The Bayesian Quantile Regression and Rought Set Classification Analysis of Taiwanese People's Satisfaction Level with the Government in 2014

SHIANGHAU WU
School of Business
Macau University of Science and Technology
Avenida Wai Long, Taipa, Macau
CHINA
shwu@must.edu.mo

Abstract: The general public's satisfaction is vital to the success of the governance. In this paper, we propose to utilize the Bayesian quantile regression to find the variables that affect the satisfaction of the general public at a specific quantile of service satisfaction. The result of the Bayesian quantile regression analysis provides the public opinion toward the Kuomintang(KMT)-ruled central government in 2014 at different quantiles of governance satisfaction. We also compare the prediction accuracy of regression models at different quantiles of service satisfaction and two algorithms of the rough set classification. The experiment result shows the major factors which have the positive relationship with the people who have higher satisfaction level with the central government. These factors include:(i)satisfaction with the uncorrupted performance of the central government; (ii)evaluation of your household's economic condition one year after; (iii)satisfaction with the central government's measures on food safety.(iv) satisfaction with the twelve years primary education reform.

Key–Words: Bayesian Quantile Regression, satisfaction; rough set

1 Introduction

In recent years,there is a growing concern on the measurement of government performance. Since the government performance can not be objectively measured, some studies focused on the usage of citizen surveys[1][2].The factors which affect the government performance were diverse in recent studies. However, many scholars agree that the increasing efficiency and government transparency are pivotal factors of good government performance[3]. In contrast to the studies related to the impact of ideological factors on Taiwan issues, the study only focused on the domestic affairs issues on the Taiwanese people's satisfaction with the government[4].
Among the numerous studies on Taiwan people's satisfaction with the government, this study sources the data of Taiwans Election and Democratization Study for analysis [5]. In this study, we attempted to analyse the variables that affected the levels of Taiwanese people's satisfaction with the government with Bayesian quantile regression on 841 effective survey respondents of samples in Taipei city, Taiwan at the different quantiles of satisfaction and also used two algorithms for rough set classification. The results provided the Taiwan's ruling Kuomintang party's fail in the general election in 2014 with information required to justify the reasons why citizens were discontent at different quantiles of service satisfaction . Finally, we make the comparison of the performance of quantile regressions at different quantiles in terms of three different measures for performance evaluation and also compare the classification results among Bayesian quantile regression and results of two algorithms in roughsets classification.

The remainder of this paper is organized as follows. Section 1 introduces the research motivations and purposes. Section 2 is the literature review on the Bayesian quantile regression model. Section 3 discusses the sample data and the results of Bayesian quantile regression. Section 4 offers results of two algorithms in rough set classification. Section 5 is the comparison of classification results in the three models. Section 6 is the conclusions and suggestions of this paper.

2 Research Method

2.1 Quantile Regression

Koenker and Bassett (1978) proposed the quantile regression method[6]. Quantile regression aims at exploring the effects of explanatory variables (X) on the explained variable (Y) at different quantiles of the values of the explained variable. It is different from conventional linear models, such as the least squares method, which predict the mean of the explained variable given a specific value of each explanatory variable. The quantile regression model can be represented as:

$$Y_i = X_i\beta_\theta + \mu_{\theta i} \tag{1}$$

$$X_i\beta_\theta = (Quantile)_\theta(Y_i \mid X_i) \tag{2}$$

The quantile regression method can predict the value of the explained variable at a specific quantile of the explained variable. Therefore, quantile regression allows an understanding of the effects of different explanatory variables on the explained variable at different quantiles of the values of the explained variable. Therefore, quantile regression has been widely used in many applications[7][8][9].

As for the model of binary dependent variable, the equation of quantile regression can be represented as[10]:

y_i*=$x'_i\beta_\tau + \mu_i$

y_i =1 if $y_i^* \geq$0, y_i =0 otherwise (3)

y_i* is a latent continuous variable which determines the binary dependent variable y_i.Xi is a 1×k vector of the explanatory variables, and $\beta\tau$is a $k\times 1$ vector of unknown parameters to be estimated for different values of $\tau$and $\mu_i$is a random error term that is independently and identically distributed. In the original formulation, the only requirement of the distribution of $\mu$is that the $\tau-$th quantile equals 0. The $\tau-$th quantile of y_i* on the explanatory vector Xi can be represented as:

$$Q_\tau(y_i^* \mid x_i) = x'_i\beta_\tau \tag{4}$$

Kordas (2006) stipulated the probabilistic prediction method for the binary quantile regression model. By changing the value $\tau$from 0 to 1 it is possible to get the predicted distribution of y* given x and then we can estimate the probability of y which takes value of 0 and 1 [11].

Variable	Mean	Std	Minimum	Maximum
Y	1.86564	0.75421	1.00000	4.00000
X2	2.03686	0.83371	1.00000	4.00000
X3	1.73960	0.73515	1.00000	3.00000
X4	1.65636	0.83216	1.00000	3.00000
X5	1.39596	0.68933	1.00000	3.00000
X6	1.52319	0.82937	1.00000	3.00000
X7	1.42568	0.65792	1.00000	4.00000
X8	1.63734	0.71246	1.00000	4.00000

Table 1: The Descriptive Statistics of the Dataset.

2.2 Sample Data and Variables

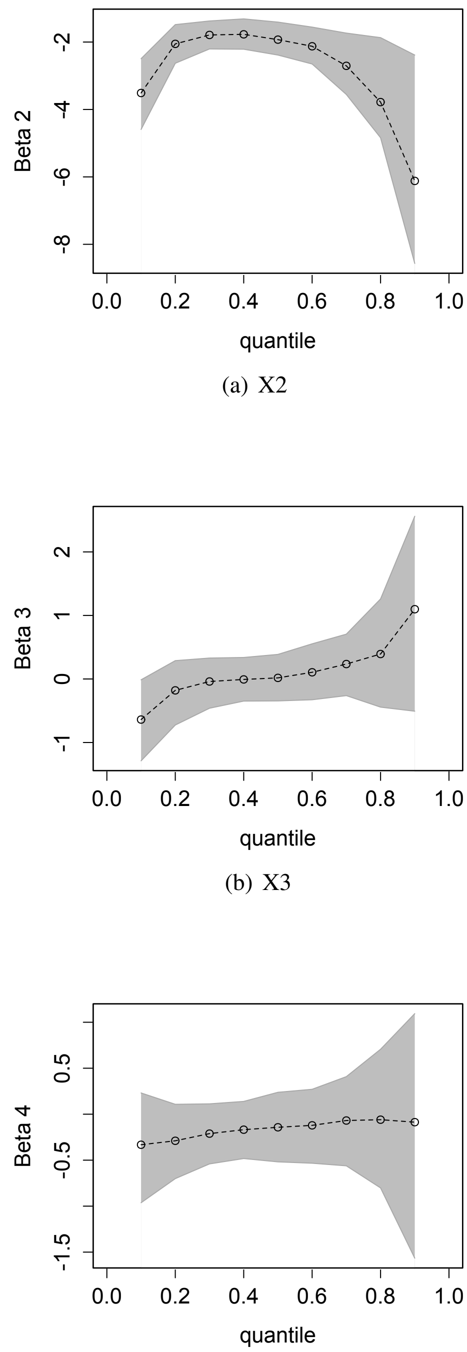
Based on the questionnaire survey of Taiwans Election and Democratization Study conducted by National Chengchi University (NCCU) on the local election of Taiwan in 2014, we analyzed the satisfaction of Taiwanese people toward the government by using the collected questionnaire data. The data set consists of interview responses from three major cities, Taipei, Taichung and Tainan. The variables designed for Taiwanese people satisfaction model include ”satisfaction with the central government in these three years” (Y, the dependent variable), ”satisfaction with the uncorrupted performance of the central government” (X2), ”evaluation of Taiwan's current economic condition as compared with a year ago” (X3), ”evaluation of Taiwan's economic condition one year after” (X4), ”evaluation of your household's economic condition as compared with a year ago” (X5), ”evaluation of your home's economic condition one year after” (X6), ”satisfaction with the central government's modus operandi on food safety” (X7), ”satisfaction with the central government's policies on the twelfth years primary education reform” (X8). The study chose the Taipei city data as the sample data. The descriptive statistics of the dataset included 841 survey responds was shown on table 1.The satisfaction level ranges from 1 (the lowest) to 4(the highest).

3 Empirical Analysis: Bayesian Quantile Regression

The study used the Bayesian quantile regression method to analyze the effects of exploratory variables from X2 to X8. And the study classified the independent variable data as two groups; the satisfaction level of 1 and 2 belongs to group 1 and the satisfaction level of 3 and 4 belongs to group 0. The study analyzed the regression parameters for the quantile levels $\tau$from 0.05 to 0.95. The study computed 90% pointwise Bayesian credible interval from the marginal posterior of each parameter. The results showed that the confidence intervals of regression parameters X3, X4, and X5 overlap the value of zero on practically all quantile levels in accordance with Figure 1 and Figure 2. It means the regressors X3, X4 and X5 are not important for the estimation of the Bayesian quantile regression model.

Table 2 to Table 4 represent the Bayesian quantile regression coefficients at 0.25, 0.5 and 0.75 quantile level. According to Figure 1, Figure 2 and Figure 3, the coefficients of X2 increases from the 0.1 quantile level to 0.4 quantile level (-3.5144 to -1.77), but gradually decreases to -6.1194 at 0.9 quantile level. As mentioned earlier, the study coded the lowest satisfaction level as 1, highest satisfaction level as 0. The fluctuation of X2 at different quantile levels means X2 (satisfaction with the uncorrupted performance of the central government) has the decreasing positive relationship with people who have the lower satisfaction level with the central government; while X2 has the increasing positive relationship (decreasing negative coefficients) with people who have the higher satisfaction level with the central government.
The study also found the coefficients of X6 (evaluation of your home's economic condition one year after) remained at -0.2 to -0.7 in almost all quantiles, and only when the quantile reached 0.9 was the coefficient became -1.0539. It means X6 does not have significant relationship with the dependent variable. As for the coefficients of X7, it fluctuated in (-1,-2) from 0.1 to 0.7 quantile level, while decreased from 0.8 quantile level. It means the satisfaction with the central government's measures on food safety has increasing positive relationship with observing people's satisfaction level of the central government from the lowest to the highest. The study found the coefficients of X8 fluctuated in (-1,-2) from 0.1 to 0.7 quantile level, while decreased from 0.8 quantile level. It means the satisfaction with the central government's policies on the twelve years primary education reform has the increasing positive relationship when observing people's satisfaction with the central government from the lowest to the highest level.

The study also calculated the Percentage Correclty Classified (PCC) for the Bayesian quantile regression model. The histogram of predicted probabilities for the sample was shown on Figure 4. The PCC equals 0.736, which means 73.6% of the sample was correctly classified.

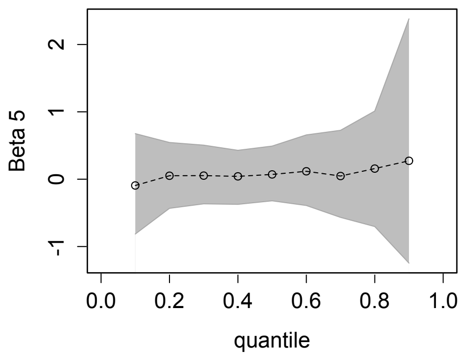


(a) X2

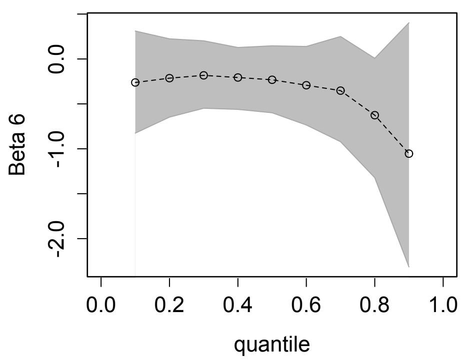
(b) X3

(c) X4

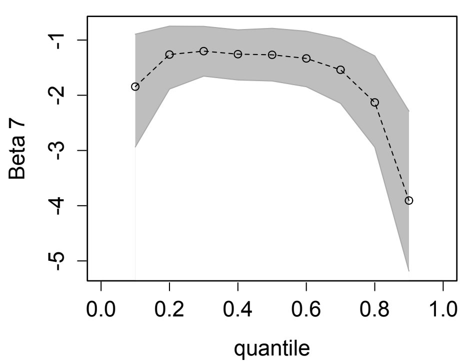
Figure 1: The Regression Parameters:X2 to X4



(a) X5



(b) X6



(c) X7

Figure 2: The Regression Parameters:X5 to X7

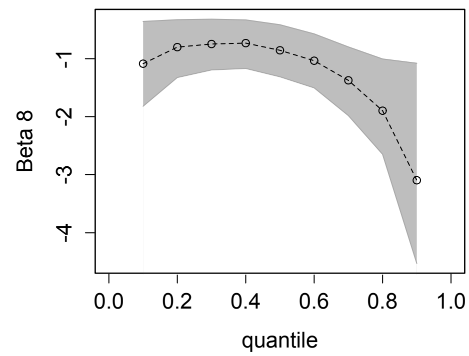


Figure 3: The Regression Parameters:X8

	Bayes Estimate	lower	upper
(Intercept)	8.4181	6.666	10.399
X2	-1.8708	-2.347	-1.415
X3	-0.0655	-0.469	0.334
X4	-0.2325	-0.602	0.153
X5	0.0619	-0.370	0.524
X6	-0.2021	-0.582	0.179
X7	-1.2231	-1.727	-0.745
X8	-0.7587	-1.228	-0.307

Table 2: Bayesian Quantile Regression Coefficients at 0.25 Quantile Level.

	Bayes Estimate	lower	upper
(Intercept)	10.3979	8.653	12.221
X2	-1.9283	-2.385	-1.409
X3	0.0172	-0.345	0.387
X4	-0.1423	-0.518	0.238
X5	0.0707	-0.321	0.491
X6	-0.2314	-0.599	0.146
X7	-1.2675	-1.741	-0.788
X8	-0.8553	-1.311	-0.415

Table 3: Bayesian Quantile Regression Coefficients at 0.5 Quantile Level.

	Bayes Estimate	lower	upper
(Intercept)	18.3442	14.247	21.694
X2	-3.1090	-3.941	-2.238
X3	0.2113	-0.366	0.861
X4	-0.0222	-0.672	0.590
X5	0.1770	-0.537	0.845
X6	-0.5188	-1.153	0.069
X7	-1.7374	-2.374	-1.049
X8	-1.5389	-2.262	-0.860

Table 4: Bayesian Quantile Regression Coefficients at 0.75 Quantile Level.

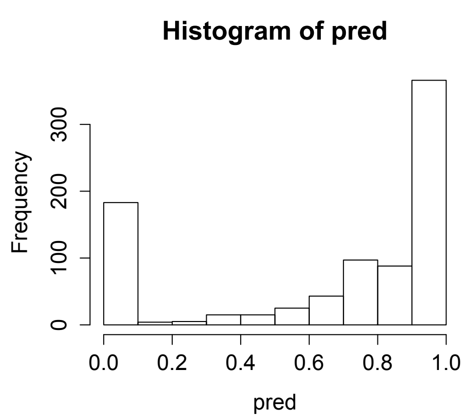


Figure 4: Histogram of Predicted Probabilities

4 Empirical Analysis of Rough Set Classification

4.1 Rough Set Rule-based Classification

There are many ways to represent the knowledge. Rule-based classifiers are one of the widely-used knowledge representation. The general structure of the rule is IF...THEN.... The IF part represents the predecessor and the THEN part represents the successor. The rule (denoted as Λ)of the Rough Sets Theory (RST) is named as the decision rule. The decision rule is expressed by IF (φ), THEN (d=ν), where $\varphi \in$C(A,V$_a$).C(A,V$_a$) is a set of pairs of conditional attributes A and their corresponding values V_a that are linked by the proportional $\wedge$ (conjunction), $\vee$ (disjunction), and $\neq$ (negation). The decision rule in Λ is true if and only if, $\|\varphi_\Lambda \| \subseteq \|d=\nu_\Lambda\|$ where in the case, $\|.\|$ is the set of objects matching the decision rule. [12]

The study applied the RoughSets package of R to make the roughsets rule-based classification. The RoughSets package creates the smallest number of decision rules that are suitable for characterizing all given cases. The package also applied the algorithms which induce all possible decision rules, while the last case that meet the pre-defined requirements.

The RoughSets package used the CN2 and LEM2 (Learning from Example Module, version 2) algorithms to induce rules from the data. It creates a minimal set of rules which is based on computing a signal local for including each concept from the decision rule.

4.2 Rough Set Rule-based Classification Results

The study utilized the CN2 and LEM2 algorithms to make the Roughsets rule-based classification. The study used 60% of the data (505 survey results) as the training data and 40% (336 survey results) as the test data.

(i) CN2 algorithm: The CN2 algorithm determines the most suitable complex to substitute for the original best complex. The computing relates to the finding of the suitable set which meets the requirements of its selectors and the probability distribution P=(p_1...p$_n$) of examples of E' among classes, whereas n is the number of classes of the training data. CN2 algorithm then used the information- theoretic entropy measure[13]

$$Entropy = -\sum_i p_i log_2(p_i) \qquad (5)$$

The study applied the CN2 algorithm to make the classification and obtained the 53 rules of the data, and the first 10 rules are as follows.

1. IF X8 is [-Inf,1] and X2 is [-Inf,2] and X3 is [-Inf,1] THEN Y is 1; (supportSize=92;laplace=0.989)
2. IF X8 is [-Inf,1] and X3 is (1,2] and X7 is [-Inf,1] and X5 is [-Inf,1] THEN Y is 1; (supportSize=62; laplace=0.9845)
3. IF X2 is [-Inf,2] and X5 is (1, Inf] and X4 is (2, Inf] THEN Y is 1; (supportSize=17; laplace=0.9473)
4. IF X7 is [-Inf,1] and X4 is [-Inf,1] and X8 is (1,2] and X6 is (1, Inf] THEN Y is 1; (supportSize=16; laplace=0.9444)
5. IF X7 is [-Inf,1] and X3 is (1,2] and X6 is [-Inf,1] and X5 is [-Inf,1] THEN Y is 1; (supportSize=21; laplace=0.9130)
6. IF X7 is [-Inf,1] and X4 is [-Inf,1] and X5 is (1, Inf] and X8 is [-Inf,1] THEN Y is 1; (supportSize=20; laplace=0.9090)
7. IF X8 is (2, Inf] and X6 is (1, Inf] and X2 is (2, Inf] THEN is 0; (supportSize=14; laplace=0.9375)
8. IF X7 is [-Inf,1] and X2 is [-Inf,2] and X3 is (2, Inf] and X6 is (1, Inf] THEN Y is 1; (supportSize=7; laplace=0.888)
9. IF X7 is [-Inf,1] and X4 is [-Inf,1] and X8 is (1,2] THEN Y is 1; (supportSize=35; laplace=0.8648)
10. IF h7 is [-Inf,1] and h8 is (2, Inf] and h2 is (2, Inf] THEN Y is 1; (supportSize=5; laplace=0.857)

The RoughSets package in R calculated the rate of correct prediction of Y is 76.48%.

(ii) LEM2 algorithm: Let B be a nonempty lower or upper approximation of a notion exemplified by a decision-value pair (d,w). Set B counts on a set T of attribute-value pairs t if and only if

$$\oslash \neq [T] = \bigcap_{t \in T} [t] \subseteq B \tag{6}$$

Set T is a minimal complex of B if and only if B counts on T and no proper subset T' of T exists such that B depends on T'. Let λ be a non-empty collection of non-empty sets of attribute-value pairs. Then λ is a local covering if and only if the following conditions are satisfied[14]:

(1) Each member of T of λ is a minimal complex of B.
(2) $\bigcap_{t \in T}[T]$=B, and
(3) λ is minimal, and λ has the smallest number of members. The usage of LEM2 can opt to consider the attribute priorities or not.
The study applied the LEM2 algorithm to make the classification and obtained the first 10 of all 140 rules:

1. IF X4 is (2, Inf] and X5 is (1, Inf] and X2 is [-Inf,2] THEN is Y=1; (supportSize=21; laplace=0.956)
2. IF X5 is (1, Inf] and h7 is [-Inf,1] and X4 is (2, Inf] and X6 is [-Inf,1] and X3 is (2, Inf] THEN Y is 1; (supportSize=2; laplace=0.75)
3. IF X2 is [-Inf,2] and X6 is [-Inf,1] and X5 is [-Inf,1] and X4 is (1,2] THEN Y is 1; (supportSize=25; laplace=0.925)
4. IF X5 is (1, Inf] and X4 is (2, Inf] and X2 is [-Inf,2] THEN Y is 1; (supportSize=21; laplace=0.956)
5. IF X2 is [-Inf,2] and X8 is [-Inf,1] and X3 is [-Inf,1] THEN Y is 1; (supportSize=92; laplace=0.989)
6. IF X8 is [-Inf,1] and X3 is [-Inf,1] and X5 is (1, Inf] THEN Y is 1; (supportSize=20; laplace=0.909)
7. IF X8 is [-Inf,1] and X2 is [-Inf,2] and X5 is [-Inf,1] THEN Y is 1; (supportSize=154; laplace=0.961)
8. IF X2 is [-Inf,2] and X5 is [-Inf,1] and X8 is [-Inf,1] THEN Y is 1; (supportSize=154; laplace=0.961)
9. IF X6 is (1, Inf] and X4 is [-Inf,1] and X3 is [-Inf,1] and X7 is (1, Inf] THEN Y is 1; (supportSize=6; laplace=0.875)
10. IF X3 is (2, Inf] and X2 is [-Inf,2] and X8 is (1,2] and X7 is [-Inf,1] THEN Y is 1; (supportSize=9; laplace=0.909)

The RoughSets package in R calculated the rate of correct prediction of Y is 78.86%.

5 Discussion

The study compared with the classification results of the Bayesian quantile regression and two algorithms of rough set classification. After cross verification of the classification results generated by the three models, we used the pROC package of R language[17] to plot the ROC curves, which were shown from figure 5 to figure 7. Bradly(1997) elaborates that the larger the area above the reference line and the curve, the more accurate the classification of the model[16]. The study also calculated the ROC analysis output result, whereas Sensitivity (Sen) means the percentage occupied the by the number with a forecast result of 1 to the number with a real number of 1, and the Specificity (Spe) means the percentage occupied by the number with a forecast number of 0 to the value with a real number of 0. Hand and Till state that Gini Index = 2*AUC-1. For all index values, the larger the better [15]. The index values of all three models were presented in the Table 5.

As seen in Table 5, the Bayesian quantile regression had the largest value for all indices. Therefore it had a better performance on alarming and detection competence.

As for the information induced from the Bayesian quantile regression, the satisfaction rate on the uncorrupted performance of the central government(X2) has the decreasing positive relationship with people who have the lower satisfaction level with the central government. The study also found the satisfaction level of governmental measures on food safety (X7) and the satisfaction level of twelve years primary education reform (X8) had the increasing positive relationship with people's satisfaction level with the central government.

Besides, according to the rough set classification results of CN2 algorithm, the classification rule with the largest support size indicated that the lowest satisfaction rate on the twelve years primary education reform (X8), the lowest satisfaction rate on the uncorrupted performance of the central government (X2) and the lowest satisfaction rate on Taiwan's current economic condition (X3) would result in the lower satisfaction rate on the central government. The study also obtained the classification rule from the LEM2 algorithm. The classification rule with the largest support size (Rule No.7 and No.8) pointed out that the lowest satisfaction rate with the uncorrupted performance of the central government (X2), the lowest satisfaction rate on the respondents' household economic condition (X5) and the lowest satisfaction rate on the twelve years primary education reform (X8) would result in the lower satisfaction rate on the central government.

Model	Sen	Spe	Auc	Gini
Bayesian Quantile Regression	0.998	1.000	0.9992	0.9984
Roughsets:CN2 algorithm	0.838	0.485	0.662	0.324
Roughsets:LEM2 algorithm	0.947	0.185	0.6485	0.297

Table 5: The Analysis Output Result of ROC Curve.

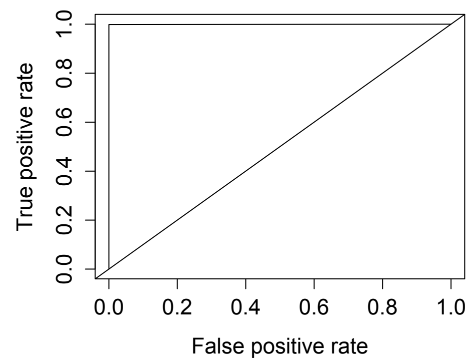


Figure 5: The ROC Curve of the Classification Results:Bayesian Quantile Regression

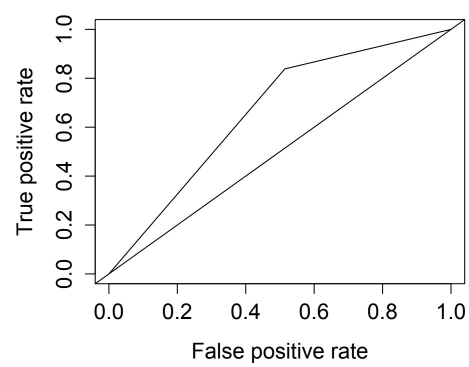


Figure 6: The ROC Curve of the Classification Results:CN2 algorithm of Roughsets classification

6 Conclusion

The study proposes to assess the Taiwanese people's satisfaction level with the central government by the Bayesian Quantile Regression and two algorithms of rough set classification in order to explore the factors which affect the people's satisfaction with

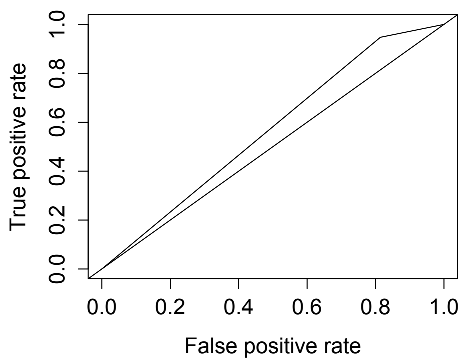


Figure 7: The ROC Curve of the Classification Results:LEM2 algorithm of Roughsets classification

the central government. The study finds that the Bayesian quantile regression classification has a better classification performance, while the LEM2 algorithm of rough set classification had the largest correct prediction of Y.

The major contribution of this paper is to find the major factors which have the positive relationship with the people's satisfaction level with the central government. These factors include:(i)satisfaction with the uncorrupted performance of the central government; (ii) the respondents' self evaluation of their household's economic condition; (iii)satisfaction with the central government's measures on food safety. (iv) satisfaction with the twelve years primary education reform.

References:

[1] G.G. Van Ryzin,An experimental test of the expectancy-disconfirmation theory of citizen satisfaction, *Journal of Policy Analysis and Management*, 32–3, 2013, pp. 597–614.

[2] G.G. Van Ryzin,Service quality, administrative process, and citizens evaluation of local government in the US, *Public Management Review*, 17–3, 2015, pp. 425–442.

[3] S.Grimmelikhuijsen, G.Porumbescu, B. Hong, and T.Im, The effect of transparency on trust in government: A cross-national comparative experiment, *Public Administration Review* 73–4, 2013, pp. 575–586.

[4] S. L. Kastner,Is the Taiwan Strait Still a Flash Point? Rethinking the Prospects for Armed Conflict between China and Taiwan, *International Security* 40–3, 2016, pp. 54–92.

[5] C. Huang,Taiwans Election and Democratization Study, 2012-2016(III): the Survey of the Nine-in-One Local Elections in 2014: 2012-2016(III): the Survey of the Nine-in-One Local Elections in 2014 (TEDS2014) (E10137), Available from Survey Research Data Archive, Academia Sinica, 2016.

[6] R. Koenker, and G.Bassett Jr, Regression Quantiles, *Econometrica: journal of the Econometric Society*, 1978, pp. 33–50.

[7] S.Frydenberg, O.H. Reiakvam, S.B. Thyness, and S.Westgaard,Hedge FundsRisk Exposure in Different Quantiles and Market Sentiments, *Econometrica: journal of the Econometric Society*, 22–3, 2013, pp. 107–134.

[8] R. Alhamzawi, K. Yu, and D. F. Benoit, Bayesian Adaptive Lasso Quantile Regression, *Statistical Modelling*, 12–3, 2012, pp. 279–297.

[9] M.Harding, C. Lamarche, Carlos, A quantile regression approach for estimating panel data models using instrumental variables,*Economics Letters*, 104–3, 2009, pp. 133–135.

[10] V. L.Miguéis, D.F. Benoit, and D. Van den Poel,Enhanced decision support in credit scoring using Bayesian binary quantile regression,*Journal of the Operational Research Society*, 64–9, 2012, pp. 1374–1383.

[11] G. Kordas,Smoothed binary regression quantiles,*Journal of Applied Econometrics*, 21–3, 2006, pp. 387–407.

[12] L.S.Riza, A. Janusz, C.Bergmeir, C.Cornelis, F. Herrera, D.Śle, J. Benítez and others, Implementing algorithms of rough set theory and fuzzy rough set theory in the R package roughsets,*Information Sciences*, 287, 2014,pp. 68-89.

[13] P. Clark, R. Boswell, Rule induction with CN2: Some recent improvements, *Machine learningEWSL-91*, 1991, pp. 151–163.

[14] J.W.Grzymala-Busse, A new version of the rule induction system LERS, *Fundamenta Informaticae*, 31–1, 1997, pp. 27–39.

[15] D.J.Hand, R.J.Till, A simple generalisation of the area under the ROC curve for multiple class classification problems, *Machine learning*, 45–2, 2001, pp.171-186.

[16] A.P.Bradly, The use of the area under the ROC curve in the evaluation of machine learning algorithms, *Pattern recognition*, 30–7, 1997, pp. 1145–1159.

[17] X. Robin and N. Turck and A. Hainard and N. Tiberti and F. Lisacek and J.-C. Sanchez and M. Muller, pROC: an open-source package for R and S+ to analyze and compare ROC curves, *BMC Bioinformatics*, 12, 2011, p.77.