# A Comparison of Classification Models for Life Insurance Lapse Risk

**Lim Jin Xong, Ho Ming Kang**

***Abstract:** Insurer usually incurs expenses such as policy issuance cost, commission and administrative costs after the launching of an insurance product. When a policyholder decided to lapse a policy, the insurer have to seek for alternative to cover all the losses, which might liquidate high-yielding investments in order to satisfy their requests for the cash value or surrender value. Therefore, it is crucial to understand and develop a classification model to determine the surrender or lapse risk. In this study, four classification models such as logistic regression, k-Nearest Neighbor, Neural Network (NN) and Support Vector Machines (SVM) are used to model the life insurance lapse risk, which is the risk thatinvolving the termination of policies by the policyholders. Classification performance criterions such as prediction accuracy and area under the Receiver Operating Curve (ROC) are used to compare the performance between the models. The results showed that SVMwas outperformed than NN, logistic regression and k- Nearest Neighbor.*

***Keywords:** Logistic Regression, k-Nearest Neighbor, Neural Network, Support Vector Machines, Classification, Lapse Risk.*

## I. INTRODUCTION

Understanding the dynamic of the lapse (or surrender) rates is a crucial point for insurer. Initially, the term “lapse” meant termination of an insurance policy and loss of coverage because of the policyholder failed to pay the premium and “surrender” is used in cases where a cash surrender value is paid out to the policyholder. However, today’s insurance policies allow the policyholders to choose among many options that can significantly affect the extent of insurer’s liabilities. For example, a policyholder able to receive a surrender value or choose to discontinue premium payments when they surrender their insurance policy which will change the status of the insurance policy to be lapse.According to the Life Insurance Association of Malaysia (2017), the insurance industry registered a healthy growth in the first half of 2017 with 4.7% growth in new business weighted premium and this reflects the continued increase in awareness among the Malaysian on the importance of insurance protection. However, the life insurance industry paid out higher claims compared with the same period of the year 2016 as there is a 3.2% increase in the benefit payment for death, disability, medical, bonuses, and others. Insurance company might face losses in one day as the amount of the benefit payout is getting increase from year to year mainly due to the increasing trend for total health expenditure in Malaysia. For instance, policyholders who have adverse health or other insurability problems tend not to lapse their policies, and this makes the insurer to have more claims than expected.

**Revised Manuscript Received on January 19, 2019.**

**Lim Jin Xong,** School of Mathematics, Actuarial and Quantitative Studies, Asia Pacific University of Technology and Innovation, Malaysia.

**Ho Ming Kang,** School of Mathematics, Actuarial and Quantitative Studies, Asia Pacific University of Technology and Innovation, Malaysia.

A huge number of early surrender or insurance policy lapse causes a liquidity threat to the insurer who is subjected to interest rate risk as an increase in the interest rate will decrease the equilibrium premiums. Therefore, there is a greater chance that a new premium will be lower with the same coverage of sum insured. And this might have an impact to the existing policyholders when making a decision on surrender their insurance policy. So they are able to benefit from lower premium that available in the market or a higher yields after the rise of the interest rate (Milhaud et al, 2011).

The decision to lapse a life insurance policy can have far-reaching effects on the issuing insurance company. A high lapse rate might be having a negative impact on insurer’s reputation which might result in even more policyholder lapsing or choose to surrender the insurance policy, as well as harm new business (Eling & Kochanski, 2013). From the perspective of an insurer, the excessive policy with lapse status adversely give impacts to costs, mortality experience, investment returns, each of which will negatively affect the financial stability and well-being of the insurance company (Fier & Liebenberg, 2013). Besides, lapses also will diminish the effectiveness of risk pooling and lapse risk accounts for half of the capital requirements in the life insurance underwriting. Insurer should have a thorough understanding of lapse dynamics in order to define the reasonable capital standards. From the customers’ perspective, lapses can be one of the main indicators to review the life insurance company’s product and service quality. Companies with above average lapse rates might provide less services than competitors to the customer or offer more expensive products. Customers are able to utilize this indicator as an additional information before the decision is made (Kiesenbauer, 2012).

To reducethelapse risk and losses that might be occurred in the future, insurershave to take initiative to develop an effective classification model in order to classify the life insured into two distinct groups, which is lapsed or not lapsed.Support vector machines (SVM) was found to provide a good result, with high accuracy in classification (Min, 2005; Lee, 2007; Crook et. al., 2007; Huang et. al., 2008; Bellotti & Crook, 2009; Trustorff et. al., 2010; Memić, 2015; Mohammadi and Zangeneh, 2016). According to Stąpor (2016), SVM are successful in building the credit scoring model, however, support vector machines with non-linear kernel does not give the significant improvement over the simpler models like discriminant analysis. However, Karaa and Krichene’s (2012) confirmed the superiority of the multilayer Neural Network (NN) to SVM.


*Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication*

They commented that NN is the best prediction model compared to SVM in classification rates and reduction of Type I Error. While the radial basis function scoring model (neural network) was superior in identifying the default client, however, the logistic regression performed slightly better than the radial basis function model in terms of the overall accuracy rate (Bacconi, 2002). Furthermore, Nazari and Alidadi (2013) discussed the customers' credit risk measurement by using NN and able to estimate an average of 80% correct in classifying bad and good customers. Lee and Choi (2013) suggested that the prediction of bankruptcy using NN was better than multivariate discriminant analysis (Prediction accuracy of 81.43% and 74.82% for NN and multivariate discriminant analysis respectively) and identified that NNworked better in capturing the nonlinear pattern between the independent variables and the dependent variable. Brown and Mues (2012) compared the several classification techniques such as NN, logistic regression, linear and k- nearest neighbor together with least square SVM. They found out logistic regression was reasonably competitive with the more complex techniques and even when the samples became much more imbalanced whereas thek- nearest neighbor performed significantly worse than the SVM. In addition, k- nearest neighbor classification is one of the most fundamental and simple classification methods (Parvin et. al., 2010). According to the Babu & Satish (2013), k- nearest neighbor can be describe as the benefit from reducing the cost of analysis. For instance, in the study of Abdelmoula(2015) he has found that the k- nearest neighbor classifier algorithm was conducted with 88.63% classification accuracy in tackling the question of default prediction of short term loans.

Apart from that, there are some findings show that the key determinants of the lapse are very similar across all product categories. The study of Kiesenbauer (2012) showedthat the main lapse determinants are including distributional channel, insured's age, and participation rate spread which applies mainly to saving products. Milhaud et al. (2011) applied logistic regression, classification and Regression Tree model and found out that senior insurance policyholder who have a premium frequency in annual term, tend to surrender more than others, while the gender of the policyholder does not seem to be significant in this case study. In the study that carried out by Cerchiara et al. (2017), they proposed a Generalized Linear Models which typically used in general insurance to investigate the risk factors that underlying the underwriting risk of life insurance. Their study shows the product class, calendar year of exposure, policy duration, and policyholder age are the determinants of lapse risk.

To develop a classification model and determine the lapse or surrender risk, independent variables such as Premium Frequency, Entry Age, Policy Term, Sum Assured and Gender are included in the analysis. The data collected is updated with policy status as at 30 June 2017. The targeted population is the insurance policyholder who aged between 21 to 65 years old as they are the active generation and have a stable income for the daily consumption in their life (Assumed the retirement age is at 65 years old). A total of 800 sample data are obtained from the Malaysia based insurance company with the characteristics of the insurance policy such as Policy Status, Sum Insured, Premium Frequency, Entry Age, Policy Term and Gender.

Definition of the variables used in this study:-

**a. POLICY STATUS**(*PStatus*)

Generally, there are four types of the policy status for the insurance policy including, Inforce, Lapse, Surrender and Death. A life insurance policy will lapse when premium payments are missed and cash surrender value of the particular insurance policy is exhausted on a life insurance policy. An inforce policy status means that the policy is paid up and there is an active protection to the policyholder. The sample data collected consists of the insurance policy status "Inforce", or "Lapsed or Surrender".

$$\text{PStatus}=\begin{cases} 0, & \text{Not Lapsed} \\ 1, & \text{Lapsed} \end{cases}$$

**b. PREMIUM FREQUENCY**(*PFrq*)

There are some options of premium frequency that allow insurance policyholder to choose in order to pay the premium. Insured can choose to pay the insurance premium in various term period such as "monthly", "quarterly", "semi-annually" or "annually".

$$\text{PFrq}=\begin{cases} 1, & \text{Monthly} \\ 2, & \text{Quarterly} \\ 3, & \text{Semi-annually} \\ 4, & \text{Annually} \end{cases}$$

**c. ENTRY AGE**(*EAge*)

The age of the policyholder which he/she purchased the insurance and eligible to be part of contributor in the insurance fund pool and receive benefit of payment whenever the claims is made.

**d. POLICY TERM**(*PTerm*)

Policy Term indicates the period of the coverage and protection which provided by the insurance policy and insurer.

**e. SUM ASSURED**(*SumA*)

The amount that the insurer that agree to pay on the occurrence of an event in the future.

**f. GENDER**(*Gder*)

Gender of the insurance policyholder is either "Male" or "Female".

$$\text{Gder}=\begin{cases} 0, & \text{Male} \\ 1, & \text{Female} \end{cases}$$

**Classification Models**

**Logistic Regression (LR)**

Logistic regression is the most common method in classifying the participant into two distinct groups. This study will be targeting on the binary response of whether an insurance policy status become "Lapsed" $(Y = 1)$ or "Not Lapsed" $(Y = 0)$ in status. The logistic regression model is written as:

$$logit(p_i) \equiv ln\left(\frac{p_i}{1-p_i}\right) = b_0 + \sum_{i=1}^{n} b_i x_i + \varepsilon$$

where

$\boldsymbol{p_i}$ is the probability of the insurance policy status become "Lapse".



Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$\boldsymbol{x_i}$ , $i = 1,2,3,4,5$ is representing the characteristics of the particular insurance policy such as policy status, sum insured, premium frequency, entry age, policy term and gender.

$\boldsymbol{b_0}$is the intercept of the equation.

$\boldsymbol{b_i}$ where $i = 1,2,3,4,5$ are the coefficient that associated with the corresponding predictor $x_i$.

$\boldsymbol{\varepsilon}$is the error term.

The odds express the likelihood of an event occurring relative to the likelihood of an event not occurring. There are some assumptions apply to the logistic regression such as:

1. The dependent variable must be dichotomous in nature.
2. Assume linearity between independent variables and log odds. Explanatory variables should linearity related to the log odds.
3. No multicollinearity among independent variables, which occur when two or more independent variables that highly correlated with each other.

**_k_- Nearest Neighbor (kNN)**

The $k$- nearest neighbor classifies a particular data by taking the majority vote of its $k$ most similar data. This classifier is memory- based and it requires no model to be fit. The most common choice, a Euclidean distance metric is used to measure the similarity between each training example and a new example (Babu & Satish, 2013). For each new observation, this method explores the pattern space for $k$-nearest neighbor that are closest to new observation. Consider the test point $X_{\text{(test)}}$ and the point in the training set $Y$ where $Y = \{y_1, y_2, \dots, y_n\}$, the $k$ training points from the training set $Y$ are found that has closest distance to $X_{\text{(test)}}$. Then the test point is classified as using the majority vote among the $k$ neighbors where test point is assigned to the class which it's most $k$- nearest neighbor belong to the class. Euclidean distance is given by:-

$$d(\mathbf{x}_i, \mathbf{x}_j) = \|\mathbf{x}_i - \mathbf{x}_j\| = \left[(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{x}_i - \mathbf{x}_j)\right]^{\frac{1}{2}}$$

The $k$ number is an important parameter. The result might be sensitive to the noisy data when there is a too small $k$ number, whereas, if the $k$ number is too high, the $k$ nearest neighbor can include instances from different classes. Therefore, with the cross- validation approach, where algorithm is repeated for different odd number of $k$, the best value of $k$ can be estimated. Starting with $k$ = 1, a test is used to estimate the error rate of the classifier. The $k$ number with the minimum error rate is selected, the larger the training samples is, and the large of the value $k$ will be (Abdelmoula, 2015).

**Neural Network (NN)**

Neural networks isa mathematical representation modelled on the functionality of the human brain. The added benefit of a neural network is its flexibility in modelling virtually any non-linear association between input variables and target variable. According to Nazari & Alidadi (2013), it plays an increasingly important role in financial applications for such tasks as pattern recognition, classification, and time series forecasting.There are three main components which are the input data layer, the hidden layer(s) and the output measure(s) layer. It is also known as multilayer perceptron (MLP). An "input" layer's units is linked to a layer of "hidden" units, which is connected to "output" layer's units. Stimuli are transmitted by the neurons through connections. Every connection is associated to a weight which it multiplies itself upon receiving a stimulus. This is best seen graphically. Next, each neuron contributes for the activation function to determine the output stimulus. The hidden(s) layer contain two processes, transformation function and weighted summation functions (typically used in a feedforward/ back propagation neural network model). Both functions relate the value from the input data to the output measures.In the data normalization process, the input data are normalized between 0 and 1. To find out the best neural network, multilayer perceptron (MLP) neural network is trained using Resilient Backpropagation As activation function, the sigmoid function and hyperbolic tangent were chosen in this study.There are input signal, $X_i, i = 1,2,3,4,5$ and corresponding weights $\boldsymbol{w_i}, i = 1,2,3,4,5$ and limit being $k$. The level of activity in this model is given by:

$$a = \sum_{i=1}^{5} W_{ij} X_j$$

Meanwhile the output layer consists of two neuron that produce the binary output, $y$ like,

$$y = \begin{cases} 1 \ or \ \text{"Lapsed"} & , a \geq k \\ 0 \ or \ \text{"Not Lapsed"} & , a < k \end{cases}$$

**Support Vector Machines(SVM)**

Developing and implementing of algorithm for support vector machines are currently of great interest to theoretical researchers in machines learning. Support vector machines are available in both linear and nonlinear, which involves optimization of convex loss function under given constraints and so are unaffected by the problems of local minima (Izenman, 2013). Besides, it has become an increasingly popular non- parametric methodology for developing classification models (Nikis et al., 2011). In SVM, it separates a binary classified data by a hyperlane such that the margin width between the hyperlane and the examples is maximized. By maximizing the margin width, the complexity of the model can be reduced.From the training data set $\{(\boldsymbol{x}_k, y_k)\}_{k=1}^{N}$ with the dependent binary variable is labelled as

$$y_k = \begin{cases} 1 & , \quad \text{"Lapsed"} \\ 0 & , \text{"Not Lapsed"} \end{cases}$$

and the input data $\boldsymbol{x}_k \in \mathbb{R}^n$ which represents the independent variables of this study. According to Trustorff et al. (2010), a separating hyperplane, $w^T\varphi(\boldsymbol{x}) + \boldsymbol{b} = 0$ satisfies for a non-linear case

$$\boldsymbol{w}^T\varphi(\boldsymbol{x}_k) + \boldsymbol{b} \geq +1 \quad \text{if} y_k = 1 \ \text{ and}$$
$$\boldsymbol{w}^T\varphi(\boldsymbol{x}_k) + \boldsymbol{b} \leq -1 \quad \text{if} y_k = -1$$

With that, it is equivalent to

$$y_k[\boldsymbol{w}^T\varphi(\boldsymbol{x}_k) + \boldsymbol{b}] \geq 1$$

For a non- separable case, a stack variable, $\xi$ will be introduced to tolerate the misclassification in the set of margin constraints. When $\xi > 1$, then there is occurring the misclassification, the classification constraints is followed as:-

$$y_k[\boldsymbol{w}^T\varphi(\boldsymbol{x}_k) + \boldsymbol{b}] \geq 1 - \xi_k$$

The objective function for the optimization problem is considered as

$$\min \mathcal{J}_p\,(\boldsymbol{w},\xi) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + \gamma\sum_{k=1}^{N}\xi_k$$

subject to

$$y_k[\boldsymbol{w}^T\varphi(\boldsymbol{x}_k) + \boldsymbol{b}] \geq 1 - \xi_k,\;\; \xi_k \geq 0, k = 1, \dots, n$$

where $\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}$separating the hyperplane and maximize the margin between the two classes, while $\gamma\sum_{k=1}^{N}\xi_k$ minimize the error of the classification and the parameter $\gamma$ controls how much error- tolerant the machine is. Next, the solution function $f(x)$ with the Lagranga multiplier, α and choosing a particular kernel function, $K(\boldsymbol{x},\boldsymbol{x'})$ is as

$$f(x) = sgn\left(\sum_{k=1}^{N}\alpha_k y_k K(\boldsymbol{x},\boldsymbol{x'}) + b\right)$$

The kernel function $K(\boldsymbol{x},\boldsymbol{x'}) = \varphi^T(\boldsymbol{x})\varphi(\boldsymbol{x'})$ returns a real valued number characterizing the similarity of a particular test data $x$ and each training data point. Four kernel functions are considered: linear, polynomial (with degree d), radial basis function (with parameter $\sigma$) and sigmoid function, where $d, r \in \mathbb{N}$ and $\gamma \in \mathbb{R}^+$.

**Model Performance Assessment Criteria**

**Classification Accuracy**

From Table 1, TP, FP, FN and TN are True Positive, False Positive, False Negative and True Negative respectively. Classification accuracy or score of hits is the proportion of the cases that classified correctly to the total data available.

**Table 1 Confusion Matrix for Classification Performance.**

| CLASSIFICATION METHOD (LR, *k*- NN, NN & SVM) | | PREDICTED | | TOTAL |
|---|---|---|---|---|
| | | LAPSED | NOT LAPSED | |
| ACTUAL | LAPSED | TP | FN | TP + FN |
| | NOT LAPSED | FP | TN | FP + TN |
| TOTAL | | TP + FP | FN + TN | N |

The accuracy of the classification algorithm is then considered the ratio of True Positive ($TP$) and True Negative ($TN$) on the total data set. By referring to the Table 1, the classification accuracy can be defined as:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

The classification accuracy ignores the different cost of the both Type I error and Type II error where the insurance policies' status are actually "lapse" in for which the model predict "not lapse" or vice- versa.

**Receiver Operating Characteristic (Roc) Curve and Area Under Curve (AUC)**

One of the method of evaluating the performance of the classification algorithm is to note its True Positive (TP) rate and False Positive (FP) rate. Both True Positive rate and False Positive rate are the percentage of target samples that are correctly classified as target samples and non-target samples that are incorrectly classified as the target samples respectively (Woods & Bowyer, 1997) When it comes to performance, the AUC is a good measurement criterion for classifiers (Ghodselahi & Amirmadhi , 2011).For a better classifier, the ROC curve will go straight to the $y$-axis along $x$-axis. Consequently for referring to Figure 1, the greater Area Under Curve (AUC) which lies below the ROC curve, the greater the power of a model in prediction of the particular classifier.

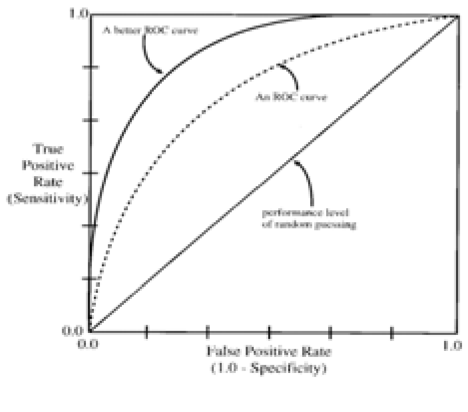



**Figure 1: Receiver Operating Characteristic Curves (Woods & Bowyer, 1997).**

## II. RESULTS

**A. Classification Performance Evaluation**

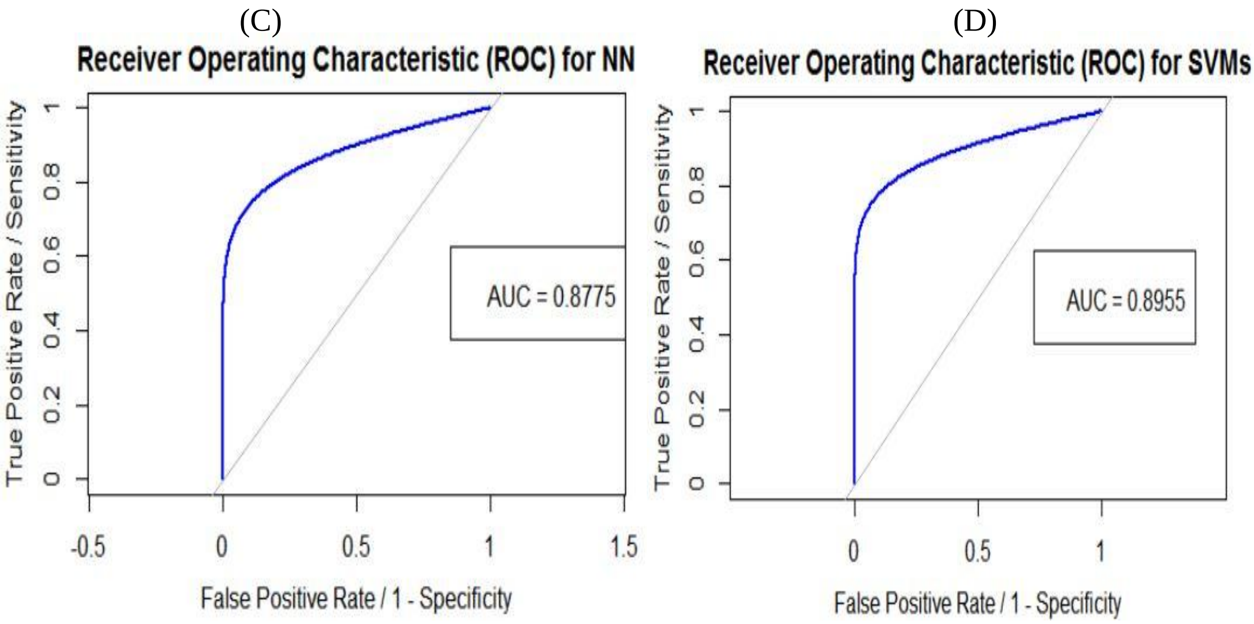
Figure 2 shows the performance of four classification models. The hit percentages in training are very similar in both*k*- nearest neighbor (86.25%) and neural network (86.09%) meanwhile logistic regression model shows the lowest (81.56%). The classification model with highest accuracy in training is SVM model. However in the testing data, neural network shows the highest accuracy (85%) as compared to other models.

Area under the Receiver Operating Characteristic (ROC) curve is applied to evaluate the classification performance of each model. Since the area under curve is a portion of the area of unit square, its value is always within the range [0, 1]. The larger the value of the area under curve, the more accurate of the classification model. For instance, area under curve will equal to 0.5 while an area of 1 reflects a perfect model. Hence, the receiver operating characteristic curve that has a larger area under curve is preferred. In this study, all the models have a good discriminating capacity as area under curve value were greater than 70% (Cholongitas et al., 2006). Based on the Figure 3, the area under curve criterion for logistic regression, *k*- nearest neighbor, neural network and SVM are 81.09%, 82.28%, 87.75% and 89.55% respectively.It can be concluded that the SVM outperformed than other classification models by having the largest area under curve value. On the other hand, although the ROC for logistic regression exceeds 80%, but the performance is still the lowest as compared withother classification models.


Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication


**Figure 2:Classification Result of Four Classification Techniques**

Figure 3Receiver Operating Characteristic for (A) logistic regression, (B) k-nearest neighbour, (C) Neural Network, (D) SVM




## III. CONCLUSION

Lapse rate modelling has been active field of research in previous years (Eling & Kochanski, 2013). This study presented an empirical comparison of four classification algorithms such as logistic regression, k- nearest neighbor, neural network and support vector machines for constructing life insurance lapse risk assessment. This study is using a large sample of 800 observations which obtained from an insurance company based in Malaysia. The best classification method will be defined based on the correct or good classification results together with other criterions which will be discussed in this study. The classification power of these classification algorithms was assessed using prediction accuracy on training and testing, and area under ROC curve. Based on the results above, both SVM and NNproduced high prediction accuracy in both training and testing, and with the largest ROC value compared to logistic regression and *k*-nearest neighbour. Underwriting is the process of the risk assessment evaluation and selection based on the identification of factors which contribute to “good” and “bad” risk. Insurer should apply the proposed classification models in order to classify their client (insured) accurately. Apart from classifying the insured appropriately, the resultsalso helpthe life insurance underwriters to charge the applicant at a reasonable premium with


*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*

respect to the lapse risk level. For instance, the higher the lapse risk level of an insured, the higher the insurance premium.

## REFERENCES


1. Abdelmoula, A. (2015). Bank Credit Risk Analysis with K-Nearest Neighbor Classifier: Case Of Tunisian Banks. *Accounting and Management Information Systems*, 14(1), pp.79 - 106.
2. Babu, R. and Satish, A. (2013). Improved of K-Nearest Neighbor Techniques in Credit Scoring. *International Journal For Development of Computer Science & Technology*, 1(2), pp.1 - 4.
3. Bacconi, E. (2007). Credit Risk Analysis Applying Logistic Regression, Neural Networks and Genetic Algorithms Models. *Research Gate*, pp.1-48.
4. Bellotti, T. and Crook, J. (2009). Support vector machines for credit scoring and discovery of significant features. *Expert Systems with Applications*, 36(2), pp.3302-3308.
5. Brown, I. and Mues, C. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications*, 39(3), pp.3446-3453.
6. Cerchiara, R., Edwards, M. and Gambini, A. (2017). Generalized linear models in life insurance: decrements and risk factor analysis under solvency ii. pp.1-18.
7. Cholongitas, E., Senzolo, M., Patch, D., Shaw, S., Hui, C. and Burroughs, A. (2006). Review article: scoring systems for assessing prognosis in critically ill adult cirrhotics. *Alimentary Pharmacology and Therapeutics*, 24(3), pp.453-464.
8. Crook, J., Edelman, D. and Thomas, L. (2007). Recent developments in consumer credit risk assessment. *European Journal of Operational Research*, 183(3), pp.1447-1465.
9. Eling, M. and Kochanski, M. (2013). Research on lapse in life insurance: what has been done and what needs to be done?. *The Journal of Risk Finance*, 14(4), pp.392-413.
10. Fier, S. and Liebenberg, A. (2013). Life Insurance Lapse Behavior. *North American Actuarial Journal*, 17(2), pp.153-167.
11. Ghodselahi, A. and Amirmadhi, A. (2011). Application of Artificial Intelligence Techniques for Credit Risk Evaluation. *International Journal of Modeling and Optimization*, pp.243-249.
12. Huang, C., Liao, H. and Chen, M. (2008). Prediction model building and feature selection with support vector machines in breast cancer diagnosis. *Expert Systems with Applications*, 34(1), pp.578-587.
13. Izenman, A. (2013). *Modern multivariate statistical techniques*. New York: Springer, p.369.
14. Karaa, A. and Krichene, A. (2012). Credit–Risk Assessment Using Support Vectors Machine And Multilayer Neural Network Models A Comparative Study Case Of A Tunisian Bank. *Accounting and Management Information Systems*, 11(4), pp.587-620.
15. Kiesenbauer, D. (2012). Main Determinants of Lapse in the German Life Insurance Industry. *North American Actuarial Journal*, 16(1), pp.52-73.
16. Lee, S. and Choi, W. (2013). A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis. *Expert Systems with Applications*, 40(8), pp.2941-2946.
17. Lee, Y. (2007). Application of support vector machines to corporate credit rating prediction. *Expert Systems with Applications*, 33(1), pp.67-74.
18. Life Insurance Association of Malaysia (2017). *Healthy Growth for Life Insurance Industry in First Half 2017*. [online] Available at: https://www.liam.org.my/index.php/newsmedia-room/media-releasepress- statements/english/967-healthy-growth-for-life-insurance-industry-in-first-half-2017 [Accessed 2 Oct. 2017].
19. Memic, D. (2015). Assessing Credit Default using Logistic Regression and Multiple Discriminant Analysis:Empirical Evidence from Bosnia and Herzegovina. *Interdisciplinary Description of Complex Systems*, 13(1), pp.128-153.
20. Milhaud, X., Loisel, S. and Maume-Deschamps, V. (2011). Surrender Triggers In Life Insurance What Main Features Affect The Surrender Behavior In A Classical Economic Context. *Bulletin Francais d'Actuariat, Institut des Actuaires*, 11(22), pp.5-48.
21. Min, J. and Lee, Y. (2005). Bankruptcy Prediction Using Support Vector Machine with Optimal Choice Of Kernel Function Parameters. *Expert Systems with Applications*, 28(4), pp.603-614.
22. Mohammadi, N. and Zangeneh, M. (2016). Customer Credit Risk Assessment using Artificial Neural Networks. *International Journal of Information Technology and Computer Science*, 8(3), pp.58-66.
23. Nazari, M. and Alidadi, M. (2013). Measuring Credit Risk of Bank Customers Using Artificial Neural Network. *Journal of Management Research*, 5(2), p.17.
24. Niklis, D., Doumpos, M. and Gaganis, C. (2011). A Comparison Of Alternative Methodologies For Credit Risk Evaluation. *Journal of Computational Optimization in Economics and Finance*, 3(3), pp.169 -184.
25. Parvin, H., Alizadeh, H. and Minati, B. (2010). A Modification on K-Nearest Neighbor Classifier. *Global Journal of Computer Science and Technology*, 10(14), pp.37 - 41.
26. Stąpor, K. (2016). A Critical Comparison Of Discriminant Analysis And Svm-Based Approaches To Credit Scoring. *Uniwersytetu Ekonomicznego w Katowicach*, pp.59-70.
27. Trustorff, J., Konrad, P. and Leker, J. (2010). Credit risk prediction using support vector machines. *Review of Quantitative Finance and Accounting*, 36(4), pp.565-581.
28. Woods, K. and Bowyer, K. (1997). Generating ROC Curves For Artificial Neural Networks. *IEEE Transactions on Medical Imaging*, 16(3), pp.329-337.


## AUTHORS PROFILE


**Lim Jin Xong** is working in School of Mathematics, Actuarial and Quantitative Studies, Asia Pacific University of Technology and Innovation, Malaysia

**Ho Ming Kang** is working in School of Mathematics, Actuarial and Quantitative Studies, Asia Pacific University of Technology and Innovation, Malaysia