JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

CPRAMLCT: Design of a model for Credit Prediction and Risk analysis on Loans Data using Machine Learning Classification Techniques

Mahesh Rathod
ME Student, Department of Computer Science & Engineering, PRMIT&R
Badnera-444701, Maharashtra, India

Dr V.M. Deshmukh
Asso. Professor, Department of Computer Science & Engineering, PRMIT&R
Badnera-444701, Maharashtra, India

Abstract: Credit Prediction requires analysis of multiple user-level features including their previous transactions, spending patterns, billing patterns, previous line of credit status, etc. To analyse the risk associated with credit sanction, various Machine Learning Models (MLMs) are proposed by researchers, but very few of them are capable of confidently estimating true credit limits for different users. Thus, banking firms & financial advisory firms are not able to confidently recommend credit amounts on a per user basis. To overcome these limitations, this paper discusses design of a novel Credit prediction model that can perform Risk Minimization via Ensemble Learning process. The proposed model uses a large-scale dataset that combines user's personal details like Education Level, Gender, Marital Status, and Age with details about previous spending & payment patterns to estimate credit default probability, along with maximum amount that can be guaranteed with minimum risk & maximum retention rates. The model uses combination of Naïve Bayes (NB), Support Vector Machine (SVM), Deep Random Forest (DRF), Multilayer Perceptron based Neural Network (MLP), & Logistic Regression (LR) to predict user's credit scores. These scores are combined with a correlation-based model which assists in estimation of maximum amount of credit that can be extended to the customer with confidence of higher retention rates. Due to a combination of these methods, the proposed CPRAMLCT Model is capable of high accuracy, high precision, better recall, and low delay operations. The classification models were combined using a unique union-based 'mode' method, which assisted in identification of common & correctly classified samples. This reduces classification errors, which assists in improving classification accuracy by 6.5%, recommendation precision by 15.5%, recommendation recall by 8.5%, and recommendation speed by 14.5% under different real-time scenarios. Due to these performance enhancements, the model is useful for a wide variety of applications.

Keywords: Credit Prediction, Risk Analysis, Tracking, Recommendation, Ensemble, NB, SVM, LR, MLP, DRF, Accuracy

[1] Introduction

Credit Prediction and Risk analysis is a multi-domain task that requires efficient modelling of blocks that include, data collection on a per user basis, pre-processing of data for outlier detection, segmentation of data for extraction of useful information, feature extraction & selection for efficient representation of data, classification into valid, outlier, & other types, and post-processing based on application-specific operations. A typical credit score analysis model [1] that uses Machine Learning Models (MLMs) for identification of default probability is depicted in figure 1, wherein different payment behaviour systems along with credit risk calculation, credit risk management, etc. are observed. The model also uses multiple processing feedbacks that include developmental & accuracy feedback for continuous optimization of model performance. These optimizations include adjustment of payment terms, optimization of managerial services, etc. Based on these optimizations, the model is able to produce better results for future evaluations.

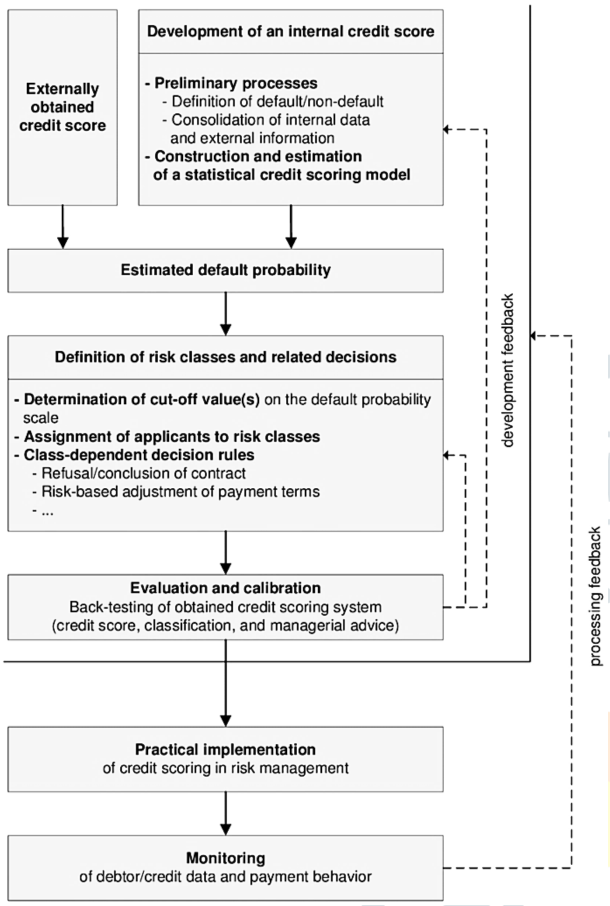

Figure 1. A typical credit score management model with continuous feedback operations

The model is also able to identify different credit score levels, which assist in estimating outlier requests. Similar models that use different ML techniques [2], [3], [4] along with their characteristics are discussed in the next section of this text. Based on this discussion, readers will be able to identify various context-level nuances, application-level advantages, functional limitations, and deployment-level future research scopes. This will assist in identification of optimum model sets of their application-specific & performance-specific use cases. But most of these models do not produce a confidence score, and do not consider multiple previous transaction records while estimating credit scores. Due to this limitation, they cannot be used for large-scale credit prediction applications. To overcome this limitation, section 3 proposes design of a novel Credit Prediction and Risk analysis model with Risk Minimization via Ensemble Learning process. The model uses a combination of multiple high-efficiency classifiers, which assist in classifying input credit requests into 'Valid' or 'Default' categories. The model is cascaded with a correlation analysis method, which assists in estimating maximum amount of credit that can be extended to the user with minimum risk levels. Performance of this model was evaluated in section 4, in terms of accuracy, recall, precision and delay metrics. This performance was compared with various reviewed models, which assisted in observing superiority of the model w.r.t. other methods. Finally, this text is concluded with some application-specific observations about the proposed model, and also recommends various optimizations, to further improve its risk evaluation performance under real-time scenarios.

[2] Literature Review

A large number of research work has been done in the field of credit prediction and analysis, which assists banking institutions to optimize their decision taking capabilities. These models vary in terms of their deployment characteristics & real-time performance. For instance, work in [5, 6] proposes use of invoice data with different mining techniques, which assists in estimation of simplistic transactional patterns. This work doesn't provide comprehensive insights into these transactions, due to which it has minimum applicability. To improve this performance, work in [7] proposes use of supervised mining models, which can be trained as per context-sensitive requirements. Due to use of supervised learning, the models are able to continuously optimize classification performance under different use cases. These models are further extended via the work in [8, 9] which discusses use of Logistic Regression with LightGBM (LR LGBM), and SVMs for high efficiency classification process. These models are useful for unbalanced datasets, which makes them highly applicable for real-time applications. Similar models are discussed in [10, 11, 12], which propose use of Online Integrated Credit Scoring Model (OICSM), Fusion Neural Networks (FNNs), and Cost-Sensitive Neural Network Ensemble (CS NNE) that are capable of augmenting data features for high-accuracy classification applications. These models are useful for low error operations, but require larger classification delays. To improve the speed performance, work in [13, 14, 15] proposes use of Slow and Fast Learning (SFL), LightGBM, and bioinspired optimization techniques, which assist in improving credit score evaluation speed via high efficiency feature reduction processes.

Work in [16, 17, 18] further proposes use of Spy Model with Transfer Learning and Adaboost optimizations (SPY TRA), Multiple linear Clusters for Feature Selection, and use of trajectory datasets for effective classification of different datasets. These models are highly useful in analysis of multidimensional datasets, which assists in improving their deployment capabilities. Work in [19] further adds privacy to these models which assists in securely mitigating user credit requests for high performance applicability. But most of these models are highly linear in their processing capabilities, and also do not perform inter-user correlative analysis. To overcome these issues, next section proposes use of a novel model for Credit Score estimation with Risk Minimization via Ensemble Learning process. The model was evaluated on multiple datasets, and its performance was compared with various state-of-the-art methods under different scenarios.

[3] Design of the proposed model for Credit Prediction and Risk analysis via Ensemble Learning process

As per the literature review, it can be observed that existing models for estimation of credit scores are highly linear, and do not perform inter-user correlative analysis, which limits their applicability. To overcome this limitation, a novel model for Credit Prediction and Risk analysis via Ensemble Learning process is discussed in this section of the text. Flow of the proposed model is depicted in figure 2, wherein different input parameters including user's Age, Marital Status, Gender, Status of previous payments, Transactional amounts for previous bills, and Transactional amounts for previous bill payments are aggregated, and classified via different Machine Learning Models (MLMs). The outputs of these classifiers are combined to form an ensemble learning model, which assists in identification of current user request status. Based on this status, a correlation model is selectively activated, which results in a credit score that assists banks & other financial firms to grant credit loans to requesting users. The credit levels are accompanied with a credit score, which assists investment firms to optimize their credit decisions.

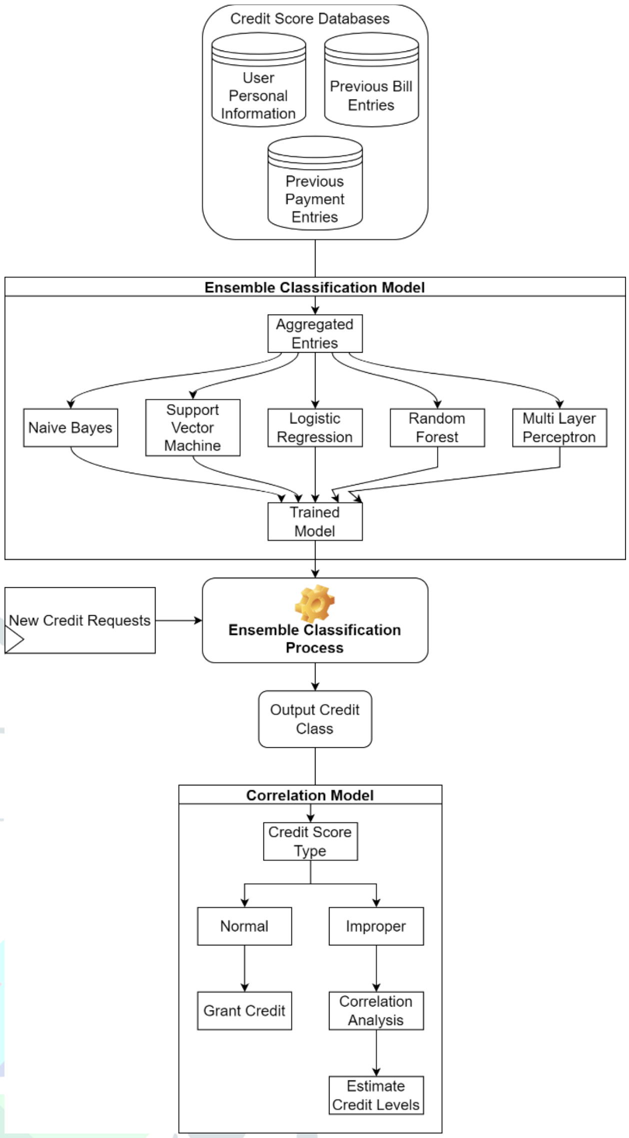

Figure 2. Overall flow of the proposed model

From the flow of model, it can be observed that input datasets are collected for user's personal information, their spending patterns, and their payment patterns. These patterns are used to train multiple ML classification techniques that include Naïve Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), and Multilayer Perceptron (MLP) for efficient classification performance. Each of these models are trained using a specific set of hyperparameters, which were selected via manual tuning process to obtain high accuracy levels. These parameters for each of the classifiers can be observed from table 1 as follows,

Classifier	Hyperparameter	Value of the parameter
NB	Smoothing Factor	10^{-10}
SVM	Error Tolerance	10^{-5}
LR	Maximum Iterations	1000
LR	Solver	Limited Memory Broyden–Fletcher–Goldfarb–Shanno
RF	Number of Estimators	100
RF	Maximum Depth	2
MLP	Solver	Stochastic Gradient Descent
MLP	Hidden Layer Size	2
MLP	Maximum Neurons Per Layer	100
MLP	Learning Rate	Adaptively Optimized

Table 1. Parameters for different classifiers

Based on these parameters, the models were trained, and their classification responses were aggregated. These responses were processed via a mode operation, which assists in estimation of frequently occurring classes. Based on this estimation, the model was able to identify credit types for new input credit requests. If the credit type is classified as 'Normal', then requested amount is approved, otherwise a correlation model is activated for estimation of maximum line of credit, that can be extended to requesting users. This correlation is calculated via equation 1 as follows,

$$C(Match) = \frac{\sum_{i=1}^{N(DB)} \left(CF_i - \sum_{j=1}^{N(DB)} \frac{CF_j}{N(DB)} \right) * \left(CF(New) - \sum_{j=1}^{N(New)} \frac{CF_j(New)}{N(New)} \right)}{\sqrt{\sum_{i=1}^{N(DB)} \left(CF_i - \sum_{j=1}^{N(DB)} \frac{CF_j}{N(DB)} \right)^2 * \sum_{i=1}^{N(New)} \left(CF(New) - \sum_{j=1}^{N(New)} \frac{CF_j(New)}{N(New)} \right)^2}} \ldots (1)$$

Where, CF represents credit features, while $N(DB)$ & $N(New)$ represents number of entries in the database, and number of entries to be processed for requesting users. This score is estimated for every database entry, and a Maximum Value Correlation Score (MVCS) is estimated via equation 2,

$$MVCS = \bigcup_{i=1}^{N(DB)} C(Match)_i \ldots (2)$$

Based on the value of MVCS, the highest matching entry is selected from the database, and its credit levels are checked. If these credit levels satisfy equation 3, then user is granted the credit amount, else, a new credit amount is calculated based on equation 4,

$$CL > \frac{CL(MVCS)}{2} \ldots (3)$$

Where, CL represents Credit Levels requested by user, while $CL(MVCS)$ represents Credit Levels obtained from the database entries.

$$A(New) = CL(MVCS) * C(Match) \ldots (4)$$

Where, $A(New)$ is the new credit amount which is accepted for the user requests. Based on this evaluation, users are granted credit levels, which assists financial firms to improve their credit risk management performance. This performance was evaluated in terms of accuracy, delay, recall & precision measures, and discussed in the next section of this text.

[4] Result analysis & comparison

The proposed model uses an aggregation of different classifiers, and combines them with a correlative analysis technique for estimation of true credit scores. This correlative model assists in identification of optimum credit amounts for requesting users. Performance of this model was evaluated using multiple datasets, which were collected from https://www.listendata.com/2019/08/datasets-for-credit-risk-modeling.html, https://github.com/JLZml/Credit-Scoring-Data-Sets, and https://www.kaggle.com/c/GiveMeSomeCredit sources. All of these sources were combined to form a total of 16k financial records which cover credit scores for different user types. The collected dataset was divided in a ratio of 60:20:20, where 60% entries were used to train the classifiers, 20% were used to test the classifiers, and remaining 20% were used to validate the classifiers. Using this strategy, the model was evaluated & values for accuracy of recommendation (A_r), precision of recommendation (P_r), recall of recommendation (R_r), and computational delay needed for recommendation (D_r) were evaluated for all these sets. The performance was compared with FNN [11], CS NNE [12], & SPY TRA, which assists in comparative evaluations. The parameters were compared w.r.t. Number of Testing (NT) entries, and can be observed from table 2, where accuracy of recognition was evaluated for different test set entries.

NT	A_R **FNN [11]**	A_R **CS NNE [12]**	A_R **SPY TRA**	A_R **CSR MEL**
50	85.80	80.56	79.20	87.71
100	86.50	82.14	80.28	88.91
150	87.37	83.71	81.45	90.19
200	88.24	85.28	82.62	91.47
250	88.76	86.06	83.25	92.15
300	89.11	85.85	83.32	92.21
350	89.46	85.38	83.26	92.16
400	89.81	85.39	83.44	92.37
450	89.79	86.07	83.75	92.73
500	89.47	86.64	83.88	92.87
600	89.28	87.10	84.00	93.00
700	89.38	87.38	84.17	93.19
800	89.71	87.57	84.42	93.47
900	90.01	87.82	84.68	93.77
1000	90.26	88.17	84.98	94.08
1100	90.45	88.57	85.26	94.38
1200	90.60	88.92	85.49	94.63
1300	90.74	89.23	85.69	94.88
1400	90.87	89.53	85.90	95.11
1500	90.97	89.84	86.10	95.32
1600	91.10	90.19	86.31	95.54
1800	91.24	90.54	86.54	95.80
2000	91.40	90.89	86.78	96.08
2200	91.57	91.24	87.03	96.36
2400	91.74	91.58	87.28	96.80
2600	91.89	91.92	87.52	97.41
2800	92.03	92.25	87.76	98.02
2900	92.17	92.60	87.99	98.45
3000	92.32	92.95	88.21	98.80
3100	92.62	93.43	88.57	99.06
3200	92.92	93.92	88.96	99.33

Table 2. Accuracy of credit prediction recommendation for different models

After the detailed observations done in table 2, it can be estimated that the proposed model is capable of achieving 6.5% higher accuracy than FNN [11], 4.9% higher accuracy than CS NNE [12], and 10.5% higher accuracy than SPY TRA under different evaluations. This is possible due to use of ensemble classification & correlation analysis, which makes the model useful for high-accuracy credit prediction and analysis applications. Based on the same datasets, precision of credit score analysis was calculated, and tabulated in table 3 as follows,

NT	P_R **FNN [11]**	P_R **CS NNE [12]**	P_R **SPY TRA**	P_R **CSR MEL**
50	84.05	79.37	77.79	86.16
100	84.81	80.90	78.89	87.37
150	85.66	82.43	80.03	88.61
200	86.34	83.58	80.91	89.56
250	86.76	83.86	81.25	89.93
300	87.10	83.52	81.26	89.94
350	87.45	83.31	81.32	90.02
400	87.61	83.64	81.56	90.30
450	87.45	84.26	81.77	90.54
500	87.20	84.75	81.89	90.67
600	87.15	85.11	82.03	90.82
700	87.37	85.34	82.24	91.05
800	87.67	85.56	82.49	91.34
900	87.94	85.85	82.76	91.63
1000	88.15	86.22	83.04	91.93
1100	88.32	86.58	83.29	92.20
1200	88.47	86.90	83.50	92.45
1300	88.59	87.20	83.70	92.67
1400	88.70	87.50	83.90	92.89
1500	88.81	87.82	84.11	93.11
1600	88.94	88.16	84.32	93.35
1800	89.09	88.50	84.55	93.61
2000	89.25	88.84	84.79	93.88
2200	89.42	89.18	85.03	94.22
2400	89.58	89.51	85.27	94.73
2600	89.72	89.83	85.50	95.33
2800	89.86	90.17	85.73	95.84
2900	90.00	90.51	85.95	96.18
3000	90.51	91.21	86.50	96.91
3100	91.53	92.39	87.55	97.92
3200	92.56	93.58	88.62	98.95

Table 3. Precision of credit prediction recommendation for different models

After the detailed observations done in table 4, it can be estimated that the proposed model is capable of achieving 5.9% higher precision than FNN [11], 4.5% higher precision than CS NNE [12], and 9.5% higher precision than SPY TRA under different evaluations. This is possible due to use of correlation analysis with ensemble classification process, which makes the model useful for high-precision credit analysis applications. Based on the same datasets, recall of credit score analysis was calculated, and tabulated in table 4 as follows,

NT	R_R FNN [11]	R_R CS NNE [12]	R_R SPY TRA	R_R CSR MEL
50	86.14	83.44	77.79	88.31
100	86.93	85.05	78.89	89.55
150	87.80	86.66	80.03	90.83
200	88.50	87.86	80.91	91.81
250	88.94	88.16	81.25	92.18
300	89.29	87.81	81.26	92.19
350	89.63	87.57	81.32	92.26
400	89.80	87.93	81.56	92.55
450	89.63	88.57	81.77	92.80
500	89.38	89.10	81.89	92.93
600	89.34	89.48	82.03	93.10
700	89.55	89.72	82.24	93.34
800	89.86	89.94	82.49	93.62
900	90.14	90.25	82.76	93.92
1000	90.36	90.63	83.04	94.23
1100	90.53	91.01	83.29	94.51
1200	90.68	91.36	83.50	94.76
1300	90.81	91.67	83.70	94.99
1400	90.92	91.98	83.90	95.22
1500	91.04	92.32	84.11	95.44
1600	91.17	92.68	84.32	95.68
1800	91.32	93.04	84.55	95.94
2000	91.49	93.40	84.79	96.22
2200	91.66	93.75	85.03	96.58
2400	91.82	94.10	85.27	97.10
2600	91.97	94.44	85.50	97.71
2800	92.11	94.79	85.73	98.24
2900	92.26	95.15	85.95	98.59
3000	92.40	95.09	86.50	98.91
3100	92.66	94.72	87.55	99.12
3200	92.94	94.36	88.62	99.35

Table 4. Recall of credit prediction for different models

After the detailed observations done in table 4, it can be estimated that the proposed model is capable of achieving 6.5% higher recall than FNN [11], 5.5% higher recall than CS NNE [12], and 10.5% higher recall than SPY TRA under different evaluations. This is possible due to use of correlation analysis with ensemble classification process, which makes the model useful for high-recall credit analysis applications. Based on the same datasets, delay needed for credit score analysis was calculated, and tabulated in table 5 as follows,

NT	*D* (ms) FNN [11]	*D* (ms) CS NNE [12]	*D* (ms) SPY TRA	*D* (ms) CSR MEL
50	3.82	4.83	4.37	2.99
100	3.79	4.73	4.31	2.95
150	3.75	4.65	4.25	2.90
200	3.71	4.57	4.19	2.87

250	3.69	4.54	4.17	2.85
300	3.68	4.55	4.16	2.85
350	3.66	4.57	4.16	2.84
400	3.65	4.56	4.15	2.83
450	3.65	4.53	4.14	2.82
500	3.66	4.50	4.14	2.82
600	3.67	4.48	4.13	2.82
700	3.67	4.46	4.12	2.81
800	3.65	4.45	4.11	2.81
900	3.64	4.44	4.09	2.80
1000	3.63	4.43	4.08	2.79
1100	3.62	4.41	4.07	2.78
1200	3.62	4.39	4.06	2.77
1300	3.61	4.37	4.05	2.76
1400	3.61	4.35	4.04	2.76
1500	3.60	4.34	4.03	2.75
1600	3.60	4.32	4.02	2.74
1800	3.59	4.31	4.01	2.74
2000	3.59	4.29	4.00	2.73
2200	3.58	4.27	3.99	2.72
2400	3.57	4.26	3.98	2.71
2600	3.57	4.24	3.97	2.69
2800	3.56	4.22	3.96	2.67
2900	3.55	4.21	3.95	2.66
3000	3.56	4.22	3.97	2.68
3100	3.58	4.27	4.04	2.74
3200	3.61	4.34	4.14	2.83

Table 4. Delay performance of credit prediction recommendation for different models

After the detailed observations done in table 5, it can be estimated that the proposed model is capable of achieving 10.5% lower delay than FNN [11], 18.3% lower delay than CS NNE [12], and 15.5% lower delay than SPY TRA under different evaluations. This is possible due to use of low complexity classification process, which makes the model useful for high-speed credit analysis applications. Thus, the model outperforms most of the recently proposed methods, thus making it applicable for real-time credit score analysis use cases.

[5] Conclusion and future scope

The proposed CPRAMLCT model is capable of reducing computational delay, while improving accuracy performance when compared with various state-of-the-art models. This is possible due to use of simplistic classification process, which allows the model to be trained & evaluated faster with better performance. This performance is further optimized via use of a correlative analysis technique, which assists in estimation of low error, and high efficiency credit scores that can be used by banking agencies to extend line of credits for different users. Due to these enhancements, the model is able to achieve 6.5% higher accuracy than FNN [11], 4.9% higher accuracy than CS NNE [12], and 10.5% higher accuracy than SPY TRA, it was also able to achieve 5.9% higher precision than FNN [11], 4.5% higher precision than CS NNE [12], 9.5% higher precision than SPY TRA, and 6.5% higher recall than FNN [11], 5.5% higher recall than CS NNE [12], and 10.5% higher recall than SPY TRA under different evaluations. This was possible due to combination of classification performance from different classifier sets. The model was also able to achieve 10.5% lower delay than FNN [11], 18.3% lower delay than CS NNE [12], and 15.5% lower delay than SPY TRA under different evaluations, which makes it useful for high-efficiency and high-speed classification applications. In future, the model's performance can be improved via use of Convolutional Neural Networks (CNNs), Q-Learning, Reinforcement Learning, and other Deep

Learning frameworks. Furthermore, the model must also be validated for larger sets, which will assist in further improving its scalability for different use cases.

[6] References

[1] N. A. Nguyen-Thi, Q. Pham-Nhu, S. Luong-Ngoc and N. Vu-Thanh, "Using novel noise filtering techniques to ensemble classifiers for credit scoring: Using noise filtering techniques to ensemble classifiers for credit scoring," 2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI), 2019, pp. 1-6, doi: 10.1109/TAAI48200.2019.8959831.

[2] A. Safiya Parvin and B. Saleena, "An Ensemble Classifier Model to Predict Credit Scoring - Comparative Analysis," 2020 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS), 2020, pp. 27-30, doi: 10.1109/iSES50453.2020.00017.

[3] K. B. Addi and N. Souissi, "An Ontology-Based Model for Credit Scoring Knowledge in Microfinance: Towards a Better Decision Making," 2020 IEEE 10th International Conference on Intelligent Systems (IS), 2020, pp. 380-385, doi: 10.1109/IS48319.2020.9199981.

[4] D. I. Calibo and M. A. Ballera, "Variable Selection for Credit Risk Scoring on Loan Performance Using Regression Analysis," 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS), 2019, pp. 746-750, doi: 10.1109/CCOMS.2019.8821664.

[5] Y. Sun, W. Jian, Y. Fu, H. Sun, Y. Zhu and Z. Bai, "A new perspective of credit scoring for small and medium-sized enterprises based on invoice data," 2021 2nd International Conference on Computing and Data Science (CDS), 2021, pp. 477-482, doi: 10.1109/CDS52072.2021.00088.

[6] A. Hassan and R. Jayousi, "Financial Services Credit Scoring System Using Data Mining," 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT), 2020, pp. 1-7, doi: 10.1109/AICT50176.2020.9368572.

[7] L. Duan, "Performance Evaluation and Practical Use of Supervised Data Mining Algorithms for Credit Card Approval," 2020 International Conference on Computing and Data Science (CDS), 2020, pp. 251-254, doi: 10.1109/CDS49703.2020.00057.

[8] B. Liu, L. Lu, Q. Zeng and Y. Li, "Implementation of credit scoring card model based on logistic regression and lightgbm," 2021 International Conference on Control Science and Electric Power Systems (CSEPS), 2021, pp. 175-178, doi: 10.1109/CSEPS53726.2021.00042.

[9] D. Tanouz, R. R. Subramanian, D. Eswar, G. V. P. Reddy, A. R. Kumar and C. V. N. M. Praneeth, "Credit Card Fraud Detection Using Machine Learning," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 967-972, doi: 10.1109/ICICCS51141.2021.9432308.

[10] Z. Zhang, K. Niu and Y. Liu, "A Deep Learning Based Online Credit Scoring Model for P2P Lending," in IEEE Access, vol. 8, pp. 177307-177317, 2020, doi: 10.1109/ACCESS.2020.3027337.

[11] T. Zhang and Y. Du, "Research on user Credit score Model based on Fusion Neural network," 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2021, pp. 1391-1395, doi: 10.1109/IAEAC50856.2021.9390672.

[12] W. Yotsawat, P. Wattuya and A. Srivihok, "A Novel Method for Credit Scoring Based on Cost-Sensitive Neural Network Ensemble," in IEEE Access, vol. 9, pp. 78521-78537, 2021, doi: 10.1109/ACCESS.2021.3083490.

[13] J. P. Barddal, F. Enembreck, L. Loezer and R. Lanzuolo, "Combining Slow and Fast Learning for Improved Credit Scoring," 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020, pp. 1149-1154, doi: 10.1109/SMC42975.2020.9283453.

[14] Z. Qiu, Y. Li, P. Ni and G. Li, "Credit Risk Scoring Analysis Based on Machine Learning Models," 2019 6th International Conference on Information Science and Control Engineering (ICISCE), 2019, pp. 220-224, doi: 10.1109/ICISCE48695.2019.00052.

[15] P. Diaconescu and V. -E. Neagoe, "Credit Scoring Using Deep Learning Driven by Optimization Algorithms," 2020 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), 2020, pp. 1-6, doi: 10.1109/ECAI50035.2020.9223139.

[16] Q. Wei, Y. Liu and K. Wu, "Transfer Learning Based Credit Scoring," 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2021, pp. 1251-1255, doi: 10.1109/CSCWD49262.2021.9437749.

[17] L. Alfat, M. Rizkinia, R. F. Sari and D. M. Romano, "Feature Selection of Credit Score Factor Based on Smartphone Usage using MCFS," 2019 4th Technology Innovation

Management and Engineering Science International Conference (TIMES-iCON), 2019, pp. 1-5, doi: 10.1109/TIMES-iCON47539.2019.9024462.

[18] W. Wang, Y. Yang, A. Chen and Z. Pan, "A Scoring Method for Driving Safety Credit Using Trajectory Data," 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), 2019, pp. 558-565, doi: 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00135.

[19] Y. Han, H. Chen, Z. Qiu, L. Luo and G. Qian, "A Complete Privacy-Preserving Credit Score System Using Blockchain and Zero Knowledge Proof," 2021 IEEE International Conference on Big Data (Big Data), 2021, pp. 3629-3636, doi: 10.1109/BigData52589.2021.9671494.