International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

A MULTI TASK APPROACH TO REJECT A WARE CREDIT SCORING WITH MISSING-NOT-AT-RANDOM DATA: RMT-NET

G. Prabhakar[1], A. Rajya Sree[2], A. Kalyani[3], V. Deekshitha[4]

[1]Assistant professor, School of CSE, Malla Reddy Engineering College For Women (Autonomous Institution), Maisammaguda, Dhulapally,Secunderabad,Telangana-500100

[234]UG Scholar, Department of CS, Malla Reddy Engineering College for Women, (Autonomous Institution), Maisammaguda, Dhulapally, Secunderabad, Telangana-500100

Email: prabhakarm.tech@gmail.com

ABSTRACT

Missing data is often an issue in the credit scoring industry, especially in cases where data is MNAR, which means the missing information relates to variables that are not observable. To solve the problem of MNAR data for credit scoring tasks, the proposed method here is an innovative approach known as RMT-NET (Rejection-based Multi-task Network). The three major tasks which RMT-NET simultaneously learns utilizing a multi-task learning architecture are credit score prediction, imputation of missing data, and rejection of untrustworthy data points. The model successfully identifies and removes biased or uncertain data by combining a rejection mechanism, thereby ensuring high accuracy and fairness of credit score forecasts. Employing one common feature representation for each of the tasks through separate branches for every task,. According to experimental results, RMT-NET performs better than traditional techniques in handling missing data and generates credit ratings that are more robust and reliable, especially for datasets containing a significant proportion of MNAR data. The proposed approach is a potential solution for real-world credit scoring applications with imperfect or faulty data because it not only improves the accuracy of predictions but also reduces bias.

Keywords-Credit Scoring, Missing Not At Random (MNAR), Multi-task Learning (MTL), RMT-NET, Data Imputation, Rejection Mechanism, Bias Reduction, Credit Risk Prediction, Machine Learning in Finance

I. INTRODUCTION

Credit scoring is a vital process in financial institutions to assess the likelihood of loan applicants defaulting on their credit. Machine learning models are employed to predict the default probabilities of customers based on their application data. When a customer applies for a loan, their application can either be approved or rejected. If approved, the customer may eventually be classified as either non-default (if they repay the loan on time) or default (if they fail to repay). However, for rejected applicants, no default information is available, creating a missing-not-at-random (MNAR) selection bias in the data. As rejected samples are not observed for default status, they are typically excluded from training models, which can lead to biased predictions.Traditional credit scoring models,

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

such as Logistic Regression (LR), Support Vector Machines (SVM), and XGBoost (XGB), rely primarily on approved samples, which leads to significant inaccuracies when making predictions for rejected applicants. While methods such as self-training and semi-supervised learning attempt to address this bias, they still fail to achieve optimal performance, particularly in predicting defaults for rejected applicants.

Recent research has shown that the default/non-default task and the rejection/approval task are highly correlated in credit scoring. Rejected applicants tend to have higher default rates, making rejection/approval a valuable signal for improving the prediction of defaults, especially for rejected applicants. To leverage this relationship, we propose the Reject-aware Multi-Task Network (RMT-Net), which employs Multi-Task Learning (MTL) to predict default status, impute missing data, and reject unreliable samples. RMT-Net uses a gating network to control how information is shared between the rejection/approval task and the default/non-default task, based on rejection probabilities.

We further extend RMT-Net to RMT-Net++, which handles multiple rejection/approval strategies, improving the model's generalization across different application scenarios. Our experiments on 10 datasets show that RMT-Net outperforms conventional classifiers by 47.9% and state-of-the-art reject inference approaches by 11.9%. RMT-Net++ provides an additional 5.8% improvement.

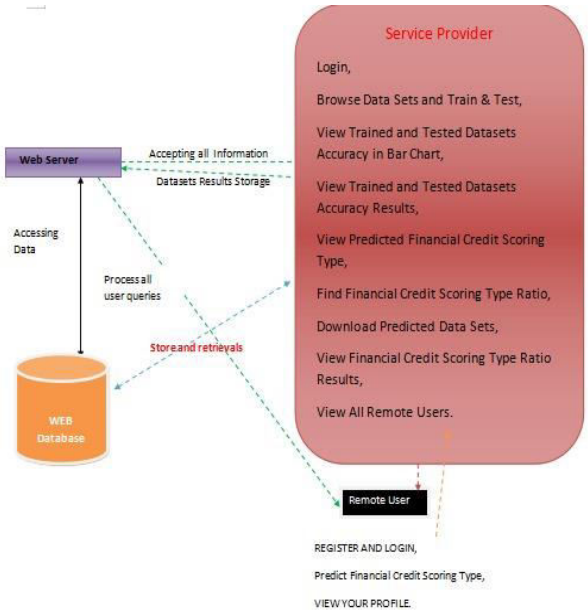

Fig 1: System Architecture

II. RELATED WORK

Reject inference in consumer credit scoring with nonignorable missing data

Authors: M. Bücker, M. van Kampen, W. Krämer

Journal: Journal of Banking & Finance, 2013

This paper addresses the challenge of reject inference in consumer credit scoring especially when the data is missing not at random. Credit scoring is biased in estimation where the future behavior of applicants who are rejected in giving loans cannot be seen; hence, biased results happen. The authors research some methods for handling non-ignorable missing data while scoring credit. In particular, they give some methods to enhance the reliability of reject inference models. With this, there is increased prediction of creditworthiness of rejected applicants.

"The economic value of reject inference in credit scoring"

Authors: G. G. Chen, T. Astebro

Institution: Department of Management Science, University of Waterloo, 2001

This paper discusses the economic benefits of using reject inference in the realm of credit scoring. Reject inference is intended to reduce missing data issues by taking the additional information available for those applicants whose applications have been rejected to enhance model estimation. This paper puts significant emphasis on the integration of reject inference in determining default probabilities for rejected applicants whose financial outcome can critically vary the outcome of lending establishments.

"Rejection of inference in application scorecards: Evidence from France"

Authors: H.-T. Nguyen, et al.

Institution: University of Paris Nanterre, EconomiX, 2016

This paper provides empirical evidence on the effectiveness of reject inference techniques in application scorecards, focusing on the French credit scoring system. Using both accepted and rejected loan applications, the authors assess several reject inference methods and their impact on improving credit scoring models. Their findings highlight that incorporating data from rejected applicants improves predictive accuracy and reduces bias in credit risk assessment.

"Can reject inference ever work?"

Authors: D. J. Hand, W. E. Henley

Journal: IMA Journal of Management Mathematics, 1993

The work, Hand and Henley, critically examine the idea of reject inference in credit scoring and ask the question of whether it could ever be successfully applied. The authors note that with the lack of observed outcomes for rejected applicants, they find it challenging to provide an accurate prediction for the same. The work hence forms a comprehensive review on the theoretical limitations and practical difficulties of applying reject inference in real-world credit scoring applications.

"Learning with a Probabilistic Teacher"

Author: A. Agrawala

Journal: IEEE Transactions on Information Theory, 1970

It brings the concept of learning from a probabilistic teacher as a fundamental idea in supervised learning. In the case of reject inference, one can relate this approach with training models that predict loan default likelihood for rejected applicants using probabilistic methods in estimating missing data. In this paper, the theory of machine learning is improved, especially in cases when the learning process has to handle uncertainty or incomplete information.

"Semi-supervised approach to reject inference in credit scoring using SVMs"

Author:S. Maldonado, G. Paredes ICDM 2010

Maldonado and Paredes suggest a semi-supervised learning approach for SVM reject inference. Their proposed methodology uses

the information found within the rejected samples of a dataset, in which only the accepted applicants were used to train the labeled set, while the entire pool was used to train the SVM model. This article is an effective method toward addressing reject inference in the field of credit scoring when powerful SVM algorithms are considered.

III. IMPLEMENTATION

A Rejection-Aware Multi-Task Network (RMT-Net) needs the following key steps for implementing rejection inference in credit scoring. Data gathering and preprocessing is required, which involves both approved and rejected loan applications. Rejected samples lack the late/non-payment labels; thus, the data are misbiased. Feature engineering is related to handling the categorical and numerical features of applicants. An important aspect of overcoming this problem will be to deal with missing non-random (MNAR) bias in rejected samples through the estimation of the probability of rejection. Probabilities allow the model to estimate the relationship between decisions over reject/approve and a default behavior.

The core of the model uses Multi-Task Learning (MTL), where two tasks are learned simultaneously: one predicting default (whether the applicant will default on the loan) and the other predicting reject/approve decisions. The architecture uses a common neural network that feeds two task-specific output layers, each corresponding to one of the tasks. To allow better learning, a gating mechanism controls the amount of information shared between tasks, hence enabling the model to make adjustments based on the rejection probability and optimize predictions for both tasks simultaneously.

It trains the model by using a joint loss function that covers the losses from both the tasks. The optimization process is done through techniques such as Stochastic Gradient Descent (SGD) and Adam Optimizer. After training, the performance of the model is evaluated based on these metrics: Accuracy, Precision, Recall, and a Kolmogorov-Smirnov (KS) statistic, which analyses the potential discriminatory ability between defaulters and non-debtors by the model. RMT-Net can also be enhanced with some different rejection strategies to facilitate its flexibility in different kinds of credit scoring scenarios as well.

IV. ALGORITHM

Decision Tree Classifiers

Decision tree classifiers are widely applied in numerous domains because of their capacity to capture knowledge regarding the decision-making procedure from the data. This approach functions by recursively dividing the set into subsets based on the maximization of a sequence of tests or features, starting from the selection of the first test that divides the data into subsets until the whole objects in any subset fall into the same class. The result is a tree-like structure where each leaf node represents a class label. Decision trees are easy to interpret and understand, making them popular for decision-making tasks in industries like finance and healthcare.

Gradient Boosting

Gradient boosting is a powerful machine learning technique used in both regression and classification tasks. It builds an ensemble of weak prediction models, typically decision

trees, in a stage-wise manner. At this point, every model learnt to correct the mistakes it has been produced by one before it. The Gradient boosting is strength in how it can optimize an arbitrarily differentiated loss function which makes that much more versatile than even in boosting methods. The outcomes are usually far better and usually better than random forest especially when weak learners taken are decision trees.

K-Nearest Neighbors

KNN is one of the simplest yet very effective classification algorithms that classifies an object based on the majority class of its K-nearest neighbors in the feature space. It is a non-parametric, lazy learning method, meaning that it does not learn any model during the training phase. Instead, it memorizes the training data and makes predictions based on the closest data points during testing. KNN is apt for problems with complex decision boundaries and does not depend on any prior assumption concerning the distribution of data.

Logistic Regression Classifiers

Logistic regression is a statistical method used to model the relationship between a categorical dependent variable and independent variables. It is widely used for binary classification tasks, such as predicting whether a customer will default on a loan. Multinomial logistic regression can be used when the dependent variable has more than two categories. Unlike discriminant analysis, logistic regression does not assume normally distributed independent variables, making it more versatile for various types of data.

Naive Bayes

The Naïve Bayes classifier is based upon Bayes' theorem assuming that features are conditionally independent given the class label. This is an implausible assumption, yet many real-world classification tasks will perform very well with it. One of the applications where Naïve Bayes seems to work particularly well for text classification and spam filtering. It is easy to apply, very fast in training time, and performs very good on large datasets. Nonetheless, performance may degrade if the assumed independence is violated.

Random Forest

Random Forest is an ensemble learning method that combines multiple decision trees to improve the accuracy of classification and regression. It mitigates the overfitting problem of decision trees by averaging the predictions of many trees, each trained on a random subset of the data. Random Forest is robust and requires little tuning, making it a "black-box" model suitable for many real-world applications. Although it usually isn't as accurate in terms of gradient boosting, it will be faster during training, and overfits much less often.

Support Vector Machines (SVM)

SVM is discriminant machine learning techniques applied for classification purposes. The objective of SVM is to find the optimal hyperplane that can separate different classes in the feature space. SVM does not require the knowledge of the underlying distribution of the data and can work with high-dimensional spaces effectively. SVM has strong performance on tasks like image recognition

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

and text classification where there exists a clear margin of separation between classes.

V.RESULTS

Fig1:Home Page

Fig2:Enter Data Set

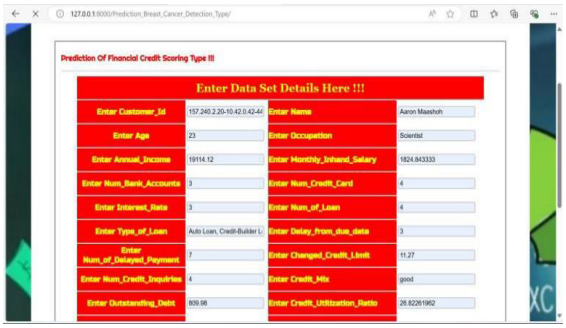

Fig3:Data Set Details

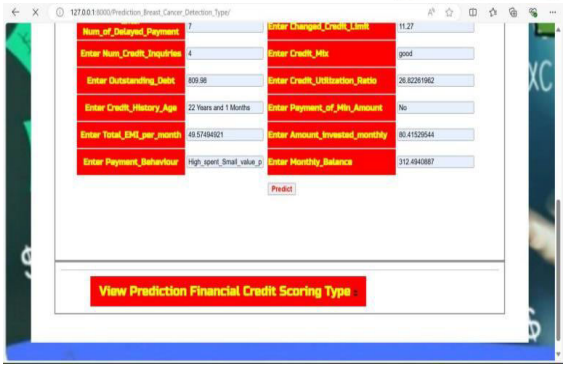

Fig4: Calculating Credit Score

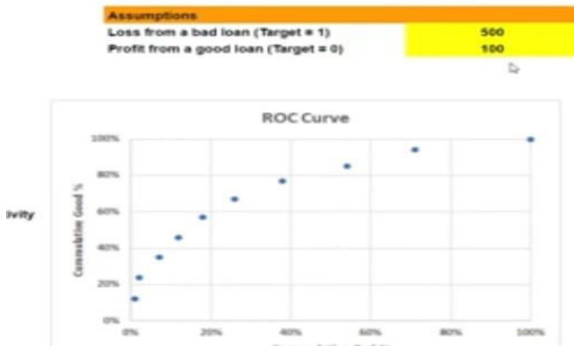

Fig5:ROC Curve

CONCLUSION

This paper focuses on addressing the challenge of modeling biased credit scoring data, where ground-truth labels are only available for approved samples, while data for rejected samples is absent. This bias introduces inaccuracies in predicting outcomes for both approved and rejected samples, thereby impacting the reliability of default predictions. Through theoretical analysis and experiments with real-world data, we observe a significant correlation between the tasks of predicting rejection/approval and default/non-default classifications in credit scoring systems.

To address this issue, we introduce **RMT-Net**, a novel approach that leverages a gating network based on rejection probabilities. This

gating mechanism determines the task weights, enabling effective information sharing between the rejection/approval and default/non-default tasks. RMT-Net represents the first solution specifically designed to handle biased credit scoring data and achieves superior performance compared to several state-of-the-art methods. Empirical evaluations across ten datasets under diverse scenarios demonstrate that RMT-Net significantly improves upon the limitations of existing multi-task learning (MTL) approaches.

Additionally, we extend this framework with **RMT-Net++**, which incorporates various rejection/approval strategies for scenario-based modeling. In experiments involving complex multi-policy scenarios, RMT-Net++ demonstrates further performance improvements by integrating multiple strategies, showcasing its versatility and robustness.

REFERENCES

[1] M. B¨ ucker, M. van Kampen, and W. Kr¨amer, "Reject inference inconsumer credit scoring with nonignorable missing data," Journal of Banking & Finance, vol. 37, no. 3, pp. 1040–1045, 2013.

[2] G. G. Chen and T. Astebro, "The economic value of reject inference in credit scoring," Department of Management Science, University of Waterloo, 2001.

[3]H.-T. Nguyen et al., "Reject inference in application scorecards: evidence from france," University of Paris Nanterre, EconomiX,Tech. Rep., 2016.

[4] D. J. Hand and W. E. Henley, "Can reject inference ever work?" IMA Journal of Management Mathematics, vol. 5, no. 1, pp. 45–55,1993.

[5] A. Agrawala, "Learning with a probabilistic teacher," IEEE Transactions on Information Theory, vol. 16, no. 4, pp. 373–379, 1970.

[6]S. Maldonado and G. Paredes, "A semi-supervised approach for reject inference in credit scoring using svms," in ICDM, 2010, pp.558–571.

[7] X. Zhu and A. B. Goldberg, "Introduction to semi-supervised learning," Synthesis lectures on artificial intelligence and machine learning, vol. 3, no. 1, pp. 1–130, 2009.

[8] D. C. Hsia, "Credit scoring and the equal credit opportunity act,"Hastings LJ, vol. 30, p. 371, 1978.

[9]J. Banasik and J. Crook, "Reject inference, augmentation, and sample selection," European Journal of Operational Research, vol. 183, no. 3, pp. 1582–1594, 2007.

[10] J. Banasik, J. Crook, and L. Thomas, "Sample selection bias in credit scoring models," Journal of the Operational Research Society, vol. 54, no. 8, pp. 822–832, 2003.

[11] N. Hassanpour and R. Greiner, "Counterfactual regression with importance sampling weights." in IJCAI, 2019, pp. 5880–5887.

[12] H. Zou, P. Cui, B. Li, Z. Shen, J. Ma, H. Yang, and Y. He,"Counterfactual prediction for bundle treatment," NeurIPS, 2020.

[13]R. Caruana, "Multitask learning," Machine learning, vol. 28, no. 1,pp. 41–75, 1997.

[14]J. Ma, Z. Zhao, X. Yi, J. Chen, L. Hong, and E. H. Chi, "Modeling task relationships in multi-task learning with multi-gate mixtureof-experts," in KDD, 2018, pp. 1930–1939.

[15] D. West, "Neural network credit scoring models," Computers & Operations Research, vol. 27, no. 11-12, pp. 1131–1152, 2000.
[16] B. Hu, Z. Zhang, J. Zhou, J. Fang, Q. Jia, Y. Fang, Q. Yu, and Y. Qi, "Loan default analysis with multiplex graph learning," in CIKM, 2020, pp. 2525–2532.
[17] Y. Liu, X. Ao, Q. Zhong, J. Feng, J. Tang, and Q. He, "Alike and unlike: Resolving class imbalance problem in financial credit risk assessment," in CIKM, 2020, pp. 2125–2128.
[18] Q. Liu, Z. Liu, H. Zhang, Y. Chen, and J. Zhu, "Mining cross features for financial credit risk assessment," in CIKM, 2021, pp. 1069–1078.
[19] D. Babaev, M. Savchenko, A. Tuzhilin, and D. Umerenkov, "Et-rnn: Applying deep learning to credit loan applications," in KDD, 2019.
[20] Z. Li, Y. Tian, K. Li, F. Zhou, and W. Yang, "Reject inference in credit scoring using semi-supervised support vector machines," Expert Systems with Applications, vol. 74, pp. 105–114, 2017.
[21] R. A. Mancisidor, M. Kampffmeyer, K. Aas, and R. Jenssen,"Deep generative models for reject inference in credit scoring," Knowledge-Based Systems, 2020.
[22] A. Ehrhardt, C. Biernacki, V. Vandewalle, P. Heinrich, and S. Beben, "Reject inference methods in credit scoring," Journal of Applied Statistics, pp. 1–21, 2021.
[23] N. Goel, A. Amayuelas, A. Deshpande, and A. Sharma, "The importance of modeling data missingness in algorithmic fairness: A causal perspective," in AAAI, 2021, pp. 7564–7573.
[24] B. M. Marlin and R. S. Zemel, "Collaborative prediction and ranking with non-random missing data," in RecSys, 2009, pp. 5–12.
[25] T. Schnabel, A. Swaminathan, A. Singh, N. Chandak, andT. Joachims, "Recommendations as treatments: Debiasing learningand evaluation," in ICML, 2016, pp. 1670–1679.