Development of Machine Learning Models for Bankruptcy Prediction of Taiwanese Companies

Georgios Iatridis
Department of Digital Systems
University of Piraeus
Piraeus, Greece
gewrgios.iatrides@gmail.com

***Abstract*—This paper analyzes data regarding the critical problem of corporate bankruptcy prediction, utilizing a dataset of Taiwanese companies [1], which is characterized by severe class imbalance. The analysis is performed using the scikit-learn library in a Python environment. To manage high dimensionality, feature selection methods such as T-test and the LASSO algorithm are applied and evaluated. Concurrently, the issue of data imbalance is addressed through sampling techniques. Within the scope of this research, various classification algorithms are examined and compared to extract results regarding their performance. Following a comparative review of the results, the optimal combination of sampling and modeling is identified, upon which hyperparameter optimization techniques are applied. Finally, appropriate metrics are determined based on the nature of the problem, and an optimal decision threshold is defined for effective credit risk management.**

***Keywords*—Bankruptcy Prediction, Class Imbalance, Feature Selection, Multilayer Perceptron (MLP), Undersampling, Threshold Adjustment, Recall.**

I. Introduction

Corporate bankruptcy prediction constitutes one of the most significant issues in the field of financial management, as it directly affects investors, credit institutions, and broader economic stability. The complexity of the problem lies, on one hand, in the large volume of financial ratios that must be analyzed and, on the other, in the rarity of bankruptcy incidents compared to healthy businesses. The present study utilizes the "Taiwanese Bankruptcy Prediction" dataset [1], which is characterized by intense class imbalance.

In international literature, this problem has been studied extensively, with research focusing both on the selection of appropriate indicators and on addressing data imbalance. Shi and Li [2] conducted a systematic literature review covering 50 years of research (1968-2017). Their study maps the evolution of the field from traditional statistical methods, such as Multivariate Discriminant Analysis (MDA), towards modern Machine Learning techniques and Deep Learning. Their main conclusion is that, although Deep Learning models are gaining ground, feature selection remains the most critical factor for prediction accuracy.

In a more recent study, Altalhan et al. [3] focus exclusively on the problem of imbalanced datasets. They thoroughly analyze preprocessing techniques, such as synthetic sample generation (SMOTE and GANs), and evaluate cost-sensitive learning algorithms. They point out that the use of standard metrics like Accuracy is misleading in such problems and propose the adoption of metrics such as AUC and G-mean for reliable model evaluation.

Veganzones and Séverin [4] investigated the interaction between feature selection and sampling. Through experiments on multiple datasets, they demonstrated that the application of feature selection methods alone is insufficient when the class ratio is extremely unequal. They propose hybrid approaches combining feature filtering with class balancing techniques to maximize predictive capability.

Of particular importance for this paper is the research by Wang and Liu [6], who specifically studied the Taiwanese bankruptcy dataset. By comparing different sampling strategies, they concluded that the method of Undersampling the majority class outperforms Oversampling on this specific data. This finding supports our methodological choice, as undersampling reduces noise and the risk of overfitting often introduced by synthetic methods.

In the same context, Lane et al. [5] developed analysis models focusing on prediction stability. Comparing algorithms such as Support Vector Machines (SVM) and Decision Trees, they found that the selection of the feature subset dramatically affects the model's sensitivity to imbalance. They argue that using simpler models with carefully selected features can often perform better than complex algorithms in high-uncertainty environments.

Alaminos et al. [7] attempted to create a global bankruptcy prediction model using Logistic Regression on data from Asia, Europe, and America. Through their analysis, they identified that while certain financial ratios have universal validity, predictive accuracy improves significantly when models are adapted to the macroeconomic characteristics of each geographical zone, highlighting the importance of local model adaptation.

Liang et al. [8] introduced a comprehensive approach combining financial ratios with corporate governance indicators, such as board structure and shareholding composition. Using Genetic Algorithms and RFE for feature selection, they proved that incorporating qualitative governance indicators significantly increases prediction accuracy compared to using purely accounting data.

In a more innovative direction, Hosaka [9] proposed converting numerical financial ratios into grayscale images. This technique allowed the use of Convolutional Neural Networks (CNNs), algorithms typically used in computer vision. The results showed that CNNs can detect complex, non-linear patterns in financial data that elude traditional classifiers.

Finally, Lahmiri and Bekiros [10] evaluated the predictive ability of various machine learning approaches using qualitative rather than quantitative data. Comparing methods such as SVM, Neural Networks (NN), and Discriminant Analysis, they concluded that non-linear techniques (such as NNs) consistently outperform others in categorizing distressed businesses, especially when data contains noise and uncertainty.

The present work proposes an integrated methodology combining findings from the literature. Initially, data cleaning and preprocessing are applied. Subsequently, two different feature selection techniques (T-test and LASSO) are compared to identify the most influential indicators. The imbalance problem is approached with three different sampling methods: *Undersampling, Oversampling*, and *Ensemble Methods with embedded balancing*. Different algorithms are trained for each method to yield the optimal combination, which is then subjected to hyperparameter tuning to achieve maximum performance. The strategic prioritization of Recall over Precision through decision threshold adjustment is decisive for minimizing credit risk.

The paper is structured as follows: Section II describes the Methodology and algorithms used. Section III presents the experimental results, ROC curves, and confusion matrices, while Section IV summarizes the conclusions and suggestions for future research.

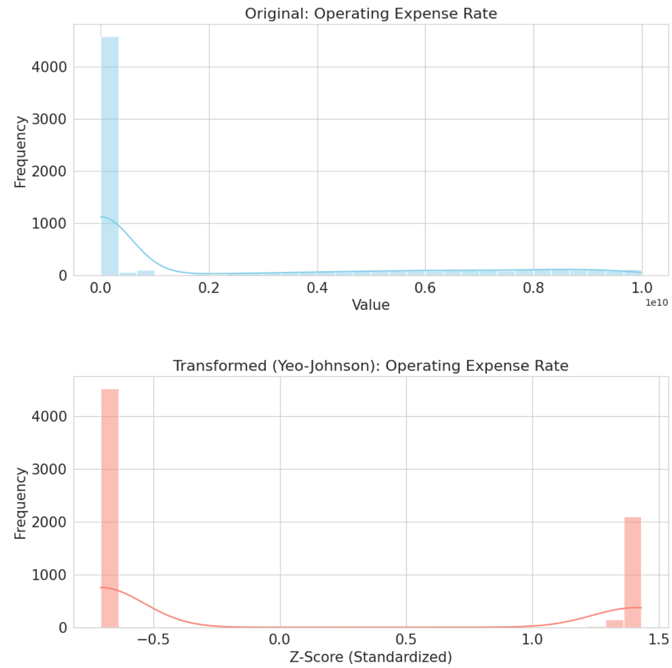

Figure 1. Comparison of distribution before and after Yeo-Johnson transformation. Skewness is reduced, but the distribution remains non-normal.

II. Methodology

The proposed methodological approach was developed in three distinct stages: preprocessing and feature selection, experimental evaluation of multiple sampling and classification combinations, and finally, optimization of the prevailing model.

A. Preprocessing and Feature Selection

In the preprocessing stage, the application of the *Yeo-Johnson* transformation [11] was investigated to address distribution skewness. However, preliminary experiments showed that its application not only failed to bring substantial improvement to the distribution of all features (Figure 1) but also did not improve the predictive ability of the models, and thus was not adopted in the final workflow.

To reduce dimensionality, a hybrid feature selection strategy was followed:

1) *Statistical Test:* Application of T-test with significance level $p < 0.05$.
2) *Embedded Method:* Use of LASSO algorithm (L1 penalty) retaining variables with a positive score.

The results for the top 30 features are depicted in Table I.

Regardless of the statistical test results, it was deemed necessary to include specific ratios that international literature perennially recognizes as fundamental for viability assessment. This addition reinforces the model's interpretability and covers aspects potentially ignored by automated methods. Specifically:

- *Working Capital to Total Assets* and *Retained Earnings to Total Assets:* These ratios constitute core components of the classic Z-score model and are highlighted in reviews by Shi and Li [2] and Alaminos et al. [7] as critical for measuring liquidity and accumulated profitability.
- *Net Income to Total Assets (ROA):* Considered one of the strongest efficiency indicators. Liang et al. [8] and Hosaka [9] emphasize that a company's ability to generate profit from its assets is the most decisive survival factor.
- *Debt ratio %* and *Total debt/Total net worth:* Leverage constitutes a structural element of credit risk. Wang and Liu [6], analyzing this specific Taiwanese dataset, as well as Veganzones and Séverin [4], underscore the importance of these ratios in imbalanced data.
- *Current Ratio:* As the most traditional measure of short-term liquidity, literature [7], [8] ranks it among the necessary indicators for assessing the ability to repay current liabilities.

The final set of 28 features (excluding the target variable Bankruptcy), as derived from the above selection processes, is presented collectively in Table II. These features cover the entire spectrum of the company's economic activity, including liquidity, efficiency, and operational activity. The table lists the English terminology of the indicators as found in the dataset, accompanied by the corresponding description.

Table I
Comparative Table of Feature Selection (T-test vs LASSO)

T-test Method			LASSO Method	
Feature	**Score**	**p-value**	**Feature**	**Score**
Net Value Per Share (B)	22.0	2.7×10^{-61}	ROA(C) before interest and depr.	14.2
Cash/Total Assets	16.3	1.2×10^{-42}	Cash/Total Assets	7.3
ROA(C) before interest and depr.	16.3	7.0×10^{-40}	Tax rate (A)	2.2
Tax rate (A)	11.1	2.7×10^{-23}	Net Value Per Share (B)	1.1
Operating Gross Margin	9.5	1.8×10^{-18}	Fixed Assets Turnover Frequency	0.8
Cash flow rate	8.8	1.5×10^{-16}	Liability-Assets Flag	0.8
Quick Assets/Total Assets	7.4	2.4×10^{-12}	Cash Turnover Rate	0.5
Total expense/Assets	5.8	2.2×10^{-8}	Quick Assets/Total Assets	0.4
Interest-bearing debt interest rate	5.6	3.3×10^{-8}	Current Assets/Total Assets	0.3
Fixed Assets Turnover Frequency	4.8	3.6×10^{-6}	Total Asset Growth Rate	0.2
Total Asset Growth Rate	4.6	8.3×10^{-6}	Quick Asset Turnover Rate	0.2
Current Assets/Total Assets	3.5	5.2×10^{-4}	Inventory Turnover Rate (times)	0.2
Non-industry income and exp./rev.	2.8	0.005	Operating Expense Rate	0.1
After-tax Net Profit Growth Rate	2.7	0.006	Operating Gross Margin	0.0
Liability-Assets Flag	2.4	0.015	Research and development exp. rate	0.0
Research and development exp. rate	2.1	0.034	Non-industry income and exp./rev.	0.0
Revenue Per Share (Yuan)	2.1	0.034	Interest-bearing debt interest rate	0.0
Quick Asset Turnover Rate	2.1	0.037	Cash flow rate	0.0
Inventory and acc. rec./Net value	2.0	0.044	Interest Expense Ratio	0.0
Cash Turnover Rate	1.8	0.071	Current Ratio	0.0
Current Liabilities/Liability	1.7	0.083	After-tax Net Profit Growth Rate	0.0
Total assets to GNP price	1.3	0.191	Net Value Growth Rate	0.0
Accounts Receivable Turnover	1.1	0.253	Realized Sales Gross Profit Growth	0.0
Contingent liabilities/Net worth	1.1	0.284	Revenue Per Share (Yuan)	0.0
Current Ratio	1.0	0.317	Revenue per person	0.0
Net Value Growth Rate	1.0	0.321	Accounts Receivable Turnover	0.0
Revenue per person	1.0	0.339	Contingent liabilities/Net worth	0.0
Current Asset Turnover Rate	1.0	0.342	Inventory and acc. rec./Net value	0.0
No-credit Interval	0.8	0.405	Long-term fund suitability ratio (A)	0.0
Total debt/Total net worth	0.7	0.461	Total debt/Total net worth	0.0

Table II
Final Feature Set and Description

Feature	Description
After-tax Net Profit Growth Rate	Growth rate of net profit after tax
Cash Turnover Rate	Velocity of cash circulation
Cash flow rate	Rate of cash flow
Cash/Total Assets	Ratio of Cash to Total Assets
Current Assets/Total Assets	Current Assets / Total Assets
Current Ratio	General Liquidity Ratio
Debt ratio %	Percentage of Debt
Fixed Assets Turnover Frequency	Frequency of fixed assets turnover
Interest-bearing debt interest rate	Interest rate on interest-bearing debt
Inventory Turnover Rate (times)	Velocity of inventory circulation
Inventory and acc. rec./Net value	Inventory & Receivables / Net Value
Liability-Assets Flag	Over-indebtedness Flag (if Liab. > Assets)
Net Income to Total Assets	Net Income / Total Assets (ROA)
Net Value Per Share (B)	Net Worth per Share (Book Value)
Non-industry income & exp./rev.	Non-operating income-expenses / Revenue
Operating Expense Rate	Operating Expense Percentage
Operating Gross Margin	Gross Operating Profit Margin
Quick Asset Turnover Rate	Velocity of quick assets turnover
Quick Assets/Total Assets	Quick Assets / Total Assets
ROA(C) before interest and depr.	Return on Assets (before interest/depr.)
R&D expense rate	R&D expense percentage
Retained Earnings to Total Assets	Retained Earnings / Total Assets
Revenue Per Share (Yuan ¥)	Revenue per Share
Tax rate (A)	Tax Rate
Total Asset Growth Rate	Growth Rate of Total Assets
Total debt/Total net worth	Total Debt / Net Worth
Total expense/Assets	Total Expenses / Total Assets
Working Capital to Total Assets	Working Capital / Total Assets

B. *Experimental Design and Model Selection*

Due to the severe class imbalance (Figure 2), an extensive evaluation experiment was designed that combined various sampling techniques and machine learning algorithms.

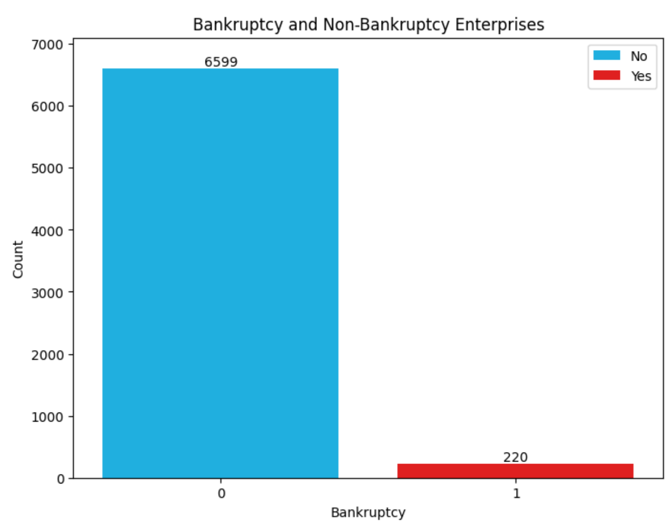

Figure 2. The number of records for each class in the dataset.

Specifically, the following sampling methods were examined:

- *Undersampling:* Random Undersampling, NearMiss, Edited Nearest Neighbours (ENN), Repeated ENN

(RENN), Neighbourhood Cleaning Rule (NCR).

- *Oversampling:* SMOTE, BorderlineSMOTE, KMeansSMOTE.
- *No sampling.*

For each sampling method, the following algorithms were trained: *Logistic Regression, MLP (NeuralNet), SVM, Naive Bayes, Random Forest, XGBoost, Balanced Random Forest, RUSBoost*, and *EasyEnsemble*.

C. Optimization and Threshold Adjustment

The selected model was further optimized via a GridSearchCV process to find the optimal hyperparameters. Finally, a decision threshold shifting strategy was applied, aiming to optimize the cost-benefit relationship between false positive and false negative predictions.

III. Results

A. Selection of Optimal Combination

From the experimental evaluation of all possible sampling and classification pairs, *Random Undersampling* with the *MLPClassifier* algorithm emerged as the optimal combination. This specific combination was chosen because it achieved the highest Recall while maintaining the highest Precision compared to other top combinations, offering the best balance of sensitivity and stability. Table III presents the summary of best performances per sampling category. For completeness, detailed results for all 57 experimental combinations are listed in Appendix A.

As shown in Table III, the majority of sampling methods highlight the Naive Bayes algorithm as the prevailing one based on Recall. However, it is observed that achieving extremely high Recall rates (above 95%) in these cases is accompanied by dramatically low Precision (approximately 4%), implying an unmanageable number of false positive results. In other words, the algorithm predicts that every record belongs to the Bankruptcy class to increase true positives, completely ignoring the existence of the Non-Bankruptcy class.

In contrast, the combination of **Random Undersampling** with the **Multilayer Perceptron (MLP)** achieves the optimal balance, keeping Recall at high levels (92.7%) while more than doubling Precision and maximizing the F2-score (0.362) compared to alternative approaches. This fact makes this specific combination the most reliable choice for minimizing credit risk while limiting noise from false alarms.

B. Hyperparameter Optimization

Through the Grid Search process, the hyperparameter space of the MLPClassifier was explored. The optimal parameters that emerged and were adopted for the final model are presented in Table IV.

Table IV
Best Hyperparameters for MLPClassifier

Parameter	Value
Activation Function	ReLU
Solver	Adam
Hidden Layer Sizes	(50, 50)
Alpha (L2 Regularization)	0.05
Learning Rate Init	0.01

C. Threshold Analysis and Performance

The evaluation of the ROC curve (Figure 3) demonstrated that the default classification threshold (0.50) did not optimally serve business objectives, as it favored Precision at the expense of Recall.

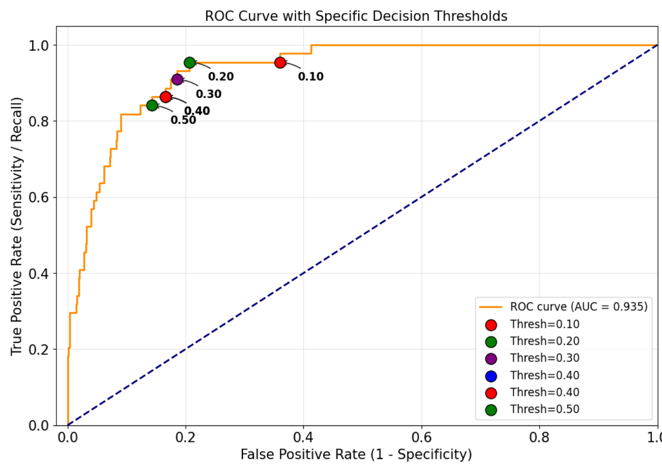

Figure 3. ROC Curve and operating points for different thresholds.

Therefore, shifting the decision threshold to **0.20** was selected. This point is considered optimal for the business application of the model in a risk management system. The cost of failing to detect a company about to go bankrupt (False Negative) is considered incomparably higher than the cost of a false alarm for a healthy business (False Positive).

This shift ensures that the model functions as an effective protection system, drastically increasing early warning capability. As shown in Table V, Recall for the bankruptcy class reaches **95%**.

The structure of predictions is depicted in the Confusion Matrix (Figure 4).

More specifically, the results are shaped as follows:

- *True Positives (TP): 42.* The model successfully identified 42 out of 44 bankrupt companies.
- *False Negatives (FN): 2.* The model failed to identify 2 bankruptcy cases, which were incorrectly classified as viable.
- *True Negatives (TN): 1021.* 1021 healthy companies were correctly identified.
- *False Positives (FP): 299.* An increase in false positive results was observed, as a consequence of lowering the decision threshold to minimize credit risk.

Table III
Summary of Best Performances by Sampling Method

Sampling Method	Best Algorithm	AUC	Recall	Precision	F1	F2	Time (s)
None (Baseline)	Naive Bayes	0.886	0.973	0.038	0.073	0.164	0.24
SMOTE	Naive Bayes	0.882	0.968	0.041	0.079	0.175	0.27
NCR	Naive Bayes	0.885	0.932	0.066	0.124	0.257	1.19
Random Undersampling	**MLP (Neural Net)**	**0.899**	**0.927**	**0.106**	**0.190**	**0.362**	**0.51**
Borderline SMOTE	Naive Bayes	0.861	0.923	0.048	0.091	0.197	0.31
ENN	Naive Bayes	0.882	0.914	0.084	0.154	0.305	1.09
Repeated ENN	Naive Bayes	0.879	0.909	0.091	0.165	0.323	6.96
KMeans SMOTE	Naive Bayes	0.853	0.886	0.072	0.133	0.270	0.38
NearMiss	XGBoost	0.786	0.809	0.061	0.113	0.233	0.78

Table V
Classification Report for Threshold 0.20

Class	Precision	Recall	F1-Score	Support
Non-Bankruptcy	1.00	0.77	0.87	1320
Bankruptcy	**0.12**	**0.95**	**0.22**	**44**
Accuracy			0.78	1364
Macro Avg	0.56	0.86	0.54	1364
Weighted Avg	0.97	0.78	0.85	1364

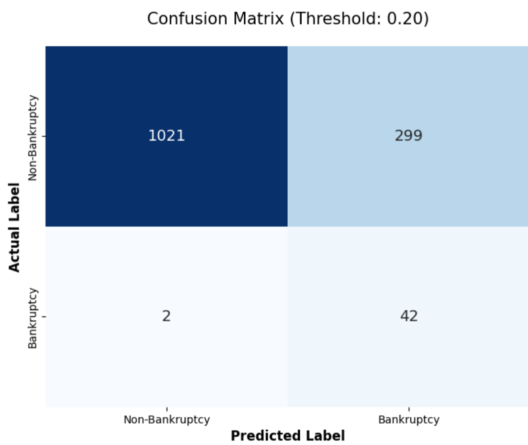

Figure 4. Confusion Matrix for Threshold 0.20.

IV. Conclusions

The present work highlighted the importance of combined use of hybrid feature selection and undersampling techniques (Random Undersampling) for effective corporate bankruptcy prediction in highly imbalanced data. The application of the optimized Multilayer Perceptron (MLP), combined with the targeted adjustment of the decision threshold to 0.20, led to a model of high sensitivity.

The main conclusion of the study is that the strategic prioritization of Recall over Precision constitutes the optimal approach for credit risk management. Shifting the threshold ensures that the model acts as a safety valve, increasing the early warning capability for 95% of bankruptcy cases. Although this strategy increases the number of false positive results, this cost is deemed acceptable in order to minimize the catastrophic risk of unforeseen bankruptcy.

Future extensions of the research could be oriented towards the application of advanced *Deep Learning* algorithms, such as Long Short-Term Memory (LSTM) networks. Unlike the present Multilayer Perceptron (MLP) which approaches the problem statically, such architectures have the ability to model time series, identifying longitudinal patterns of deteriorating financial health that gradually lead to bankruptcy.

Appendix A
Detailed Experimental Results

This appendix details the results for all combinations of sampling methods and classification algorithms examined within the scope of the research (Table VI).

Table VI
CONSOLIDATED EXPERIMENTAL RESULTS

Sampling Method	**Algorithm**	**AUC**	**Recall**	**Precision**	**F1**	**F2**	**Time (s)**
Borderline SMOTE	Naive Bayes	0.861	0.923	0.048	0.091	0.197	0.31
Borderline SMOTE	Logistic Regression	0.914	0.759	0.200	0.316	0.485	0.39
Borderline SMOTE	SVM	0.919	0.605	0.249	0.350	0.467	43.92
Borderline SMOTE	MLP (Neural Net)	0.905	0.523	0.298	0.375	0.449	12.83
Borderline SMOTE	Random Forest	0.924	0.495	0.398	0.438	0.470	26.41
Borderline SMOTE	XGBoost	0.916	0.495	0.387	0.430	0.466	3.42
ENN	Naive Bayes	0.882	0.914	0.084	0.154	0.305	1.09
ENN	XGBoost	0.920	0.386	0.404	0.392	0.388	3.50
ENN	Logistic Regression	0.920	0.345	0.434	0.382	0.359	1.12
ENN	Random Forest	0.918	0.345	0.494	0.404	0.366	8.57
ENN	MLP (Neural Net)	0.914	0.314	0.462	0.361	0.329	3.60
ENN	SVM	0.880	0.195	0.479	0.271	0.220	8.24
KMeans SMOTE	Naive Bayes	0.853	0.886	0.072	0.133	0.270	0.38
KMeans SMOTE	Logistic Regression	0.909	0.664	0.244	0.353	0.488	0.48
KMeans SMOTE	SVM	0.925	0.532	0.286	0.368	0.449	29.06
KMeans SMOTE	MLP (Neural Net)	0.910	0.468	0.318	0.376	0.425	8.00
KMeans SMOTE	XGBoost	0.920	0.409	0.432	0.419	0.413	2.45
KMeans SMOTE	Random Forest	0.920	0.336	0.430	0.373	0.349	28.41
NCR	Naive Bayes	0.885	0.932	0.066	0.124	0.257	1.19
NCR	XGBoost	0.918	0.377	0.464	0.409	0.388	2.39
NCR	Logistic Regression	0.921	0.314	0.483	0.378	0.336	1.20
NCR	Random Forest	0.925	0.286	0.503	0.363	0.313	10.20
NCR	MLP (Neural Net)	0.911	0.191	0.473	0.268	0.216	3.36
NCR	SVM	0.870	0.105	0.580	0.173	0.124	8.40
NearMiss	XGBoost	0.786	0.809	0.061	0.113	0.233	0.78
NearMiss	SVM	0.679	0.791	0.045	0.085	0.182	0.78
NearMiss	Random Forest	0.778	0.786	0.055	0.103	0.215	1.52
NearMiss	Logistic Regression	0.828	0.782	0.084	0.152	0.294	0.33
NearMiss	MLP (Neural Net)	0.596	0.764	0.036	0.068	0.150	0.67
NearMiss	Naive Bayes	0.502	0.350	0.036	0.066	0.128	0.32
No Sampling	Naive Bayes	0.886	0.973	0.038	0.073	0.164	0.24
No Sampling	Easy Ensemble	0.927	0.868	0.152	0.258	0.446	7.89
No Sampling	Balanced RF	0.935	0.818	0.201	0.322	0.505	3.25
No Sampling	RUSBoost	0.888	0.686	0.136	0.216	0.350	0.33
No Sampling	XGBoost	0.916	0.227	0.500	0.310	0.254	2.21
No Sampling	Logistic Regression	0.917	0.191	0.591	0.287	0.220	2.72
No Sampling	Random Forest	0.923	0.136	0.615	0.222	0.161	8.12
No Sampling	MLP (Neural Net)	0.819	0.068	0.632	0.117	0.082	1.94
No Sampling	SVM	0.848	0.036	0.550	0.068	0.045	9.39
Random Undersampling	**MLP (Neural Net)**	**0.899**	**0.927**	**0.106**	**0.190**	**0.362**	**0.51**
Random Undersampling	SVM	0.926	0.868	0.155	0.263	0.452	0.42
Random Undersampling	Random Forest	0.926	0.864	0.158	0.267	0.456	2.02
Random Undersampling	Logistic Regression	0.919	0.845	0.156	0.263	0.448	0.26
Random Undersampling	XGBoost	0.917	0.845	0.155	0.262	0.446	0.84
Random Undersampling	Naive Bayes	0.880	0.423	0.197	0.232	0.274	0.25
Repeated ENN	Naive Bayes	0.879	0.909	0.091	0.165	0.323	6.96
Repeated ENN	XGBoost	0.919	0.459	0.396	0.424	0.444	6.12
Repeated ENN	Logistic Regression	0.927	0.423	0.408	0.413	0.419	5.58
Repeated ENN	Random Forest	0.926	0.386	0.444	0.410	0.395	13.69
Repeated ENN	MLP (Neural Net)	0.899	0.309	0.478	0.343	0.317	8.79
Repeated ENN	SVM	0.893	0.286	0.450	0.343	0.306	10.64
SMOTE	Naive Bayes	0.882	0.968	0.041	0.079	0.175	0.27
SMOTE	Logistic Regression	0.917	0.805	0.165	0.273	0.452	0.36
SMOTE	SVM	0.919	0.686	0.186	0.292	0.444	62.66
SMOTE	MLP (Neural Net)	0.905	0.514	0.306	0.381	0.450	14.34
SMOTE	Random Forest	0.925	0.514	0.352	0.415	0.468	25.78
SMOTE	XGBoost	0.912	0.468	0.364	0.407	0.440	2.78

References

[1] "Taiwanese Bankruptcy Prediction," UCI Machine Learning Repository, 2020, DOI: https://doi.org/10.24432/C5004D.

[2] Y. Shi and X. Li, "An overview of bankruptcy prediction models for corporate firms: A systematic literature review," *Intangible capital*, vol. 15, no. 2, pp. 114–127, 2019.

[3] M. Altalhan, A. Algarni, and M. T.-H. Alouane, "Imbalanced data problem in machine learning: A review," *IEEE Access*, 2025.

[4] D. Veganzones and E. Séverin, "An investigation of bankruptcy prediction in imbalanced datasets," *Decision Support Systems*, vol. 112, pp. 111–124, 2018.

[5] P. C. Lane, D. Clarke, and P. Hender, "On developing robust models for favourability analysis: Model choice, feature sets and imbalanced data," *Decision Support Systems*, vol. 53, no. 4, pp. 712–718, 2012.

[6] H. Wang and X. Liu, "Undersampling bankruptcy prediction: Taiwan bankruptcy data," *Plos one*, vol. 16, no. 7, p. e0254030, 2021.

[7] D. Alaminos, A. Del Castillo, and M. Á. Fernández, "A global model for bankruptcy prediction," *PloS one*, vol. 11, no. 11, p. e0166693, 2016.

[8] D. Liang, C.-C. Lu, C.-F. Tsai, and G.-A. Shih, "Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study," *European journal of operational research*, vol. 252, no. 2, pp. 561–572, 2016.

[9] T. Hosaka, "Bankruptcy prediction using imaged financial ratios and convolutional neural networks," *Expert systems with applications*, vol. 117, pp. 287–299, 2019.

[10] S. Lahmiri and S. Bekiros, "Can machine learning approaches predict corporate bankruptcy? evidence from a qualitative experimental design," *Quantitative Finance*, vol. 19, no. 9, pp. 1569–1577, 2019.

[11] I.-K. Yeo and R. A. Johnson, "A new family of power transformations to improve normality or symmetry," *Biometrika*, vol. 87, no. 4, pp. 954–959, 2000.