Email: editor@ijerst.com or editor.ijerst@gmail.com

Int. J. Eng. Res. & Sci. & Tech. 2025

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

https://doi.org/10.62643/ijerst.2025.v21.i2.pp1268-1276 Vol. 21, Issue 2, 2025

MULTI-TASK DEEP LEARNING FOR CREDIT RISK ASSESSMENT WITH REJECT-AWARE MNAR DATA HANDLING

[1]*Shaik Shafath, MCA Student, Department of MCA*

[2] *B Purushotham, M.Tech, Assistant Professor, Department of MCA*

[12]*Dr KV Subba Reddy Institute of Technology, Dupadu, Kurnool*

ABSTRACT

Financial credit scoring determines whether loan applications are accepted or denied. Missing-not-at-random selection bias results from the fact that we are only able to witness default/non-default labels for authorised samples and are unable to view rejected samples. Such biassed data will always lead to faulty machine learning models. Based on both theoretical analysis and real-world data investigation, we discover a strong correlation between the default/non-default classification task and the rejection/approval classification problem in this work. Therefore, rejection/approval may help with default/non-default learning. As a result, we initially suggest using Multi-Task Learning (MTL) to model the skewed credit score data. In particular, we present a unique Reject-aware Multi-Task Network (RMT-Net) that uses a gating network based on rejection probability to learn the task weights that govern the information passing from the rejection/approval task to the default/non-default task. RMT-Net makes use of the relationship between the two tasks, which states that the default/non-default task must learn more from the rejection/approval task the higher the rejection probability. Additionally, in order to simulate situations with different rejection/approval techniques, we expand RMT-Net to RMT-Net++. Numerous datasets are the subject of extensive testing, which firmly confirms RMT-Net's efficacy on both accepted and rejected samples. Furthermore, RMT-Net++ enhances RMT-Net's functionality.

I. INTRODUCTION

CREDIT scoring measures the likelihood that borrowers will fail on their credit loans using machine learning techniques. [1, 2], [3, 4], [5]. Financial organisations like banks and internet lenders have the authority to accept or deny credit loan applications based on the assessed credit.

The approval or rejection of a customer's credit loan application is a possible outcome. The consumer will get the loan if the application is accepted and becomes an authorised sample. It will be a non-default sample if the consumer repays the credit loan on time after a certain amount of time; if not, it will be a default sample. On the other hand, the consumer will not get a credit loan if the application is denied, making it a rejected sample. We are unable to determine if a rejected sample would default or not since they are not given any loans. Figure 1 depicts the aforementioned procedure. Since we lack ground-truth default/non-default labels for rejected samples, credit scoring models are often built using accepted samples. [6] [7] [8] [9]. Approved and rejected samples have varied feature distributions because the rejection/approval procedures are often machine learning models or expert guidelines based on customer attributes. This forces us to confront the bias in data that comes from missing-not-at-

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1268-1276
Vol. 21, Issue 2, 2025

random selection [9] [10] [11]. Credit scoring models must, however, infer loan application credits from feature distributions of both accepted and denied samples when they are used online. The model parameters are biassed as a result of training models with such biassed data [12], meaning that the expected relationship between input characteristics and default probability is not accurate. Significant financial losses result from applying such models to samples with different data distributions [7] [13] [14]. Therefore, in addition to modelling accepted samples, we also need to account rejected samples and infer their genuine credits for accurate credit scoring [15].

Credit rating data is often modelled using machine learning models such as XGBoost (XGB), Multi-Layer Perceptron (MLP), Support Vector Machines (SVM), and Logistic Regression (LR). However, their ability to provide accurate and dependable predictions is impacted by the missing-not-at-random bias in the data. Some current methods handle selection bias and perform reject inference from many viewpoints in order to overcome this issue. In order to retrain the model [17], some methods use the self-training algorithm [16], which progressively adds rejected samples with a higher default probability as default samples. This method is semisupervised [18]. Additionally, credit scoring systems also use Semi-Supervised Gaussian Mixture Models (SS-GMM) [7] and Semi-Supervised SVM (S3VM) [6]. From a different angle, several methods try to mimic unbiased data by re-weighting the accepted training samples [14] [19] [20] [21]. These methods are comparable to counterfactual learning, which aims to eliminate bias in data by re-weighting observed samples [10] [11] [22] [23].

Even while some of the aforementioned methods have improved credit score datasets somewhat [7] [14], they fall short of optimum results because they overlook certain important aspects. In particular, based on both real-world data research and theoretical analysis in Sec. 3, we conclude that the default/non-default classification task and the rejection/approval classification job are substantially connected in actual credit scoring applications. It makes sense that when a credit approval system is working well, authorised clients have lower default percentages and rejected customers have higher ones. As a result, learning of rejection/approval may help with learning of default/non-default. Therefore, using Multi-Task Learning (MTL) [24] to model biassed credit rating data may be beneficial.

These days, adaptively learning task weights in a mixture-of-experts framework is the primary focus of state-of-the-art MTL techniques. [26, 27; 28; 29] [25]. As a result, task weights vary between samples, allowing tasks to adaptively communicate information that is helpful but not conflicting. These MTL techniques show promise in a range of situations. However, we do not get good results, and we even obtain subpar results in default prediction on rejected samples, when we use state-of-the-art MTL techniques for modelling the default/nondefault task and the rejection/approval job. This might be as a result of the fact that during model training, we did not see default/non-default labels for rejected data. In the feature distribution of rejected samples, the task weights—which determine how much information is exchanged between the two tasks—are not optimally optimised. Therefore, we need a new and specifically created MTL technique since existing ones are ineffective at modelling the biassed credit score data.

II. LITERATURE SURVEY

"Models for neural network credit scoring,"
D. West

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1268-1276
Vol. 21, Issue 2, 2025

Five neural network models—the multilayer perceptron, mixture-of-experts, radial basis function, learning vector quantisation, and fuzzy adaptive resonance—are examined in this study for their accuracy in credit scoring. Ten-fold cross-validation is used to evaluate the neural network credit scoring models utilising two real-world data sets. The outcomes are compared to more conventional techniques being considered for commercial applications, such as decision trees, logistic regression, k closest neighbour, kernel density estimation, linear discriminant analysis, and logistic regression. The findings show that both the mixture-of-experts and radial basis function neural network models should be taken into consideration for credit scoring applications, and that the multilayer perceptron may not be the most accurate neural network model. The most accurate of the conventional techniques is determined to be logistic regression.

"Analysis of loan default using multiplex graph learning,"

Y. Fang, Q. Jia, Y. Hu, Z. Zhang, J. Zhou, J. Fang, Q. Yu, Y. Qi,

Industrial efforts primarily try to use a traditional classifier with a complex feature design for prediction in order to successfully identify loan default in the Mobile Credit Payment Service. These systems, however, overlook the essential fundamental characteristics of loan default detection—communicability, complementation, and induction—and do not take use of the multiplex linkages present in the financial situations. In order to solve these problems, we create a brand-new attributed multiplex graph-based method for detecting loan defaults that successfully incorporates multiplex relations in financial situations. An Attributed Multiplex Graph (AMG) is suggested as a way to simultaneously describe different relations and objects, as well as the rich characteristics on nodes and edges, in light of the complexity of financial scenarios. In order to include crucial information obtained from local structure in each element of AMG, we carefully develop relation-specific receptive layers with adaptive breadth functions. We also stack numerous propagation layers to investigate high-order connectivity information. Additionally, during end-to-end training, a relation-specific attention mechanism is used to highlight pertinent information. The efficiency of the suggested model in comparison to the state-of-the-art is confirmed by extensive tests carried out on the sizable real-world dataset. Furthermore, after a successful deployment in the Alipay APP, AMG-DP has also shown a 9.37% increase in performance on the KS measure in recent months.

"Differing and similar: Addressing the issue of class imbalance in financial credit risk assessment,"

Q. Zhong, J. Feng, J. Tang, Y. Liu, X. Ao, and Q. He,

In order to take early action before the real financial crisis occurs, financial credit risk assessment acts as the catalyst for evaluating a customer's credit admittance or prospective company collapse. Usually expressed as a binary classification issue, its goal is to forecast the likelihood that a consumer would be a member of a high-risk category. However, the current models have a serious class-imbalance issue since there aren't enough high-risk examples. This issue may be resolved by oversampling those high-risk users, but noise instances also have a greater impact. In order to address the class imbalance issue in financial credit risk assessment, we provide a unique adversarial data augmentation technique in this study. To distinguish between authentic and fraudulent instances, we use a discriminator to train a generator for creating synthetic samples. In addition, the generator and an auxiliary risk discriminator are trained together to evaluate the

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1268-1276
Vol. 21, Issue 2, 2025

credit risk. Results from experiments on three real-world datasets show how successful the solution is.

"Assessing financial credit risk through cross-feature mining,"

J. Zhu, Y. Chen, H. Zhang, Q. Liu, Z. Liu,

In some domains, such as healthcare and finance, machine learning models must be universally interpretable and accurate in order to be considered reliable. Among them, credit risk assessment is one of the main ways that financial organisations utilise machine learning to analyse consumer credit and identify fraud or default. Although they are not strong enough to explain intricate nonlinear interactions among information, simple white-box models like Logistic Regression (LR) are often utilised for credit risk assessment. Complex black-box models, on the other hand, are excellent at modelling but lack interpretability, particularly global interpretability. Thankfully, automated feature crossing shows promise as a method for identifying cross features to improve the accuracy of basic classifiers without requiring extensive handmade feature engineering. However, since relevant data often comprises hundreds of feature fields, current automated feature crossover approaches are inefficient when it comes to assessing credit risk.

In this study, we discover that local interpretations of a particular feature in Deep Neural Networks (DNNs) are often not consistent across samples. We show that nonlinear feature interactions in the DNN's hidden layers are to blame for this. As a result, feature interactions in DNN may be mined and used as cross features in LR. As a consequence, cross features will be mined more effectively. As a result, we suggest DNN2LR, a brand-new automated feature crossing technique. The final model produced by DNN2LR is a white-box LR model enhanced with cross features. Our trials on public and corporate datasets from actual credit risk assessment applications demonstrate that DNN2LR performs better than a number of feature crossing techniques as well as traditional credit assessment models. Furthermore, the suggested DNN2LR approach speeds up financial credit assessment datasets with hundreds of feature fields by around 10 to 40 times when compared to state-of-the-art feature crossing techniques like AutoCross.

III. SYSTEM ANALYSIS & DESIGN

EXISTING SYSTEM

As a re-weighting technique, augmentation involves re-weighting accepted samples to reflect the whole distribution [19] [20] [21]. Reweighting based on the likelihood of acceptance or rejection is a popular method for doing this. Furthermore, a fuzzy extension of the augmentation strategy has been made [14]. Another re-weighting strategy is parcelling, in which the credit modeller adjusts the re-weighting based on the default probability by score-band [8] [21]. Notably, these re-weighting techniques are comparable to counterfactual learning studies [10] [11] [22] [23]. Re-weighting training samples is a common practice in counterfactual learning, which attempts to eliminate data bias.

In the meanwhile, the reject inference problem is also handled using semi-supervised techniques. To enhance SVM's effectiveness on credit rating, the authors in [17] use a self-training approach. The simplest semi-supervised learning technique is self-training, often referred to as self-labeling or decision-directed learning [16] [30] [31]. Using accepted samples as training data, this method classifies rejected samples with the highest default probability as default samples based on model predictions. The model is then retrained using the freshly labelled data, and this iterative process is repeated. The self-training method may promote additional classifiers including LR, MLP, and XGB, although though it is only used

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1268-1276
Vol. 21, Issue 2, 2025

to promote SVM in [17]. Additionally, S3VM [6], another semi-supervised variant of SVM, is also used in reject inference. Although S3VM has trouble fitting large-scale data, it employs both rejected and authorised samples to fit an ideal hyperplane with the largest margin [7]. In contrast, previous research has used statistical machine learning techniques for reject inference, including the Expectation-Maximization (EM) algorithm [32], Gaussian Mixture Models (GMM) [33], and survival analysis [34]. For modelling biassed credit score data, SS-GMM [7], which is based on GMM and inspired by semi-supervised generative models [35] [36], is suggested. The primary techniques for reject inference are semi-supervised learning and counterfactual re-weighting; however, neither method takes into account the relationship between learning of default/non-default and learning of rejection/approval.

MTL learns multiple tasks simultaneously in one model, and has been proven to improve performances through information sharing between tasks [24] [26]. It has succeed in scenarios such as computer vision [29] [59] [60], recommender systems [25] [26] [27] [28] [61] [62], healthcare [63], and other prediction problems [64] [65]. The simplest MTL approach is hard parameter sharing, which shares hidden representations across different tasks, and only the last prediction layers are special for different tasks [24].

However, hard parameter sharing suffers from conflicts among tasks, due to the simple sharing of representations. To deal with this problem, some approaches propose to learn weights of linear combinations to fuse hidden representations in different tasks, such as Cross-Stitch Network [59] and Sluice Network [60]. However, in different samples, the weights of different tasks stay the same, which limits the performances of MTL. This inspires the research on applying gating structures in MTL [25] [26] [27] [66]. Mixture-Of- Experts (MOE) first proposes to share and combine several experts through a gating network [66]. Based on MOE, to make the weights of different tasks varying across different samples and to improve the performances of MTL, Multigate MOE (MMOE) [25] proposes to use different gates for different tasks. Progressive Layered Extraction (PLE) further extends MMOE, and incorporates multi-level experts and gating networks [26]. Besides, attention networks are also utilized for assigning weights of tasks according to different feature representations [28] [29].

Disadvantages

• The complexity of data: Most of the existing machine learning models must be able to accurately interpret large and complex datasets to detect Financial Credit Scoring.

• Data availability: Most machine learning models require large amounts of data to create accurate predictions. If data is unavailable in sufficient quantities, then model accuracy may suffer.

• Incorrect labeling: The existing machine learning models are only as accurate as the data trained using the input dataset. If the data has been incorrectly labeled, the model cannot make accurate predictions.

PROPOSED SYSTEM

A Reject-aware Multi-Task Network (RMT-Net) is suggested by the system. Using a gating network based on rejection probability, RMT-Net learns the weights that regulate the information transfer from the rejection/approval task to the default/non-default task. More information is exchanged from the rejection/approval network and less trustworthy information may be learnt in the default/non-default network when the rejection probability is higher. This allows us to customise the information sharing weights in the feature distribution of rejected samples and take into account the connection between rejected

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1268-1276
Vol. 21, Issue 2, 2025

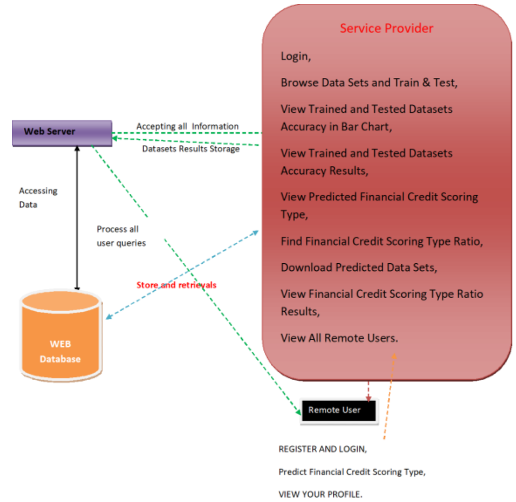
samples and default samples. Additionally, we expand RMT-Net to RMT-Net++, which models multiple rejection/approval classification jobs in the MTL framework, and take into account scenarios with multiple rejection/approval techniques.

Overall, we test RMT-Net and RMT-Net++ on ten datasets with varying configurations, demonstrating significant improvements in default prediction for both accepted and rejected samples. According to the widely used Kolmogorov-Smirnov (KS) metric1 in credit scoring, RMT-Net performs 47:9% better on average than traditional classifiers like LR, DNN, and XGB. RMT-Net comparatively enhances the performances by an average of 11:9% when compared to the most competitive reject inference techniques. Furthermore, we demonstrate in an additional experiment using several rejection/approval procedures that RMT-Net++ may further increase RMT-Net's performance by an average of 5:8%.

Advantages

For the first time, we suggest employing an MTL technique, namely RMTNet, to model biassed credit score data. We provide a number of changes to address the subpar credit scoring results of current MTL systems rather than relying just on traditional MTL techniques. Furthermore, we expand RMT-Net to RMT-Net++ and take into account various rejection/approval procedures. This makes our work suitable for a variety of real-world application settings. Ten datasets under various conditions are the subject of extensive experiments. Our suggested RMT-Net method yields significant gains on both accepted and rejected samples. Furthermore, we demonstrate that RMT-Net++ may significantly enhance performance by using different techniques.

SYSTEM ARCHITECTURE



IV. IMPLEMENTATIONS

Modules

Service Provider

The Service Provider must use a working user name and password to log in to this module. Following a successful login, he may do several tasks including browsing data sets and training and testing. See the Accuracy of Trained and Tested Datasets in a Bar Chart, See the Accuracy Results of Trained and Tested Datasets, See the Financial Credit Scoring Type Prediction, Find the Financial Credit Scoring Type Ratio, and Download the Predicted Data Sets View All Remote Users and Financial Credit Scoring Type Ratio Results.

View and Authorize Users

The administrator may see a list of all registered users in this module. Here, the administrator may see the user's information, like name, email, and address, and they can also grant the user permissions.

Remote User

A total of n users are present in this module. Before beginning any actions, the user needs register. Following registration, the user's information will be entered into the database. Following a successful registration, he must use

International Journal of Engineering Research and Science & Technology

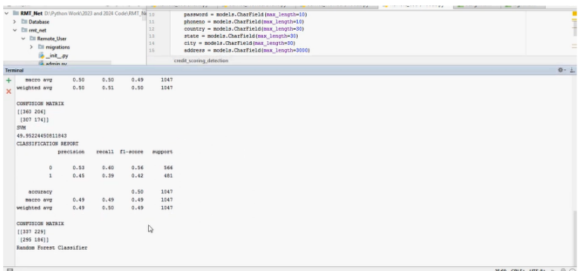
ISSN 2319-5991 www.ijerst.com

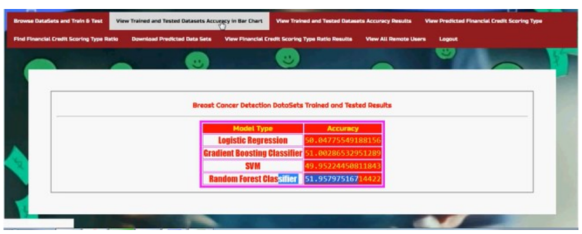
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1268-1276

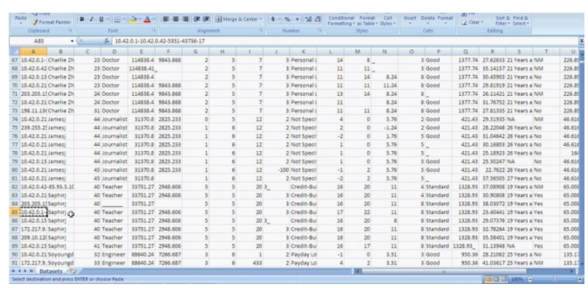
Vol. 21, Issue 2, 2025

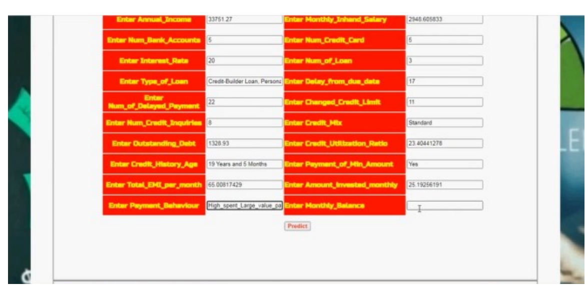
his password and authorised user name to log in. Following a successful login, the user will be able to see their profile, predict their financial credit scoring type, and register and log in.

V. SCREENSHOTS









International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1268-1276
Vol. 21, Issue 2, 2025

VI. CONCLUSION

This work focusses on modelling biassed credit score data, where rejected samples have no observations and accepted samples have just ground-truth labels. We want to increase the prediction accuracy on both approved and rejected samples since this bias compromises the default forecast's dependability. Based on both theoretical analysis and real-world data investigation, we discover that there is a strong correlation between the default/non-default classification task and the rejection/approval classification job in credit scoring apps. We present a novel RMT-Net approach, which uses a gating network based on rejection probabilities to learn the task weights that govern the information sharing from the rejection/approval task to the default/non-default task. This is the first time that we propose to use an MTL framework to model biassed credit scoring data. Empirical trials on ten datasets in various contexts show that RMT Net greatly outperforms a number of state-of-the-art techniques from several angles and improves the subpar performances of current MTL approaches. Additionally, in order to simulate situations with different rejection/approval techniques, we expand RMT-Net to RMT-Net++. An further experiment indicates that RMT-Net++ with various strategies may enhance RMT-Net's performance in a more intricate multi-policy environment.

REFERENCES

[1] D. West, "Neural network credit scoring models," Computers & Operations Research, vol. 27, no. 11-12, pp. 1131–1152, 2000.

[2] 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.

[3] 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.

[4] 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.

[5] D. Babaev, M. Savchenko, A. Tuzhilin, and D. Umerenkov, "Et-rnn: Applying deep learning to credit loan applications," in KDD, 2019.

[6] 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.

[7] R. A. Mancisidor, M. Kampffmeyer, K. Aas, and R. Jenssen, "Deep generative models for reject inference in credit scoring," Knowledge-Based Systems, 2020.

[8] 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.

[9] N. Goel, A. Amayuelas, A. Deshpande, and A. Sharma, "The importance of modeling data

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
https://doi.org/10.62643/ijerst.2025.v21.i2.pp1268-1276
Vol. 21, Issue 2, 2025
missingness in algorithmic fairness: A causal perspective," in AAAI, 2021, pp. 7564–7573.

[10] B. M. Marlin and R. S. Zemel, "Collaborative prediction and ranking with non-random missing data," in RecSys, 2009, pp. 5–12.

[11] T. Schnabel, A. Swaminathan, A. Singh, N. Chandak, and T. Joachims, "Recommendations as treatments: Debiasing learning and evaluation," in ICML, 2016, pp. 1670–1679.

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

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

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

[15] 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.

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

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

[18] 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.

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

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