
Dimensionality Reduction for Improving the Performance of Risk Calculation Using Machine Learning Algorithms

^{*1}Vaibhav A. Hiwase, ²Dr. Avinash J. Agrawal

^{1,2} Shri Ramdeobaba College of Engineering and Management, Nagpur, Maharashtra, India
Email: ^{*}hiwaseva@rknc.edu, agrawalaj@rknc.edu

Received: 09th July 2018, Accepted: 14th August 2018, Published: 31st August 2018

Abstract

The profit and loss of the insured person have been mainly associated with the risk of that person. The presence of the uneven distribution of this risk defines different characteristic and lifestyle of that person. Today because of lots of information gathering from technology and ease of availability of data, underwriting and policy-making in life insurance are possible by understanding risk factors from this uneven distribution. The presence of non-risk factors in this uneven distribution only contributes to adding the noise. This noise always increases execution time though not necessarily decrease the accuracy of machine-learning models. Time can be very crucial in real time system and its importance increases in higher dimensional feature space. This paper attempts to deal with such problem in life insurance industry without compromising the accuracy significantly in any machine-learning model. The method used to focus on analyzing hidden correlation among the features taken from different characteristic and behaviour patterns of the insured person and remove unimportant noisy features, which are not playing any significant role to increase and decrease the accuracy of risk calculation. Thus, it helps to decrease execution time substantially. This will help Life Insurance Company to take the fast decision in the competitive world. In addition to this, it will protect the company from financial damage by removing noisy features for data mining.

Keywords: Life Insurance, Adverse Selection, Risk Factor, The Null Hypothesis, Data Mining, Machine Learning, Significance Level.

Introduction

The terms and condition of Life Insurance policy and its popularity is a result of profit or loss and analysis in the claim of policyholders from historical data. Policy mainly decides the size and structure of payment. Hence, it became a challenging task to define a perfect size and structure of payment from which Life Insurance Company can prevent failure, make some profit and attract new customers. Although it is important to decide accurate size and structure of payment, the company may fail to do so due to adverse selection. Adverse selection in life insurance industry is a situation where the purchase of insurance policy is affected by the asymmetric information. This

symmetric information creates a potential threat of policyholders claiming earlier than expected period. They became a threat when they are paying low premium payment. Therefore, there has to be some kind of measure by which we can identify this potential threat by using information from historical data and charge them high premium payment. One of the measures is a risk. Correct premium payment is generally associated with risk of a person in Life Insurance Company. This risk is nothing but the life expectancy i.e. an average time a person is expected to live. This risk in life insurance can be of ordered category. If we somehow able to understand the association of this ordered risk category with characteristic information and behaviour pattern of a person, then ordered risk will become an important dependent variable to decide accurate premium payment for the insurance industry. Thus, the problem of preventing loss can narrow down to accurate classification of the policyholder in the ordered risk category.

Age and disability are key risk factors that cannot be readily substituted by alternative risk factors for many types of insurance. Risk-rated individual life insurer currently segmented in risk group by Age, Disability, Occupation, Leisure pursuits, Amount and duration of cover, Education, Income, Dwelling location, Behavioural habits (like smoking, drinking, drugs) ("Use of Age and Disability as Rating Factors in Insurance: Why Are They Used and What Would Be the Implications of Restricting Their Use?", position paper.", 2011) [1]. This factor though considered, as important factors, they might still not give any guarantee for classifying customers into the right risk group. So (Kahane, Y. e. 2007) [2] used the excess of 200 attributes which further classified into four groups to build a predictive model which is useful for underwriting and rate making. They used a two-stage approach, which involves survival analysis, and linear regression model. This helps them to estimate the risk level of each customer and the proneness to file a claim. Their results show that riskiest people on an average 12 times more expensive than least risky people in motor vehicles insurance. This is a clear indication that excesses of 200 attributes help them in classifications of the people.

One of the biggest problems in working with high dimensional data is the presence of irrelevant features

in a cluster and correlation of relevant features in different clustering. This is the main challenge in clustering and can be thought as a curse of dimensionality. Zimek, 2008 [3] suggests some of the interesting approaches for solving curse of dimensionality, which can be tackled individually, or in the combination with a different approach. On the other hand, (Devi, 2016) [4] examine how to get meaningful input variables for creating a model to do that. They use rule-based cluster model for motor policies.

Finkelstein, 2014 [5] tried to use asymmetric information and identify individual characteristics that are risk relevant and correlated with insurance demand, but still unused by insurance companies. Their results show that political economy of insurance regulation may play an important role in determining pricing function and so we can expect to find some other interesting knowledge. On the other hand, (Rahman, 2017) [6] apply attribute selection techniques to properly classify the data and prove that classification techniques are very useful in classifying customers according to their attributes.

The compulsory purchase of Life Insurance policy in the market from the government can make Life Insurance Company too big to fail and hence they can tolerate adverse selection. However, this method is not applicable everywhere. Therefore the calculation of ordered risk of each policyholder from lots of characteristic and behavioural information may be helpful.

Proposed system helps us to understand these important risk factors by eliminating irrelevant features that are not having statistical significance and test whether they can contribute to the identification of the important features. This system will help to enhance execution time of any machine-learning model by identifying statistically significant features without compromising the significant accuracy of the model.

Experimental Set-Up

Collection of Dataset and Preprocessing

The dataset was taken from Prudential Life Insurance Assessment available on kaggle, which is a platform for predictive modelling and analytics competitions. This data was available to make buying life insurance easier. In this dataset, over a hundred variables describing attributes of life insurance applicants are provided. The task is to predict the "Response" variable for each Id in the test set. "Response" is an ordinal measure of risk that has 8 levels. There are total 60 categorical features and 13 continuous variables, 5 discrete variables and 48 dummy variables in this dataset. We pre-processed 60 categorical features are into 767 dummy variables without

creating dummy variable trap. Therefore, total 833 predictors are used in the analysis.

Proposed System

In the proposed system the predictors are the features extracted from the dataset. They are treated as a null hypothesis. This will imply that there is no relationship between dependent variable and one or more combination of independent feature variable(s). Rejecting or disproving the null hypothesis gives a ground to believe that there is a significant relationship between dependent and independent variable(s). Whereas accepting the null hypothesis only provides the insignificant relationship between them. Hence, it is better to remove predictors if they are incapable to contribute a significant increase in accuracy of any machine-learning model. If not removed, then their presence will only increase the execution time of machine learning models and will make model worst. So, it becomes important to understand this relationship by running test of the Null hypothesis on some method by which we can have a guarantee to reduce type I error with a goal of making the power of test close to one. One such method is known as the multiple linear regression.

The multiple linear regression studies the relationship between a dependent variable and one or more independent variable(s). The multiple linear regression equations can be written as follows:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k + \varepsilon$$

Where y is the dependent variable i.e. response variable, the x 's are the independent variables, the b 's are the regression coefficients, b_0 is the constant (intercept) and ε is the residual variable. The multiple linear regression has the same goal of linear regression i.e. to explain the behaviour of response variable based on the independent variable(s).

Following assumptions should be check before running multiple linear regression:

1. All the variables are continuous and dichotomous.
2. The relationships between the dependent variable and the independent variables are linear, both for each independent variable and globally.
3. There are no significant outliers in the data series.
4. There is the independence of errors (i.e. there is no relationship between the independent variables and the residual variable).
5. The dependent variable has the same variance for all the values of the independent variables (i.e. the assumption of homoscedasticity).
6. The residual variable is approximately normally distributed.
7. The independent variables are not strongly correlated with one another (i.e. we don't have important multicollinearity). Dummy variable trap is a situation in which independent variables are multicollinear. If variables are multicollinear

then they are highly correlated with each other. Each variable then can be predicted from remaining variables and thus model cannot able to distinguish between them from one another. Therefore it is very important to eliminate dummy variable trap to train model perfectly. The only solution for eliminating dummy variable trap is to drop any one of the dummy variable or alternatively, drop the intercept constant.

Feature Extraction

One of the following ways can do feature extraction in multiple linear regression:

1. Backward Elimination
2. Forward Selection
3. Bidirectional Elimination

Backward Elimination

In this method, all features are selected to fit the model and the one, which is most statistically insignificant in the selected feature, is removed. In the next iteration, the model is fitted without this variable. This iterative process helps us to identify features that do not play an important role because they act as a garbage within the data that increases time complexity of the machine learning models. In addition, they do not contribute to significantly increase in model accuracy as they fail to explain significant variance in the response variable. Backward elimination is simplest and fastest method to implement compared to other two. The steps used in the backward elimination algorithm are given below.

Step 1: Select a significance level to stay in the model (e.g. significance level = 0.05).

Step 2: Fit the full model with all possible predictors.

Step 3: Consider the predictor with the highest p-value. If $p >$ significance level, go to step 4, otherwise finish and model is ready.

Step 4: Remove the predictor.

Step 5: Fit the model without variable.

Forward Selection

This method is opposite of backward elimination. In this method, feature variable is added to the model one by one. From the added features the one which is most statistically significant is used to fit the model. Once one feature is added to the model then the next feature is selected from the remaining features out of which the most statistically significant feature is used in combination with the already fitted features within the model. The steps used in the forward selection algorithm are given below.

Step 1: Select a significance level to enter the model (e.g. significance level = 0.05).

Step 2: Fit all simple regression model. Select the one with the lowest p-value.

Step 3: Keep this variable and fit all possible models with one extra predictor added to the one you already have.

Step 4: Consider all the predictor with the lowest p-value. If $p <$ significance level, go to step 3, otherwise go to finish and model is ready.

Bidirectional Elimination

This method is the combination of the above two methods. In each step, either the feature variable is included or it is excluded based on the testing result. Here we keep on adding variables using a method of forward selection and then check its significance level with the one we already have. Then in the next step, we try to remove variable which is most statistically insignificant by backward elimination method. This will help to separate out the entire variables that are statistically insignificant from the statistically significant one. The steps used in the bidirectional elimination algorithm are given below.

Step 1: Select a significance level to enter and to stay in the model (e.g. significance level = 0.05).

Step 2: Perform the next step of forward selection.

Step 3: Perform all step of backward elimination.

Step 4: No new variable can enter and no old variable can exit.

Coefficient of Determination (R^2 or r^2):

The coefficient of determination also known as “R squared” is the proportion of the variance in the dependent variable that is predictable from the independent variable(s). This statistic explains how well the predicted outcome is explained by the model based on the variation of response variable. In multiple linear regression, the linear function is used to explain variation. The value of R^2 ranges from 0 to 1. Negative values of R^2 indicates fitting of a non-linear function to data. In this case, the mean of the data provides a better fit to data. R^2 states the goodness of fit of the regression model. It is equal to one minus the ratio of the sum of squared estimated errors (the deviation of the actual value of the dependent variable from the regression line) to the sum of squared deviations about the mean of the dependent variable. The variability of the data set can be measured using three sums of squares formulas:

1. The total sum of squares (SS_{tot}) (proportional to the variance of the data) is defined as being the sum, over all observations, of the squared differences of each observation y_i from the overall mean \bar{y} :

$$SS_{tot} = \sum_i (y_i - \bar{y})^2$$

- The regression sum of squares (SS_{reg}), also called the explained sum of squares is the sum of the squares of the deviations of the predicted values f_i from the mean value of a response variable \bar{y} :

$$SS_{reg} = \sum_i (f_i - \bar{y})^2$$

- The sum of squares of residuals (SS_{res}), also called the residual sum of squares is the sum over all observations, of the squared differences of each observation y_i from the deviations of the predicted values f_i :

$$SS_{res} = \sum_i (y_i - f_i)^2 = \sum_i e_i^2$$

The most general definition of the coefficient of determination is

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

In general, $SS_{tot} = SS_{reg} + SS_{res}$.

One drawback of R squared is that once we added the feature to the model, the regression process will either find a way to give the coefficient to the newly added feature variable such that the SS_{res} will be minimized or the regression process will make the coefficient of that newly added feature variable equal to zero and thus excluded from the regression equation. In the former case, the SS_{tot} will not change significantly but the SS_{res} minimized so that the R^2 value always increases. Whereas in latter case we do not add any new feature to the regression model. Hence, R^2 value does not change. In both the cases, R^2 value never decreases. Therefore R squared cannot be a useful measure to clearly understand the significance of important features.

The drawback of R squared is eliminated by Adjusted R^2 . Adjusted R^2 has a penalization factor. It gives penalty by adding independent variables that don't help the model. If p is the number of regressors and n is the sample size, then $Adj R^2$ is defined as follows:

$$Adj R^2 = 1 - (1 - R^2) \frac{n - 1}{n - p - 1}$$

If we add more regressors, the $Adj R^2$ will start decreasing but if the independent variable is significantly important then R^2 will start increasing. So in $Adj R^2$ there is a battle between the independent variable and its significant importance. Therefore $Adj R^2$ only increases if newly added feature has significant importance in the regression model. In addition, if $Adj R^2$ decreases then it signifies that newly added feature tries to degrade model accuracy. If $Adj R^2$ neither increases nor decreases then the accuracy of model does not increase significantly, but we can still decrease model execution time by

deselecting unimportant features using feature extraction algorithms in multiple linear regression explained earlier.

The proposed system used ordinary least squares (OLS) or linear least squares method. This is a method for estimating the unknown parameters in a linear regression model. In addition, $Adj R^2$ was used to evaluate important feature combination. Thus, the proposed system will also help researchers by narrow down consideration of features to identify hidden patterns in the data. These patterns will have a good correlation with the response variable.

Out of three feature extraction methods explained earlier, proposed system uses only backward elimination algorithm because of its simplicity and fast processing. The steps are as follows:

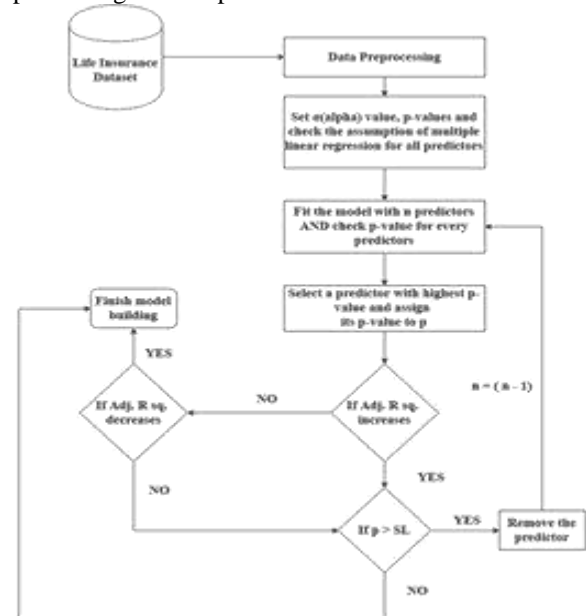


Figure 1: Flow Chart Describing Steps of Proposed System

Steps of Proposed System:

Step 1: Data Preprocessing:

Real-world data generally are inconsistent, incomplete and lacking in certain behaviour and trends. Hence preprocessing is necessary. Data goes through series of steps during preprocessing like data cleaning, data integration, data transformation, data reduction, data discretization and handling categorical explanatory variables in regression models.

Step 2: Set α (alpha) value, p-values and check the assumption of multiple linear regression for all predictors:

Alpha value should be decided depending upon the accuracy and selection of model. Assumptions of regression test i.e. Linearity, Homoscedasticity, Multivariate Normality, Independence of errors, lack

of multicollinearity should be checked before proceeding further.

Step 3: Fit the model with n predictors AND check p-value for every predictor:

Use model to run the Multiple Linear Regression of the null hypothesis keeping the power of test close to one. This will ensure we will not accept the false negative results. Our Null hypothesis initially include all predictors. These predictors decrease by one in every iteration. Thus, the Null hypothesis changes in every iteration and is equal to the difference between total number of predictors and total number of iteration. Use risk as a dependent variable for life insurance data and characteristic features as the independent variables as our null hypothesis. Null hypothesis assume that there is no significant linear relationship between characteristic independent variables and risk response variable. This can be ensured by checking p-value for every predictor. If $p >$ significance level, we can accept the Null hypothesis and remove some predictors. Rejection of the Null hypothesis will give significantly important features and help us to increase accuracy when we can find strong correlation of predictor with response variable.

Step 4: Select a predictor with highest p-value and assign its p-value to p :

The p-value of each predictor is calculated in step 3. Out of the calculated p-values of predictors, the one with the highest p-value is used. It is denoted by p in step 3. If p-value is too high then it is not useful for machine learning models. Hence, the possibility of removing that predictor becomes high. If we remove all high p-value predictors in each iteration then we will be left with predictors that may have statistically significant strong correlation with response variable.

Step 5: Remove the predictor:

If the null hypothesis is rejected, then we will be left with predictors that are statistically significant. These predictors then can be useful for training machine-learning model. This rejection of the null hypothesis is also called as acceptance of alternative hypothesis. If the null hypothesis is accepted, then system will remove the predictors that are responsible for acceptance of null hypothesis. When this will happen, we should go for next meaningful predictors.

The acceptance or rejection of the null hypothesis is decided by following two ways:

1. **Using $Adj R^2$:** If $Adj R^2$ decreases then we should stop removing predictors from previous iteration and stop building model. If $Adj R^2$ increases or same as that of predictor in the previous iteration, then we should check its p-value.

2. **Using p-value:** If p-value is not statistically significant or p-value is greater than significance level, then we should remove that predictor by accepting the null hypothesis and should not use it in the next iteration. If p-value is statistically significant or p-value is lesser or equal to significance level, then we can finish building model by rejecting final null hypothesis. The purpose of this model is to build important predictors from features.

If this model fails to increase accuracy substantially, then also it will be useful to provide lesser but important predictors contributing to decrease in execution time of any machine-learning model.

Model Evaluating Parameters k-fold Cross-Validation

In k-fold cross-validation technique the original dataset sample is partitioned into k equal size subsamples. Out of k -subsamples, a single subsample is retained for testing the model and remaining $k - 1$ subsamples are used for training the model. This process is repeated for k times and each time different subsamples are used for testing the model. Thus, all k subsamples are used to test the accuracy of the model. Since accuracy is taken from all k subsamples of the dataset, this process helps to overcome the problem commonly known as overfitting. In overfitting, the model is too closely fit by training dataset that it cannot be able to give accurate prediction of the new data points other than training dataset. In addition, sometimes model learns too much of the noise and random fluctuations from training dataset in such a way that it fails to generalize itself for testing dataset. If the added features or predictors from proposed system negatively learn these concepts and overfit the model from noise and random fluctuations, then model accuracy will decrease for testing data and increase for training data. In order to minimize this opposite effect 10-fold cross validation is used to evaluate model's performance.

The proposed system can evaluate the problem of under fitting by testing the predictors generated from multiple linear regression model on different machine-learning models. In addition, the model with the most accurate result with the small execution time should be selected ideally. If there is a tradeoff between the accuracy and the time, then decision should be based on domain knowledge and the purpose of building machine learning models. If purpose is to build a real time system then preference should be given to time and vice versa for accuracy.

Results and Discussion

The accuracy and execution time is calculated for training of the model as well as prediction of the model. Initially we use all features for the above calculations and then it is compared with statistically significant features. The comparison of the percentage of time required to train and tests the machine-learning models are shown from figure 2 to figure 7. This comparison also indicates the accuracy along with the different number of predictors which is decreasing along the horizontal axis along with significance level.

From proposed system we select all 833 predictors and determine execution time for training and testing of some classification machine learning model. The classification was done on 8 risk level of the response variable. Out of 833 predictors the one with the highest p-value was selected and removed to see the increase in $Adj R^2$ value. If we get the increase or no change in $Adj R^2$ value then we go check for p-value to determine statistical significance. If the combination of all 833 selected predictors are statistically significant then we finish model building. If $Adj R^2$ decreases then we go for finish model building without checking significance level since the model accuracy will then start decreasing significantly.

Our result did not show any increase or decrease in $Adj R^2$ value. Therefore, we go for significance level of p-value. Whenever we accept the Null hypothesis, we got p-value >Significance level. At that time, we removed the highest p-value predictor and start building the model excluding that predictor. This is the process of backward elimination algorithm. We tried to run the model on different significance level viz. 10%, 5% and 1%. In addition, we do check the increment in $Adj R^2$ value every time. When an increment in $Adj R^2$ value takes place, we selected the increment, which only happened closer towards p-value of greater than 0.10 in which the selected number of predictor decreases from 833 to 332. Similarly, the increment of $Adj R^2$ value happened closer towards p-value of greater than 0.05 decreases the selected number of predictors from 215 to 175 and finally the increment of $Adj R^2$ value happened closer towards p-value of 0.01 decreases the selected number of predictors from 168 to 155. The significance level of the p-value of 0.10, 0.05 and 0.01 had a number of predictors equal to 215, 168 and 106 respectively.

The experiment was performed on six classification machine learning model viz. Bernoulli Naive Bayes model, Gaussian Naive Bayes, Decision Tree Classification model, K-Nearest Neighbours model, Logistic Regression model and Random Forest model.

Out of the six models, we found out that K-Nearest Neighbours and Random Forest model was showing consistent increase in accuracy up to 1% significance level of p-value. In addition, random forest took the least amount of execution time when compared to remaining models. Some models are showing decrease in accuracy but since that decrease is not greater than 0.6% we can rely on proposed method because it reduces approximately 80% and 87% unimportant predictors at 5% and 1% significance level respectively. This decrease in predictors plays a very important role to reduce execution time of the machine learning model in real time application.

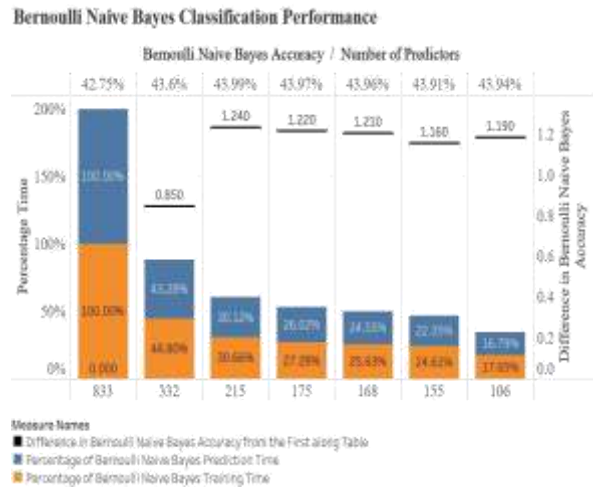


Figure 2. Bernoulli Naive Bayes Classification Performance.

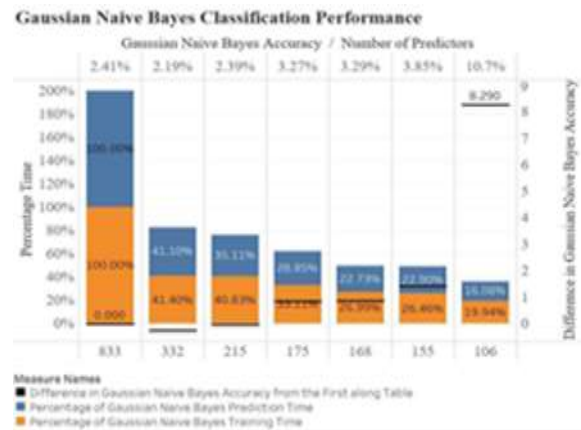


Figure 3. Gaussian Naive Bayes Classification Model Performance.



Figure 4. Decision Tree Classification Model Performance.



Figure 5. K-Nearest Neighbours Classification Model Performance.



Figure 6. Logistic Regression Model Performance.



Figure 7. Random Forest Classification Model Performance.

Conclusion

Traditionally segmentation was used in life insurance with the assumption that all individuals in the segment are “alike”. The problem was small segments suffers from lack of statistical significance and poor prediction whereas large segments may not be homogenous for decision-making. Clustering had a humble origin in life insurance and its application seems to give promising results. Today, it is important to know the behaviour of clustering in higher dimensional feature space as insurance industry needs to be more accurate to perform well in the competition. In addition, the adverse selection is a critical situation caused due to asymmetric information in the database and unshared private information from insurance buyers about their risk factors. To understand this effect and behaviour of the features in higher dimensional feature space, we cannot rely only on clustering. Proposed system removes unimportant predictors from analysis and decreases execution time of machine learning models. This will enable the firm to understand the risk factors in higher dimensional feature space so that they can be useful for accurate prediction of the risk of policyholder on an individual basis in Life Insurance Company. These risk factors can then become as the counterforce against adverse selection and useful for enhancing the loyalty of policyholder and profit as well.

Future Scope

Proposed system did not try increase the accuracy of machine-learning model by understanding correlation and by deriving new unknown feature from available dataset. Since this decision is completely depend of domain knowledge, proposed system cannot rely on that. Possible solution can comes from brute-force method, but for higher dimensional data, it is very time consuming job. Therefore, the task of increasing the accuracy by deriving new features from correlation is the future area of this research.

References

- [1] (2011). "Use of Age and Disability as Rating Factors in Insurance: Why Are They Used and What Would Be the Implications of Restricting Their Use?, position paper.". Européen, Groupe Consultatif Actuariel.
- [2] Kahane, Y. e. (2007). "Applying data mining technology for insurance rate making: an example of automobile insurance.".Asia-Pacific Journal of Risk and Insurance 2.1.
- [4] Zimek, A. (2008). "Correlation Clustering". Dissertation, LMU München: Faculty of Mathematics, Computer Science, and Statistics. Devi, O. (2016). "Portfolio rule-based clustering at automobile insurance in Portugal. Diss". APA.
- [5] Finkelstein, A., & Poterba, J. (2014). Testing for asymmetric information using "unused observables" in insurance markets: Evidence from UK annuity market. *Journal of Risk & Insurance*, 81(4),709-734.
- [6] Rahman, M. S. (2017). "Analyzing Life Insurance Data with Different Classification Techniques for Customers' Behavior Analysis." *Advanced Topics in Intelligent Information and Database Systems*. Springer International Publishing, 15-25.
- [7] Ansari, A. a. (2016). "Customer clustering using a combination of fuzzy c-means and genetic algorithms." *International Journal of Business and Management* 11.7, 59.
- [8] Bezdek, J. C. (1984). "FCM: The fuzzy c-means clustering algorithm.". *Computers & Geosciences* 10.2-3. 191-203.
- [9] Roiger, R. J. (2017). *Data mining: a tutorial based primer*. CRC Press.
- [10] Lemaire, J. (1990). "Fuzzy insurance.". *ASTIN Bulletin: The Journal of the IAA* 20.1, 33-55.
- [11] Madeira, S. a. (2002). "Comparison of target selection methods in direct marketing." *European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems*.
- [12] Ostaszewski, R. A. (1995). "Fuzzy Techniques of Pattern Recognition in Risk and Claim Classification". *Journal of Risk and Insurance*, vol. 62, Issue 3, 447-482.
- [13] Vaibhav A. Hiwase & Dr Avinash J Agrawal. (2018). Review On Application of Data Mining in Life Insurance Accepted in International Conference on Recent Trends in Engineering & Science (ICRTES). To be published in coming issues Vol.8 of International Journal of Engineering and Technology (UAE). E-ISSN: 2227-524X.
- [14] Agrawal, A. J. (2008, April). Using domain specific question answering technique for automatic railways inquiry on mobile phone. In *Information Technology: New Generations*, 2008. ITNG 2008.

Fifth International Conference on (pp. 1111-1116). IEEE.

[15] Dixit, S., & Agrawal, A. J. (2013). Survey on review spam detection. *Int J Comput Commun Technol ISSN (PRINT)*, 4, 0975-7449.