linear regression method is forced to analyze data whose response variables have the characteristics as mentioned above, there will be a violation of the Gauss-Markov assumption [8]. Haridanti et al. [9] state that the logistic regression analysis is a regression analysis used to describe the relationship between the response variable (outcome or dependent) and a set of predictor variables (explanatory or independent), where the response variable is binary or dichotomous. The response variable is dichotomous qualitative data with a value of 1 (one) to indicate the occurrence of an event and a value of 0 (zero) to indicate the non-occurrence of an event. The general form of the logistic regression equation model is formulated as follows:

$$\pi_{(j)} = P(Y = 1|X) = \frac{e^{\beta_0 + \beta_{10}x_{11} + \beta_{2xj2} + \ldots + \beta_p X_{jp}}}{1 + e^{\beta_0 + \beta_{10}x_j + \beta_{22j} + \ldots + \beta_p X_p}} \tag{1}$$

The link function used is logit, with the logit of π ie;

$$\text{logit}(\pi) = \log\left(\frac{\pi_j}{1 - \pi_j}\right) = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots + \beta_p x_{jj} \tag{2}$$

The logistic regression model is used to determine the probability or the occurrence of a situation. Logistic regression is a non-linear regression, used to explain the non-linear relationship between X and Y [7]. In contrast to Ordinary Least Square (OLS) or linear regression, logistic regression does not require a linear relationship between the independent and dependent variables. The data also do not have to be normally distributed [10].

This study aims to determine the application of binary logistics regression in analyzing the factors that influence the adoption process of innovative technology in vegetable farmers. The results of the research after some analysis of stages such as the formation of the initial model test, simultaneous tests, partial test, and model match test obtained.

2 Methodology

2.1 Data Collection Method

The data used in this study are primarily data resulting from the as a study in the Ngudi Rahayu Farmers Group in Samirono Village Semarang, Central Java Province in 2020. The sample was determined by purposive sampling with a total sample of 60 farmer members.

Analysis of the data used in the implementation of the assessment is descriptive analysis and binary logistic regression analysis. Descriptive analysis is used to determine the stages of adoption and the level of adoption that occurs. The stages of adoption, according to Harinta [11] consist of stages: awareness, interest, evaluation, trial, and adoption. While the adoption rate is classified into three criteria, namely low (0.0–33.3%), moderate (33.4–66.7%), and high (66.8–100%) [12]. The adoption measurement activity refers to Abdullah's [13] research which measures adoption time into three time stages, starting from 0–1 weeks, more than 1–3 weeks, and 3–5 weeks.

2.2 Research Variable

Based on the data sources obtained, data on technology adoption, characteristics of women farmers, and regional accessibility are as follows:

Technology adoption (Y) is a categorical variable as an adoption variable

Y = 1 to adopt/accept

Y = 0 for not adopting/not accepting

As for the predictor variable (X)

1. Age (X1) is a continuous variable
2. Education (X2) is a categorical variable, with values: 1 = SD, 2 = SMP, 3 = SMA, 4 = S1/D4
3. The number of family dependents (X3) is a continuous variable
4. Farming experience (X4) is a continuous variable
5. The number of cattle/vegetable unit (X5) is a continuous variable
6. Extension intensity (X6) is a continuous variable
7. Courage in taking risks (X7) is a categorical variable, worth: 1 = not daring, 2 = doubtful, 3 = daring
8. Distance from settlement to road (X8) is a continuous variable
9. Distance from settlement to the source of innovation (X9) is a continuous variable
10. Distance from settlement to sources of capital (X10) is a continuous variable.

2.3 Data Analysis Stages

The data processing uses IBM SPSS Statistics 22.0 Software with the following stages of analysis:

- Inputting data on technology adoption, characteristics of women farmers, and regional accessibility of women farming members
- Analyzing a binary logistic regression model of technology adoption process data
- Conduct Simultaneous Test using the Likelihood Ratio Test
- Performing Partial Test uses Wald Test
- Analyzing with Model Fit Test uses Hosmer and Lemeshow test.

2.4 Logistic Regression Model of Technology Adoption

$$\begin{aligned}\mathrm{Ln}\frac{Pi}{(1-P)} =& \alpha + \beta_1 \text{ LnUmur } + \beta_2\text{LnPnd} + \beta_3\text{LnTgKlrg} + \beta_4 \text{ Pglmn} \\ &+ \beta_5 \text{ JlTrnk } + \beta_6 \text{ ItPyl } + \beta_7 \text{ Kbrmian } + \beta_8 \text{ JrkJalanry } n \\ &+ \beta_9 \text{ Jrksmbinvs } + \beta_{10} \text{ Jrksmbmodl } + \mathrm{e}n\end{aligned}$$

Note:
Ln Y = Adoption of farm women, (1 = adopted, 0 = Not adopted)
α = Constant
βi = regression coefficient (i = 1,2,3,…..,10)
LnUmr = Age (years)
LnPnd = Education (level; SD, SMP, SMA, S1/D4)
LnTgKlrg = Family dependent (soul)
LnPglmn = Farming experience (years)
LnJlTrnk = Number of cattle/vegetable (heads/units)
LnItPyl = Extension intensity (meetings)
LnKbrnian = Courage (level; not daring, indecisive, daring)
LnJrkJalanry = Distance of settlement to highway (m)
LnJrksmbinvs = Distance from settlement to the source of innovation (m)
LnJrksmbmodl = Distance from settlement to sources of capital (m)
e = disturbance term or disturbing factor (error).

3 Results and Discussion

3.1 Data Description

3.1.1 Characteristics of Women Farmers Group

Overall, the respondents' age consisted of 85% in productive age (15–60 years) and the remaining 15% in unproductive age, with education level of 56.67% being elementary school graduates, 18.33% high school graduates, 13.33% S1/D4 graduates, and 11.67% junior high school graduates. The number of dependents in the family is dominated by the number of dependents 1 person by 71.67%, (another 26.67% the number of dependents of the family is 2–3 people and 1.66% the number of dependents of the family 4. For the experience of raising the respondent, it is 70% respondents have experience in livestock 10 years and the remaining 30% 10 years The number of livestock owned by respondents ranges from 0–4 heads with a percentage of 78.33% and others 5–9 heads by 15%, 10–14 heads by 5%, and 15 tails by 1.67%. Based on the intensity of the extension, only 21.66% of respondents attended extension 2 meetings and the courage to take risks only 6.67% of respondents stated their courage in taking risks which was considered from courage in taking risks, a readiness to take risks, and liking in trying.

3.1.2 Area Accessibility

Overall the accessibility of the area in Samirono Village is in good condition. The accessibility of settlements to roads is generally conducive, because geographically, the Samirono Village area is traversed by alternative provincial roads that connect other districts. Respondents' access to sources of innovation is also not too difficult, it's just that the contours of the mountains cause the roads to be taken sometimes are different. As for the distance to the capital is about 2.7 km. It can be said that it is relatively close because of the decent road conditions and light traffic.

3.2 Binary Regression Model for Technology Adoption Process

3.2.1 Early Model

The estimation of logistic regression parameters used in this initial model is the maximum likelihood method followed by Newton Rapshon iteration. The estimated value of the initial model parameters is explained as follows:

Based on the estimated value (B) of each variable in Table 1. The initial model can be carried out as follows:

$$\pi(X) = \frac{e^{g(x)}}{1 + e^{g(x)}} \tag{3}$$

With the value of g(x) is:
g(x) = −18,219 – 0,37age + 0,619 education (1)
+0,442 education (2) – 0,395 education (1)
+0,476Sum_dep_family – 0,01Experince_farm
+0,399Sum_catveg + 2,044 _Intensity_extens
+1,598Courage_take_risk (1)
–1,788598Courage_take_risk (2)
–0,012Distance_street
+0,010Distance_source_inovat
+0,006Distance_tocapital

3.2.2 Simultaneous Test

Simultaneous testing uses the likelihood ratio to test the feasibility of the model obtained from the parameter estimation results. The simultaneous test aims to determine whether the predictor variables contained in the model have a significant effect on the whole or

Table 1. Estimated value of initial model parameters

Step 1[a]	B	S.E.	Wald	df	Sig.	Exp(B)
Age	−0,037	0,052	0,512	1	0,474	0,963
Education			0,720	3	0,869	
Education (1)	0,619	1,188	0,271	1	0,602	1,857
Education (2)	0,442	1,243	0,126	1	0,722	1,555
Education (3)	−0,395	1,315	0,090	1	0,764	0,673
Sum_dep_family	0,476	0,420	1,280	1	0,258	1,609
Experience_farm	−0,001	0,039	0,001	1	0,982	0,999
Sum_cattveg	0,399	0,146	7,440	1	0,006	1,491
Intensity_extens	2,044	0,990	4,262	1	0,039	7,723
Courage_take_risk			3,527	2	0,171	
Courage_take_risk (1)	10,598	1,160	1,900	1	0,168	4,945
Courage_take_risk (2)	−10,788	2,589	0,477	1	0,490	0,167
Distance_street	−0,012	0,007	3,026	1	0,082	0,988
Distance_sources_inovat	0,010	0,004	5,377	1	0,020	1,010
Distance_tocapital	0,006	0,002	6,716	1	0,010	1,006
Constant	−18,219	6,755	7,274	1	0,007	0,000

Source: Processed Primary Data 2021

Table 2. Statistical value of likelihood ratio test

Step	−2 Log likelihood	Cox & Snell R Square	Negelkerke R Square
1	51,017[a]	0,382	0,519

Sources: primary data 2021

not. Simultaneous testing can use the likelihood ratio test with the G test statistic that follows a chi square distribution with degrees of freedom of one [6].

- Hypothesis
  H0: $\beta 1 = \beta 2 = \ldots = \beta 10 = 0$
  H1: at least there is value $\beta j \neq 0$ with $j = 1,2,3,\ldots,10$
- Significance level $\alpha = 5\% = 0,05$
- Test Statistic

Base on a value of −2 log likelihood = 51,017 on the Table 2 obtained:

$$G = -2\ln\left(\frac{likelihood\_\tan pa\_\text{variable\_bebas}}{likelihood\_dengan\_\text{variable\_bebas}}\right)$$

G = 79,881 − 51,017
G = 29,864
Critical area_Reject H0 if G > $X^2_{(10;0,05)}$.
Decision: H0 rejected because $G\cdot = -29,864 > -X^2_{(10,0,05)} = 16,92$.

Conclusion According to the table can be concluded that at a significance level of 5% the predictor variables contained in the model are: age, education level, number of family dependents, farming experience, number of livestock and vegetable, intensity of extension, courage in taking risks, distance of the settlement to the highway, distance of the settlement to the source of innovation, the distance of settlement of capital has a significant effect simultaneously.

3.2.3 Partial Test

Partial test using the Wald test, has the aim of knowing the significance of the parameters on the predictor variables.

Hypothesis
H0: βj = 0
H1: βj ≠ 0 where j = 1,2,3,….,10
Significance level

$$\alpha = 5\% = 0,05$$

$$Wj = \left\{\frac{\beta j}{se(\beta j)}\right\}^2 \quad (4)$$

- Critical area
  Reject H0 if Wj > or reject H0 if sig. < α
- Decision
  Based on the Wald value for each variable and its significance in Table 3, the results obtained from the Wald test are as follows:

Conclusion at a significance level of 5%, the variable number of livestock and vegetable, intensity of extension, distance of the settlement to the source of innovation, and distance of settlement of capital has a significant effect on the process of adopting technology partially. Meanwhile, the variables of age, education level, number of family dependents, experience in raising livestock and vegetable, courage in taking risks, and the distance of the settlement to the highway have no significant effect on the technology adoption process.

Table 3. Decisive result of the wald test

Variable	Wald	Df	$X^2(10;0,5)$	Sig.	Decision
Age	0,512	1	3,84	0,474	H0 accepted
Education	0,720	3	7,81	0,869	H0 accepted
Education (1)	0,271	1	3,84	0,602	H0 accepted
Education (2)	0,126	1	3,84	0,722	H0 accepted
Education (3)	0,090	1	3,84	0,764	H0 accepted
Sum_dep_family	1,280	1	3,84	0,258	H0 accepted
Experience_farm	0,001	1	3,84	0,982	H0 accepted
Sum_cattveg	7,440	1	3,84	0,006	H0 rejected
Intensity_extens	4,262	1	3,84	0,039	H0 rejected
Courage_take_risk	3,527	2	5,99	0,171	H0 accepted
Courage_take_risk (1)	1,900	1	3,84	0,168	H0 accepted
Courage_take_risk (2)	0,477	1	3,84	0,490	H0 accepted
Distance_street	3,026	1	3,84	0,082	H0 accepted
Distance_sources_inovat	5,377	1	3,84	0,020	H0 rejected
Distance_tocapital	6,716	1	3,84	0,010	H0 rejected

Source: Primary data 2021

Table 4. Value of hosmer and lemeshow test statistics

Step	Chi-square	df	Sig.
1	6,212	8	0,623

Sources: Primary Data 2021

3.2.4 Model Fit Test

The suitability test of the model used is the Hosmer and Lemeshow test, which aims to test the suitability of the binary logistic regression model.

Hypothesis
H0: suitable model
H1: the model does not fit
Level of significance
$\alpha = 5\% = 0.05$
Test statistics
Based on Table 4 obtained the value that
$\hat{C} = 6{,}212$
Sig. = 0,623

Table 5. Classification accuracy

Observed		Predicted		
		Adoption Decision		
		Not Accepted	Accepted	Percentage Correct
Step 1 Adoption decision	Not Accepted	32	5	86,5
	Accepted	8	15	65,2
Overall Percentage				78,3

Sources: Primary data 2021

- Critical area
  Reject H0 if $\widehat{C} > X^2_{(10,0,05)} = 16{,}92$ or sig $< \alpha$.
- Decision
  $\widehat{C} = 6{,}212 < X^2_{(10;0,05)} = 16{,}92$ or sig $= 0{,}623 > \alpha = 0{,}05$, then H0 accepted.

  Conclusion at the 5% significance level the binary logistic regression model is appropriate. After testing the significance of the model, both simultaneously and partially, the results obtained that the variable number of livestock, intensity of extension, distance of settlement to the source of innovation, and distance of settlement of capital has a significant effect on the process of adopting the technology.

3.2.5 Size Classification Accuracy

The classification table explains how well the model grouping cases into two groups, both those who accept technology adoption and those who do not accept technology adoption (Table 5).

Based on the classification accuracy, it can be seen that the overall prediction accuracy is 78.3%. While the prediction accuracy for those who do not accept technology adoption is 86.5% and the prediction accuracy for those who accept the technology adoption process is 65.2%.

3.3 Factors Affecting of the Respondents Characteristics and Area Accessibility to the Technology Adoption Process

The factors that influence the results of the binary logistic regression test are the number of livestock with a significance value of 0.006 ($p < 0.05$), the intensity of extension factor with a significance value of 0.039 ($p < 0.05$), the distance factor of the female farmers' settlement to the source innovation significance value is 0.020 ($p < 0.05$), and the distance factor of women's farmer settlements to sources of capital with a significance value of 0.010 ($p < 0.05$). For the interpretation of the model in binary logistic regression can use the value of the odds ratio. The value of the odds ratio aims to determine how much the number of livestock factors, intensity of extension, distance of settlements to sources of innovation, and distance of settlements to sources of capital affect the technology adoption process. The interpretation of the value of the logit regression coefficient (B) is carried out by calculating the antilog of B or the exponential power of the logit coefficient (exp(β)) [14].

4 Conclusion and Recommendation

4.1 Conclusion

The implementation of binary logistic regression analysis shows that the factors of influence the technology adoption process are the number of livestock and vegetable farm with a significance value of 0.006 ($p < 0.05$. The extension intensity factor with a significance value of 0.039 ($p < 0.05$). The factor of distance women's farmer settlements to sources of innovation with a significance value of 0.020 ($p < 0.05$), and the factor of distance of women's settlements from sources of capital with a significance value of 0.010 ($p < 0.05$). The research showed that the probability factor influencing of technology adoption process base on the distance farmer from the capital and source innovation, the intensity of extensions and also number of livestock and vegetable.

4.2 Recommendation

- To find out the factors that really influence by the binary logistic regression model, needed further tests by using the stepwise method.
- The results of the analysis of the influencing factors in the binary logistic regression model can be used in determining and considering targets for future extension plan activities.

References

1. M. Rusliyadi, and A.B. Jamil, "Participatory Poverty Assessment Effort in Food Security and Extension Policy: Evidence From Indonesia." In V. Erokhin & T. Gao (Eds.), Handbook of Research on Globalized Agricultural Trade and New Challenges for Food Security, pp. 481–499, 2020. IGI Global. https://doi.org/10.4018/978-1-7998-1042-1.ch024.
2. Padmanagara, Salmon, Bapak Penyuluhan Pertanian Pengabdi Petani Sepanjang Hayat, p. 4, Jakarta: PT. Duta Karya Wisata, 2012.
3. M. Rusliyadi, A. Bin Hj. Mohd. Jamil, A. Maseleno, R. Tri Kumalasari, M. Othman, "Agricultural extension policy, agricultural growth and poverty reduction in Indonesia," IACSIT International Journal of Engineering and Technology, 7(4), 5539–5550, 2019. https://doi.org/10.14419/ijet.v7i4.1333.
4. Isbandi, Penyuluhan Untuk Pembaharuan Perilaku. Semarang: Badan Penerbitan Universitas Diponegoro, 2011.
5. Burhansyah, Rusli. Faktor-Faktor yang Mempengaruhi Adopsi Inovasi Pertanian PUAP dan Non PUAP di Kalimantan Barat (Studi Kasus: Kabupaten Pontianak dan Landak). Jurnal Informatika Pertanian vol. 23, no. 1, pp. 65–74, Pontianak: Balai Pengkajian Teknologi Pertanian Kalimantan Barat, 2013.
6. D.W. Hosmer, and S. Lemeshow, "Applied Logistic Regression," John Wiley and Sons Inc., New York, 2000.

7. R. Hendayana, Penerapan Metode Regresi Logistik dalam Mengenalisis Adopsi Teknologi Pertanian. Jurnal Informatika pertanian, vol. 22, no. 1, pp. 1–9, Bogor: Balai Besar Pengkajian dan Pengembangan Teknologi Pertanian, 2012.
8. M.H. Kutner, C.J. Nachtsheim dan J.Neter. "Applied Linear Regression Models. Fourth Edition," The Mc Graw-Hill Companies Inc., Singapore, 2004.
9. S. Haridanti, A. Robiatul, S.A. Gita, F.Y.A. Putri, and Z. Ulfa, "Analisis Regresi Non Linear Model Logistik (Studi Kasus: Lembaga Pelatihan Kerja Kabupaten Slemman, Yogyakarta)." Jurnal Seminar Nasional Inovasi dan Teknologi di Industri 2018, p. 63, Malang: Institut Teknologi Nasional Malang, 2018.
10. F. Janssen, "Do Managers' Characteristics Influence the Employment Growth of SMEs?." PP. 293–315. Journal of Small Business & Entrepreneurship, 19 (3), 2006.
11. Y.W. Harinta, "Hubungan Sosial Ekonomi Petani dan Kualifikasi Penyuluh Terhadap Adopsi Teknologi Pengelolaan Tanaman dan Sumber daya Terpadu Budi daya Padi (Oryza sativa)." Jurnal Widyatama, 19(2), 2010.
12. A.S. Romdon, S. Supardi, & L.A. Sasongko, "Kajian Tingkat Adopsi Teknologi Pada Pengelolaan Tanaman Terpadu (PTT) Padi Sawah (Oryza sativa L) di Kecamatan Boja Kabupaten Kendal," Mediagro: Jurnal Ilmu-Ilmu Pertanian, 8(1), 2012.
13. M.A Abdullah, S.S. Dlay, W.L Woo & J.A. Chambers, "Metode segmentasi iris yang kuat berdasarkan gaya kontur aktif baru dengan normalisasi nonsirkular." *Transaksi IEEE pada sistem, manusia, dan sibernetika: Sistem*, *47* (12), 3128–3141, 2016.
14. A. Widarjono, "Analisis Statistika Multivariat Terapan," Yogyakarta: Unit Penerbitan dan Percetakan STIM YKPN, 2010.