Metadata of the chapter that will be visualized in SpringerLink

Book Title	Information and Communication Technology and Applications	
Series Title		
Chapter Title	Credit Risk Prediction in Commercial Bank Using Chi-Square with SVM-RBF	
Copyright Year	2021	
Copyright HolderName	Springer Nature Switzerland AG	
Author	Family Name	**Alabi**
	Particle	
	Given Name	**Kayode Omotosho**
	Prefix	
	Suffix	
	Role	
	Division	Department of Computer Science
	Organization	Kwara State University
	Address	Malete, Nigeria
	Email	
Author	Family Name	**Abdulsalam**
	Particle	
	Given Name	**Sulaiman Olaniyi**
	Prefix	
	Suffix	
	Role	
	Division	Department of Computer Science
	Organization	Kwara State University
	Address	Malete, Nigeria
	Email	
Corresponding Author	Family Name	**Ogundokun**
	Particle	
	Given Name	**Roseline Oluwaseun**
	Prefix	
	Suffix	
	Role	
	Division	Department of Computer Science
	Organization	Landmark University
	Address	Omu-Aran, Nigeria
	Email	ogundokun.roseline@lmu.edu.ng
	ORCID	http://orcid.org/0000-0002-2592-2824
Author	Family Name	**Arowolo**
	Particle	
	Given Name	**Micheal Olaolu**
	Prefix	

Suffix	
Role	
Division	Department of Computer Science
Organization	Landmark University
Address	Omu-Aran, Nigeria
Email	arowolo.olaolu@lmu.edu.ng

Abstract	Financial credit risk analysis management has been a foremost influence with a lot of challenges, especially for banks reducing their principal loss. In this study, the machine learning technique is a promising area used for credit scoring for analyzing risks in banks. It has become critical to extract beneficial knowledge for a great number of complex datasets. In this study, a machine learning approach using Chi-Square with SVM-RBF classifier was analyzed for Taiwan bank credit data. The model provides important information with enhanced accuracy that will help in predicting loan status. The experiment achieves 93% accuracy compared to the state-of-the-art.

Keywords (separated by '-')	Credit risk - Chi-square - SVM - Bank - Machine learning

Author Proof

Check for updates

Credit Risk Prediction in Commercial Bank Using Chi-Square with SVM-RBF

Kayode Omotosho Alabi[1], Sulaiman Olaniyi Abdulsalam[1], Roseline Oluwaseun Ogundokun[2(✉)], and Micheal Olaolu Arowolo[2]

[1] Department of Computer Science, Kwara State University, Malete, Nigeria
[2] Department of Computer Science, Landmark University, Omu-Aran, Nigeria
{ogundokun.roseline,arowolo.olaolu}@lmu.edu.ng

Abstract. Financial credit risk analysis management has been a foremost influence with a lot of challenges, especially for banks reducing their principal loss. In this study, the machine learning technique is a promising area used for credit scoring for analyzing risks in banks. It has become critical to extract beneficial knowledge for a great number of complex datasets. In this study, a machine learning approach using Chi-Square with SVM-RBF classifier was analyzed for Taiwan bank credit data. The model provides important information with enhanced accuracy that will help in predicting loan status. The experiment achieves 93% accuracy compared to the state-of-the-art.

Keywords: Credit risk · Chi-square · SVM · Bank · Machine learning

1 Introduction

Credit risk prediction is a critical issue faced by several sectors such as banks, telecommunications, among others. Nowadays, it helps in evaluating the loyalty of customers and helps the bank in minimizing possible loses and maximize the volume of credits [1]. Credit risk contains the peril taken by the bank, because a customer may fail to meet their loan obligations requiring prediction of the customer by using previous records and information for proper decisions [2]. Divers technology innovations have made it possible for banking sectors to incorporate efficient delivery channels, and deal with several challenges posed by the economy. Banks need to take customers relationships as an important factor that determines their success [3].

Machine learning has proven to be a method that helps in fetching relevant information from huge data with hidden pattern [4–6] that can to solving challenges of whom is worthy of giving loans to and not [7, 8]. A lot of machine methods have been proposed by different researchers, using different algorithms to improve decision-making.

In recent times, several approaches like stochastic optimization technique, evolutionary algorithms and support vector machine (SVM) have revealed capable results in terms of accuracy prediction [9–11].

In this study, a feature selection based on Chi-square algorithm is proposed with SVM classifier for credit scoring tasks, using a Taiwan bank credit dataset, for systemic

© Springer Nature Switzerland AG 2021
S. Misra and B. Muhammad-Bello (Eds.): ICTA 2020, CCIS 1350, pp. 1–12, 2021.
https://doi.org/10.1007/978-3-030-69143-1_13

Author Proof

attributed risks of the banking system, to study and analyzed the design and compare these model on evaluation areas that identify risk customers from a large number of customers and provide an effective procedure for a bank loan agreement.

2 Related Works

Several studies have been carried out in literature by numerous researchers to further enhance prediction of credit risks in banks, as loses encountered in decision making for loan issuance to customers. This section discusses several works done with approaches used to achieve better decision making.

Bank credit default prediction based on a machine learning approach was proposed [7]; it is of the essence to identify the potential risks in giving out bank loans to customers. The objective of their study was to use the classification methods in classifying loans for customers. Random Forest, Logistic Regression, SVM classification algorithms were used to analyze the bank credit data, compared, and identify risk customers from large numbers of customers by providing the solid base for bank loan approvals. Their result showed Random forest outperforming other techniques with 88%. This paper contributed, by providing an efficient experimental basis for bank loan approvals and identifying risks of issuing huge loans to customers using classification techniques, they used a small public bank credit dataset, that suggests the experiment is not well comprehensive enough.

A credit risk analysis and prediction model for bank loans using R was proposed for related risks in bank loans, to reduce the problem of capital loss. The objective of their study is to design a model to identify probable defaulters of bank loan applicants effectively, they used clustering and classification techniques in their methodology, and the class labels were predicted, their experiments achieved high accuracy of 94.3%. In their contribution, they made an effective framework for identifying the probability of customers to defaulting bank loans. Their work was limited to an open-source repository and suggested more efficient data with a better tool for the application of machine learning techniques [1].

There requires the need for the bank industry to predictive modelling issues for credit defaulters accurately. The objective of this paper was to create a credit scoring model using machine learning approach. A prediction loan status for commercial banks using machine learning classifier was proposed [2], by analyzing model for credit data, they used a min-max normalization and KNN classifiers in their method. The model designed provided important information with efficient accuracy of 75%. In their contribution, the proposed loan status model that can help predict validly and defaulting loan customers. The study's limitation suggests further study on the iterations to avoid huge losses in banks.

It is of importance for banking sectors to strategic in their mode of analysis. Data mining has proven to help discover patterns and relationship in data [13–15]. An overview of machine learning techniques and its applications in the banking sector was studied. The objective of this study was to analyze data mining techniques for decision making in banking sectors, several data mining technique methods were analyzed for fraud prevention, detection, retention, marketing and risk management [3]. The result of their

Author Proof

study showed several promising approaches, and their contributions were to suggest a combination of various techniques on bank records. Analyzing credit reis is a major problem; the objective of this study was to propose an overview for bank credit risk prediction using systemic reviews where applications of artificial intelligence like ANN were found efficient. In their result, several applications and methods for credit risk research were discussed, and they made a huge contribution by reviewing several works and suggested other approaches for a more accurate prediction approach, and systematic reviews for developing credit risk assessment approaches [16, 17].

A major problem the financial institution faces is credit risks. The objective is to suggest a credit risk prediction model based on SVM method for enterprises [18], using an index approach that considers leading enterprises credit status with neural network for alleviating credit rationing on SMEs, their result was promising, with 94% accuracy. In their contribution, they suggested an approach to help solve the problem of banks incorrect labelling of worthy creditors as a default customer and improve the credit rating status. The limitation of their study evolved around the superiority of predicting default behaviors in customers due to insufficient sample size data. Determination of the probability of defaulters in commercial banks is a big problem; several machine learning approaches have been suggested and requires good suggestions. The objective of this study was to propose a Credit risk assessment using ANN and SVM [19], a comparison was made using performance metrics with the confusion matrix to identify the most efficient model, their results showed RBF-SVM outperformed with 88% accuracy, which is a very efficient model to improve Tunisia credit risk management for monitoring and managing credits in banks. The major contribution of their study was to check through the ability of machine learning models for prediction of credit risks in customers seeking bank loans. The study suggests a hybridize model due to the limitation of the scoring model studied corporate account data.

Analyzing and predicting loan payback is a huge problem in the banking sector. The major objective of this study was to find the customers nature of the application for loans by exploring the analytical techniques of the record for prediction of loans was proposed [20], to find the nature of clients applying for personal loans using an exploratory data analysis technique in dealing with problems. The study used several analytical methods such as the customer income, classification of customers, terms of loan, categories, experience and job. The result analysis showed short term loans. The main contribution of the study is by classifying and analyzing the nature of the applicant's loan. The limitation of the study suggested an extension of the study into a higher level and consideration of other factors for loan prediction.

Credit risk management and its analytical approaches is a major issue in decision making for the financial sector. The objective of this study is to construct a credit scoring model. Credit scoring using feature selection with deep learning method [9], to evaluate applicants credit score from applicants input features on Australian and German credit data, their result showed an efficient prediction rate with comparable performance with 86% overall accuracy. The contribution of this work focused on determining higher acceptable accuracy. The workload for credit evaluation is a limitation as there was less computational intensiveness.

Author Proof

Problems encounter in financial sectors for identifying defaulters and granting credits is alarming. This study proposed model types to address credit scoring. Credit scoring metaheuristic approaches have been proposed [21], with artificial intelligence techniques to address assessment procedures with a compilation of past experiments and results. This study majorly contributed to the formulation of a credit scoring model using SVM and MA approach. Various limitations were observed in the assessment procedure and suggested futuristic directions.

Risks in the banking sector is a problem and require a better analytical approach. The objective of this study is to apply machine learning technique to predict banking risk systems using SVM [22], was suggested to predict bank risks in an attempt to suggest an innovative model with better explanation stability, SVM classifier with Chinese bank data projected an efficient result of 85% accuracy. The major contribution of the study was to validate the prediction performance using SVM classifier and to construct prediction indicators. The empirical analysis of the prediction accuracy can be further enhanced due to the limitation in the prediction model. A problem associated with SVM evaluation of variable predictors requires more work. The objective of this study was to extend an RFE feature selection approach for proposing the visualization of iterations and identification of relevant predictors for non-linear kernel methods. There result proved RFE-pseudo-samples outperformed other methods. The contribution of their study is majorly on the interpretation of variables relevant for classification, with a more computational complexity as a limitation of the study [23].

Business failure prediction is a challenge. Machine learning algorithms [24] were applied to business failure prediction [25]. The objective of the study was to investigate efficient algorithms for commercial bank distress. They proposed a cost-effective prediction variable selected from variable eliminating phases and sensitive attribute evaluators, by considering Naïve Bayes, KNN, ANN, SVM, C4.5, CHAID and CRT methods. Their experiment showed an evaluation performance with enhancement and comparisons of 91% overall accuracy. The major contribution of this study was to efficiently utilize machine learning algorithms for financial distress prediction with a limitation of not assessing financial structure, which can be misleading in prediction methods.

The credit scoring problems in banks requires improvement of predictive approaches that are relevant. The objective of this study is to improve prediction analysis by introducing a credit scoring machine learning technique, using PCA and Binary logistic regression [27] was proposed to improve the predictive power of prediction credit applicant status and compared, with an accuracy of 72%. The major contribution of this study is by using binary logistic regression to predict customers status in the bank. The limitation of the output requires futuristic works to enhance the accuracy and performance of the study.

The problem of predicting factors in the financial institution has drawn the attention of researchers, the objective of this study is to predict microfinance credit defaulter in a Ghana bank [28], by examining factors predicting the bank's credit fault, by using chi-square with classification, the result achieved 83%. This study contributed by predicting loan determinant through customers record and requires further study by consulting other machine learning approach, due to the limitation in the record used.

Author Proof

There is a problem existing in the influencing of banking adoptions. Chi-square approach was used for banking adoption as an approach [29], to investigate factors influencing adoption of bank credits and reveal more insights.

3 Materials and Methods

3.1 Bank Credit Scoring Benefits

A lot of benefits come along with credit scoring, for both lenders, borrowers and the banking industry. It helps in reducing discrimination by providing objective analysis of creditworthiness to customers. It enables credit providers in focusing on the relevant information relating to credit risk and avoiding of personal subjective of an analyst such as; formal loan application, interests, terms and conditions, charges, among others [27].

3.2 Data Mining

Datamining is the discovery of the knowledge of huge amount of data, which helps in discovering interesting patterns within the data for decision making, that can help predicts and classify the behaviour of the model [13, 15, 30]. Converting the information into meaningful form is a needful competitive intelligence. Several techniques have existed in literature, and have been used for creating predictive models. Several dimensionality reduction, clustering and classification approach for datamining have been used to model the prediction of chur customers in the banking sector. Some of these techniques used to achieve these models are the decision trees, SVM, Neural networks, random forest, k-means, PCA, ANOVA [31, 32], among others.

3.3 Dataset

In this study, a customer credits risk prediction was proposed for determining the churn rate of the financial sector. The predictive model uses an open-source Taiwan bank credit dataset with 30,000 instances, and 25 attributes [26], Chi-square and SVM approach is proposed to estimate the credit scoring model.

3.4 Method

In this study a Taiwan credit data is used for the credit risk prediction, Chi-Square algorithm is used as a feature selection algorithm to eliminate unwanted information in the data, the SVM classifier is used to evaluate performance. Figure 1 shows the workflow for the study bypassing the high dimensional data into chi-square, the reduced data is classified, using SVM, and the results of the experiment are evaluated.

Author Proof

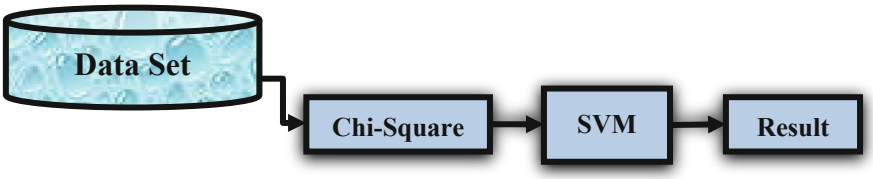


Fig. 1. The framework of the experiment

3.5 Chi-Square Algorithm

Chi-square is a non-parametric algorithm in machine learning used for the testing hypothesis of an association between groups or population, for testing observed distribution [33, 34]. The chi-Square approach consists of specifying the hypothesis, devise an analytical plan, examine the sampling data, and deduce the results [35].

$$\sum X_{i-j}^2 = \frac{(O-E)^2}{E} \tag{1}$$

Chi-square compares obtained values for the frequencies of classes, by splitting predictable frequencies of the class. Samples N_{ij} are the number of samples of C_i classes in the jth intervals. Samples of jth interval is represented by Nij. The predictable frequency of N_{ij} is the $E_{ij} = M_{ij}|C_i|/N$.

Algorithm 1: Chi-Square Algorithm [27]

Chi-squared
Original

```
    /*
        chi_array: two-dimensional array. First column contains feature indexes in DS and
                   second column contains Chi-squared values for the features and class labels.
    */
1:  chi_array ← ∅
2:  For i←1 To N Do
3:      chi_val ← chi.squared(DS[i], CL)   // calculate Chi-squared value between features in DS and CL
4:      Append (i, chi_val) To chi_array
5:  End For
6:  Sort chi_array by the second column (Chi-squaref value) in descending order
7:  Store first column values of chi  array To CHOSEN
8:  Return CHOSEN
```

Chi-square is a discrete data hypothesis approach; it evaluates correlations between variables and fixes the independents that are correlated positively or not. Chi-Square hypothetical test methods for counting huge data belongs to the non-parametric test category. Its idea is to majorly compare degrees of frequencies and fitness.

3.6 Support Vector Machine (SVM)

SVM is a supervised machine learning algorithm that solves problems of assigning labels to information where labels are assigned from finite element sets [7, 19, 35].

Author Proof

SVM was introduced by Vapnik. It is a valuable technique for solving classification problems. Compared to other classifiers, it has prominent functions such as local minima absence, high simplification competence, and fitness for dataset samples. Given a dataset A = {(xi, yi)|xi ϵ yi ϵ {−1, +1}, i = 1, . . . N}, where x_i is a sample of the d-dimension, yi is the corresponding class name, and N is the number of samples. The discriminant functions of the SVM takes:

$$F(x) = \text{sgn}\left(\sum_{i=1}^{av} \alpha_i y_i C(x, x_i) + b\right) \tag{2}$$

In Eq. 1, *av* is the support vector numbers, α_i is the langrange multipliers, b is the optimum bias hyperplane classification, while C(x, x_i) is the kernel functions. SVM-RBF was used as the SVM classification kernel for this study. AQ3

$$C(x_i, x_j) = \exp\{-\frac{|x_i - x_j|^2}{2\sigma^2} \tag{3}$$

SVMs uses the feature space hyperplane that is optimal to maximum marginal principles, to make decision boundaries. The kernel functions change the shapes of the hyperplane into non-linear, with two hyperplane parameters which are the parameter regularization C and the kernel width parameter σ. The training vectors are mapped in higher dimensional space function. In predicting the binary classification, the discriminant between objects are one of two groups known as the positive or the or negative. The SVM concepts solves problem by using the huge margin separation and kernel purposes.

An SVM maps input into high-dimensional feature space by finding separating hyperplanes which exploits the margin classes. Lacking somewhat knowledge of mapping, SVM discovers the optimal hyperplane using point product functions in feature space known as the kernels. The optimal hyperplane solution is a grouping of few input points [35]. SVMs are kernel methods of algorithms that are contingent on a given data through dot-products. The dot product is replaced using kernel functions for computing the dot product in high dimensional feature space. SVM comprises of four major kernels known as the; Linear, Quadratic, Polynomial and the Radial Basis Function (RBF).

Selecting suitable kernel function is significant, as kernel function describes the feature space that training set samples are classified. In this study, SVM-RBF is adopted for its high accuracy potentials and its capability to pact with gene expression high-dimensional data.

3.7 Evaluation

Performance evaluation of this experiment, 10-fold cross-validation was used, evaluate the performances. The dataset was reduced using chi-square, a classifier was trained and tested. Measure the performance of the SVM classification algorithm, and evaluation metrics use Accuracy, Precision, and Recall for evaluating the classification [36–39]. This study proposes a Chi-Square feature selection algorithm in selecting relevant features in high-dimensionality customer retention data. An SVM classifier is used as an algorithm to evaluate the performance of the experiment. This study uses a Taiwan bank dataset

Author Proof

with 30000 observations and 25 attributes. MATLAB tool is used to analyze and carry out the experiment for this study and achieved a credit scoring investigational results.

In this study the classification confusion matrix is used for the performance evaluation of the model. It comprises of the correctly and incorrectly classified values with the precise test data results. It comprises of the:

True Positive (TP): results that are correctly predicted positively in the class model.
True Negative (TN): results that are correctly predicted negatively in the class model.
False Positive (FP): results that are incorrectly predicted positive in the class model.
False Negative: results that are incorrectly predicted negative in the class model.

There are several measures used in the confusion matrix to determine the performance of the models such as:

Accuracy: this measures the bias, using the formula (TP + TN)/(TP + TN + FP + FN).
Sensitivity: this is called a recall; it is the true positive rate which is the proportion of the actual positives correctly identified as positive. TP/(TP + FN).
Specificity: this is the selectivity of the true negative rate, measuring the proportion of the actual negatives correctly identified as negatives. TN/(FP + TN).
Precision: this is the prositive predictive value. TP/(TP + FP).
F1 Score: this is the measure of the tested accuracy that harmonizes the mean of precision and the recall. 2TP/(2TP + FP + FN).

4 Results and Discussions

In this study Chi-Square was used on the Taiwan Bank Credit data to fetch out relevant information in the data, 11088 features were achieved as a subset and passed into the classification using the SVM-RBF algorithm, a confusion matrix was used as the predictive performance evaluation. Figure 2 shows the confusion matrix used to evaluate the performance of the experiment. It comprises of the true positive and negative, false positive and negative, which is a standard for finding the performance metrics.

The performance evaluation metrics of Chi-Square and SVM-RBF is calculated as follows:

Correctly Classified Instances (accuracy) $= \frac{TP+TN}{TP+TN+FP+FN} = \frac{4+10}{4+10+0+1} * 100 = 93.33\%$
Sensitivity (Recall) $= \frac{TP}{TP+FN} = \frac{4}{4+1} = 0.8 = 80\%$
Specificity $= \frac{TN}{FP+TN} = \frac{10}{0+10} = 1 = 100\%$
Precision $= \frac{TP}{TP+FP} = \frac{4}{4+0} = 1 = 100\%$
F-measure $= \frac{2*Recall*Precision}{Recall+Precision} = \frac{2*0.8*1}{0.8+1} = 0.89 = 89\%$

Author Proof

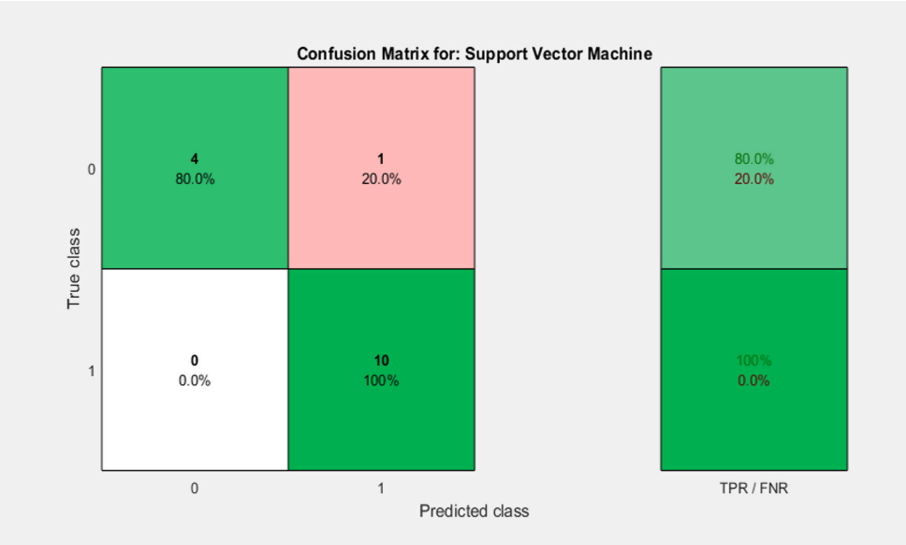


Fig. 2. Confusion matrix for chi-square + RBF-SVM classification TP = 4; TN = 10; FP = 1, FN = 0

In this study, the evaluation of the experiment is calculated using the confusion matrix formulas. Table 1 shows the performance evaluation summary.

Table 1. Predictive performance summary for chi-square with RBF-SVM.

Performance metrics (%)	Experimental results
Accuracy	93.33
Sensitivity	80
Specificity	100
Precision	89
F-measure	*89*

In this study, a feature selection and classification approaches were carried out on a Taiwan Bank credit dataset, using RFE feature selection algorithms to fetch out relevant information on the given data, for customer credit risk management, SVM was used as a classification, the result shows that RFE outperforms in terms of accuracy, the results are tabulated in Table 1 above. However, this study has shown a relevant output result that can be proposed for banks on loan concerns.

5 Conclusion

From this study, feature selection method using Chi-Square and SVM-RBF classification was used to develop a predictive model in a financial industry using Taiwan dataset to

Author Proof

predict and classify loans applicant who is eligible or not eligible for loan collection. The model was implemented using MATLAB application. The result shows that RBF and SVM-RBF accuracy was 93% shown in the predictive performance table. This study confines with the prediction of credit loan. Future research proposes to analyze the retention strategies by choosing fitting variables from the dataset.

References

1. Sudhamathy, G.: Credit risk analysis and prediction modelling of bank loans using R. Int. J. Eng. Technol. **8**(5), 1954–1966 (2016)
2. Arutjothi, G., Senthamarai, C.: Prediction of loan status in commercial bank using machine learning classifier. In: International Conference on Intelligent Sustainable Systems. IEEE (2017)
3. Chitra, K., Subashini, B.: Data mining techniques and its applications in banking sector. Int. J. Emerg. Technol. Adv. Eng. **3**(8), 219–226 (2013)
4. Ogundokun, R.O., Awotunde, J.B.: Machine learning prediction for COVID 19 pandemic in India. medRxiv (2020)
5. Adebiyi, M.O., Ogundokun, R.O., Abokhai, A.A.: Machine learning-based predictive farmland optimization and crop monitoring system. Scientifica **2020** (2020). Article ID AQ4 9428281
6. Adegun, A.A., Ogundokun, R.O., Adebiyi, M.O., Asani, E.O.: CAD-based machine learning project for reducing human-factor-related errors in medical image analysis. In: Handbook of Research on the Role of Human Factors in IT Project Management, pp. 164–172. IGI Global (2020)
7. Ibukun, E., Ayodele Adebiyi, A., Ambrose, A., Okesola, O.: Fraud prediction in bank credit administration: a systematic literature review. J. Theor. Appl. Inf. Technol. **97**(11), 3147–3169 AQ5 (2019)
8. Verma, H., Nandal, R.: A review paper on prediction analysis: acceptance of banking. J. Emerg. Technol. Innov. Res. **5**(1), 746–748 (2018)
9. Ha, V.-S., Nguyen, H.-N.: Credit scoring with a feature selection approach based deep learning. In: MATEC Web of Conferences, vol. 54 (2016)
10. Abolade, R.O., Famakinde, S.O., Popoola, S.I., Oseni, O.F., Atayero, A.A., Misra, S.: Support vector machine for path loss predictions in urban environment. In: Gervasi, O. (ed.) Computational Science and Its Applications – ICCSA 2020: 20th International Conference, Cagliari, Italy, July 1–4, 2020, Proceedings, Part VII, pp. 995–1006. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58820-5_71
11. Behera, R.K., Rath, S.K., Misra, S., Leon, M., Adewumi, A.: Machine learning approach for reliability assessment of open source software. In: Misra, S. (ed.) Computational Science and Its Applications – ICCSA 2019: 19th International Conference, Saint Petersburg, Russia, July 1–4, 2019, Proceedings, Part IV, pp. 472–482. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24305-0_35
12. Ying, L.: Research on bank credit default prediction based on data mining algorithm. Int. J. AQ6 Soc. Sci. Humanit. Invention **5**(6), 4820–4823 (2018)
13. Oladele, T.O., Ogundokun, R.O., Kayode, A.A., Adegun, A.A., Adebiyi, M.O.: Application of data mining algorithms for feature selection and prediction of diabetic retinopathy. In: Misra, S. (ed.) Computational Science and Its Applications – ICCSA 2019: 19th International Conference, Saint Petersburg, Russia, July 1–4, 2019, Proceedings, Part V, pp. 716–730. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24308-1_56

Author Proof

14. Adebiyi, M.O., Falola, D., Olatunji, O., Ogundokun, R.O., Adeniyi, E.A.: Computational prediction of corynebacterium matruchotii protein's 3D structure reveals its capacity to bind AQ7 to DNA domain site in the malaria vector, Anopheles
15. Omolewa, O.T., Oladele, A.T., Adeyinka, A.A., Oluwaseun, O.R.: Prediction of student's academic performance using k-means clustering and multiple linear regressions. J. Eng. Appl. Sci. **14**(22), 8254–8260 (2019)
16. Ogwueleka, F.N., Misra, S., Colomo-Palacios, R., Fernandez, L.: Neural network and classification approach in identifying customer behavior in the banking sector: a case study of an international bank. Hum. Factors Ergon. Manuf. Serv. Ind. **25**(1), 28–42 (2015)
17. Assef, F.M., Steiner, M.T.A.: Ten-year evolution on credit risk research: a systemic literature review approach and discussion. Ing. E Invest. **40**(2), 1–21 (2020)
18. Zhang, L., Hu, H., Zhang, D.: A credit risk assessment model based on SVM for small and medium enterprises in supply chain finance. Financial Innov. **1**(14) (2015)
19. Khemakhem, S., Boujelbene, Y.: Artificial intelligence for credit risk assessment: artificial neural network and support vector machines. ACRN Oxford J. Finance Risk Perspect. **6**(2), 1–7 (2017)
20. Jency, X.F., Sumathi, V.P., Janani, S.S.: An exploratory data analysis for loan prediction based on nature of the clients. Int. J. Recent Technol. Eng. **7**(4), 176–179 (2018)
21. Goh, R.Y., Lee, L.S.: Credit scoring: a review on support vector machines and metaheuristic AQ8 approaches. Adv. Oper. Res. 1–30 (2019)
22. Li, S., Wang, M., He, J.: Prediction of banking systemic risk based on support vector machine. Math. Prob. Eng. 1–5 (2013)
23. Sanz, H., Valim, C., Vegas, E., Oller, J.M., Reverter, F.: SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. BMC Bioinform. **19**(432) (2018)
24. Adebiyi, M.O., Ogundokun, R.O., Abokhai, A.A.: Machine learning–based predictive farmland optimization and crop monitoring system. Scientifica **2020** (2020)
25. Aktan, S.: Application of machine learning algorithms for business failure prediction. Investment Manag. Financ. Innov. **8**(2), 52–65 (2011)
26. https://www.kaggle.com/uciml/default-of-credit-card-clients-dataset
27. Suleiman, S., Burodo, M.S., Suleman, I.: Credit scoring using principal component analysis based binary logistic regression. J. Sci. Eng. Res. **412**, 99–110 (2018)
28. Boateng, E.Y., Oduro, F.T.: Predicting microfinance credit default: a study of Nsoatreman rural bank, Ghana. J. Adv. Math. Comput. Sci. **26**(1), 1–9 (2018)
29. Mavetera, N., Moroke, N.D., Chibonda, C.: A chi-square application on the factors influencing internet banking adoption and usage in Botswana. Mediterr. J. Soc. Sci. **5**(20), 596–606 (2014)
30. Adebiyi, M.O., Adigun, E.B., Ogundokun, R.O., Adeniyi, A.E., Ayegba, P., Oladipupo, O.O.: Semantics-based clustering approach for similar research area detection. Telkomnika **18**(4), 1874–1883 (2020)
31. Ogundokun, R.O., et al.: Evaluation of the scholastic performance of students in 12 programs from a private university in the south-west geopolitical zone in Nigeria. F1000Research **8** 154 (2019)
32. Awotunde, J.B., Ogundokun, R.O., Ayo, F.E., Ajamu, G.J., Adeniyi, E.A., Ogundokun, E.O.: Social media acceptance and use among university students for learning purpose using UTAUT model. In: Borzemski, L., Świątek, J., Wilimowska, Z. (eds.) Information Systems Architecture and Technology: Proceedings of 40th Anniversary International Conference on Information Systems Architecture and Technology – ISAT 2019: Part I, pp. 91–102. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-30440-9_10
33. Rana, R., Singhal, R.: Chi-square test and its application in hypothesis testing. J. Pract. Cardiovasc. Sci. **1**(1), 69–71 (2015)
34. Mashhour, E.M., Houby, E.M.F., Wassif, K.T., Salah, A.I.: Feature selection approach based on firefly algorithm and chi-square. Int. J. Electr. Comput. Eng. **8**(4), 2338–2350 (2018)

Author Proof

35. Ikram, S.T., Cheukuri, A.K.: Intrusion detection model using fusion of chi-square feature selection and multi-class SVM. J. King Saud Univ. Comput. Inf. Sci. **29**(4), 462–472 (2017)
36. A-Harbi, O.: A comparative study of feature selection model methods for dialectal Arabic sentiment classification using support vector machine. Int. J. Comput. Sci. Netw. Secur. **19**(1), 167–176 (2019)
37. Osho, O., Mohammed, U.L., Nimzing, N.N., Uduimoh, A.A., Misra, S.: Forensic analysis of mobile banking apps. In: Misra, S. (ed.) Computational Science and Its Applications – ICCSA 2019: 19th International Conference, Saint Petersburg, Russia, July 1–4, 2019, Proceedings, Part V, pp. 613–626. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24308-1_49
38. Ogwueleka, F.N., et al.: Neural network and classification approach in identifying customer behavior in the banking sector: a case study of an international bank. Hum. Factors Ergon. Manuf. Serv. Ind. **25**(1), 28–42 (2015)
39. Arogundade, O.T., Abayomi-Alli, A., Misra, S.: An ontology-based security risk management model for information systems. Arab. J. Sci. Eng. **45**(8), 6183–6198 (2020). https://doi.org/10.1007/s13369-020-04524-4

Author Proof

Author Queries

Chapter 13

Query Refs.	Details Required	Author's response
AQ1	This is to inform you that corresponding author has been identified as per the information available in the Copyright form.	
AQ2	Please check and confirm if the authors given and family names have been correctly identified in author group.	
AQ3	The opening parenthesis does not have a corresponding closing parenthesis in "Equation 3". Please insert the parenthesis in the appropriate position.	
AQ4	Kindly provide the page range for Refs. [5, 18, 23, 24].	
AQ5	Please check the list of authors in Ref. [7].	
AQ6	Reference [12] is given in list but not cited in text. Please cite in text or delete from list.	
AQ7	Kindly provide year of the publication for the Ref. [14].	
AQ8	Kindly provide the volume id for Refs. [21, 22].	