International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

AN OPTIMIZED FEATURE SELECTION APPROACH FOR ENHANCED CREDIT SCORING PREDICTION: A HYBRID RF-L1 AND LOGISTIC REGRESSION MODEL

K. Rizwana Parveen[1], P. Thangaraju[2]

[1] Research Scholar, [2] Associate Professor, PG & Research Department of Computer Science,
Bishop Heber College (Autonomous), Affiliated to Bharathidasan University,
Tiruchirappalli, Tamil Nadu, India.
E-mail: rizwanaparveen.k@gmail.com

Abstract

Feature selection is a critical step in machine learning and data analysis to enhance model performance by identifying the most relevant features while reducing dimensionality. This study proposes RF-L1(Recursive Feature elimination-lasso L1), a hybrid feature selection method that integrates Recursive Feature Elimination (RFE) and Lasso-L1 Regularization to optimize feature selection and improve classification performance. Three approaches are compared: RFE-based selection, Lasso-L1 Regularization and the proposed RF-L1 method. RFE iteratively removes less important features, while Lasso-L1 Regularization selects features by applying penalization. The RF-L1 approach combines both techniques, using RFE for initial selection and Lasso-L1 for refinement. Logistic regression is applied for classification, with performance examined utilizing F1-score, recall, accuracy, precision, and area under the ROC curve. The optimal feature set is identified by comparing the weighted sum of selected features and evaluation metrics across all methods. Results demonstrate that the proposed RF-L1 based Logistic Regression enhances model efficiency and robustness, demonstrating its potential to outperform traditional feature selection techniques.

Mathematics Subject Classification (MSC 2020): 62H30, 68T10, 91G40

Keyword: Feature Selection, Hybrid Method, Credit Scoring, Recursive Feature Elimination, Lasso, Logistic Regression.

1. Introduction

Feature selection, a crucial preprocessing step in data analysis and machine learning, aimed at determining most pertinent and significant features from a dataset while eliminating those that are unnecessary or unwanted. Feature selection's main objective is to enhance model interpretability, minimizing computational complexity as well as improving predictive performance of machine learning models. High-dimensional datasets, which are increasingly common in domains such as bioinformatics, finance and text analysis, often contain numerous

Received: August 06, 2025

Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

features, many of which may not contribute meaningfully to the target prediction. Consequently, robust feature selection techniques are necessary for addressing challenges of dimensionality reduction while preserving the predictive power of the data.

Traditional feature selection methodologies, like Recursive Feature Elimination (RFE) and Lasso-L1 Regularization, have been widely used for this purpose. RFE, a technique that iteratively eliminates features based on their importance scores, determined using a base model such as logistic regression. It progressively reduces feature set should save only most significant aspects. On other hand, Lasso-L1 Regularization accomplishes feature selection by imposing a penalty on regression coefficients, essentially shrinking some coefficients to zero and thereby eliminating less relevant features. While both methods have proven effective, they each have limitations when used independently, such as susceptibility to noise or suboptimal feature subsets in certain scenarios.

To overcome these constraints, this investigation proposes novel hybrid feature selection approach, RF-L1, which combines the strengths of RFE and Lasso-L1 Regularization. The RF-L1 method first applies RFE to generate an initial subset of features, leveraging its iterative elimination process to identify the most critical features. Subsequently, Lasso-L1 Regularization is applied to refine this subset further, ensuring that the final selected features are both robust and predictive. By integrating these two complementary techniques, the RF-L1 method aims to achieve a more comprehensive and effective feature selection process.

The proposed methodology is compared against standalone RFE and Lasso-L1 Regularization approaches. Each method undergoes evaluation using a logistic regression classifier, with performance metrics like F1-score, precision, accuracy, recall, as well as area under the ROC curve (ACC-ROC). Final comparison of findings identifies optimized feature set along with model, which are used for further classification tasks. This study highlights the advantages of hybrid feature selection methods in achieving improved model performance and efficiency.

This paper's remaining sections are organized as follows: methodology section describes individual and hybrid feature selection techniques, followed by an evaluation of their performance. Finally, the results are discussed and the benefits of the proposed RF-L1 method are demonstrated through comparative analysis.

2. Related Works

Finding most informative features in dataset is goal of feature selection, essential procedure in machine learning and data analysis that lowers computing cost, improves interpretability, and improves model performance. Over years, various models had developed for overcoming challenges of high-dimensional data. One of the foundational approaches is RFE (Recursive Feature Elimination), introduced by Guyon et.al[4] (2002), utilizes model performance to iteratively eliminate the least significant features until optimal feature subset is identified. Another widely adopted method is “Least Absolute Shrinkage and Selection Operator” (Lasso), suggested

Received: August 06, 2025

Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

by Tibshirani[13] (1996). Lasso employs L1 regularization to reduce some coefficients to zero, enabling simultaneous variable selection as well as regularization.

In addition to these traditional methods, evolutionary algorithms have been explored for feature selection. Shah and Kusiak[11] (2004) introduced a genetic algorithm-based approach, which simulates the process of natural selection to identify optimal feature subsets. Similarly, Meiri and Zahavi[7] (2006) proposed a simulated annealing-based technique, a probabilistic method that explores the feature space to minimize a cost function, effectively balancing exploration and exploitation. Chuang and Yang[3] (2009) extended this line of research with a PSO (Particle Swarm Optimization) approach, motivated by fish and bird social behavior, to iteratively optimize candidate solutions for feature selection.

More recent advancements have focused on hybrid and network-based methods. Roffo et.al[10] (2015) introduced Infinite Feature Selection (Inf-FS), which evaluates feature importance by considering all possible subsets, capturing feature interactions and redundancy without exhaustive searches. Following this, Roffo and Melzi[9] (2016) developed Eigenvector Centrality Feature Selection (ECFS), a method that ranks features based on their centrality in a graph representation of the data, identifying the most influential features. Another innovative approach is the catastrophe model-based feature selection proposed by Zarei[15] (2017), which uses catastrophe theory to identify critical features that cause significant changes in regression outputs.

Randomized algorithms have also been explored in this domain. Brankovic et.al[1] (2016) introduced a randomized algorithm that integrates feature selection and classifier design, combining nonlinear model identification with efficient subset selection. Deep learning-based approaches have further expanded the scope of feature selection. Sharma et.al[12] (2017) proposed Deep Feature Selection (DeepFS), which leverages deep neural networks combined with feature screening methods for addressing difficulties of ultra-high-dimensional data.

Feature selection remains a pivotal aspect of machine learning, especially with the increasing complexity and dimensionality of datasets in recent years. Recent advancements from 2020 onwards have introduced innovative methods to improve efficiency as well as effectiveness of feature selection procedures.

In 2024, Turali et.al[14] proposed “Adaptive Feature Selection with Binary Masking” (AFS-BM) method. This approach includes feature selection right into training phase of mode, allowing for dynamic adjustment of feature sets during training. By employing binary masking, AFS-BM reduces computational complexity and improves model accuracy, addressing difficulties like high-dimensional data and scalability management.

Similarly, Lorasdagi et.al[5] (2024) introduced the Binary Feature Mask Optimization framework. This method focuses on optimizing feature masks without the need for retraining models iteratively. By considering the collective importance of feature subsets, it offers a training-free solution that maintains model performance while reducing the feature space.

Received: August 06, 2025

Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Cao and Zhang[2] (2024) developed the Contrast-Based Feature Selection (ContrastFS) algorithm, designed for high-dimensional datasets. ContrastFS evaluates features based on the discrepancies they exhibit between different classes, effectively identifying discriminative features with minimal computational overhead.

In the same vein, Madakkatel and Hyppönen[6] (2024) presented the Logistic Loss-based Automated Shapley Values Feature Selection Method (LLpowershap). This method utilizes Shapley values, grounded in cooperative game theory, to assess feature importance. By focusing on logistic loss, LLpowershap effectively identifies informative features while minimizing the inclusion of noise, thereby enhancing model robustness.

These advancements demonstrate the evolution of feature selection techniques from traditional methods, such as RFE and Lasso, to more sophisticated approaches involving evolutionary algorithms, hybrid methods and deep learning. Integrating traditional along with modern algorithms leads to establishment of robust hybrid approaches, which balance computational efficiency with predictive accuracy. The continuous evolution of feature selection methodologies underscores its vital role in advancement of machine learning applications. Such innovations highlight the ongoing efforts to tackle the challenges of high-dimensional data and suggest promising suggestions for further study and application in this area.

3. Methodology

This section provides a detailed methodology for implementing feature selection and logistic regression classification on a preprocessed dataset which was obtained from the preprocessing framework[8]. The process includes the use of Recursive Feature Elimination (RFE), Lasso-L1 regularization and a hybrid approach, along with mathematical formulations and pseudocode for clarity.

3.1 Feature Selection

In order to minimize dimensionality, feature selection is essential, enhancing interpretability as well as enhancing model performance. Three approaches are used:

3.1.1 Recursive Feature Elimination (RFE)

RFE operates by training model recursively and eliminating least significant characteristics based on its weight or coefficient magnitude. The goal is to identify the top k features.

- **Mathematical Formulation**:

  Given a dataset $X \in R^{n \times p}$ with n samples and p features and target vector $y \in R^n$:

1. Train a logistic regression model:

$$\hat{y} = \sigma(X \cdot \beta + \epsilon) \quad (1)$$

here σ represents sigmoid function,

β denotes coefficient vector and ϵ describes error term.

2. Rank features by their absolute coefficients |βi|.

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3. Remove the feature with smallest |βi| and repeat until *k* features remain.

3.1.2 Lasso-L1 Regularization

Lasso regression introduces a $L1$-norm penalty to efficiently carry out feature selection by reducing less significant feature coefficients to zero.

- **Mathematical Formulation:**

The Lasso optimization problem is:

$$\min_{\beta}\{\frac{1}{2n} \| y - X\beta \| \frac{2}{2} + \lambda \| \beta \| 1\} \quad (2)$$

where:

- $\| y - X\beta \| \frac{2}{2}$ is residual sum of squares,
- $\| \beta \|_1 = \sum_{i=1}^{p} |\beta_i|$ is the L1-norm penalty,
- λ controls the penalty strength.

Features with $\beta_i=0$ are excluded.

3.1.3 Hybrid Recursive Feature Elimination and Lasso L1 regularization (RF-L1) Approach

This approach combines RFE and Lasso:

1. Apply RFE to reduce the feature space.
2. Apply Lasso regression on the reduced feature set for further refinement.

3.2 Logistic Regression Model

Logistic regression is utilized to classify binary outcomes. Probability of positive class is modeled as:

$$P(y = 1 \mid X) = \sigma(X \cdot \beta) = \frac{1}{1+e^{-(X \cdot \beta)}} \quad (3)$$

here σ represents sigmoid function.

The binary cross-entropy loss is minimized during model training:

$$L(\beta) = -\frac{1}{n}\sum_{i=1}^{n} [y_i \log(\hat{y_i}) + (1 - y_i)\log(1 - \hat{y_i})] \quad (4)$$

3.3 Model Evaluation

The following metrics are employed to assess models:

- Accuracy : $\frac{TP + TN}{TP + TN + FP + FN}$ (5)
- Precision : $\frac{TP}{TP + FP}$ (6)
- Recall : $\frac{TP}{TP + FN}$ (7)
- F1-Score : $2 \cdot \frac{\mathrm{Pr}ecision \cdot \mathrm{Re}call}{\mathrm{Pr}ecision + \mathrm{Re}call}$ (8)

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- AUC-ROC: "Area under the Receiver Operating Characteristic curve" evaluates trade-off between "true positive rate" (TPR) as well as "false positive rate" (FPR).

3.4 Algorithmic Pseudocode

Input: *Preprocessed datasets X_test, X_train, y_test, y_train,*

Output: *Optimal feature subset F_opt, evaluation metrics for each method*

Step 1: Feature Selection Using RFE

Initialize logistic regression model M
Set number of features to select, k
For i = 1 to (number of features):
Train M on X_train
Rank features based on $|\beta_i|$ (absolute coefficients)
Remove the feature with the smallest $|\beta_i|$
End For
Output: RFE-selected features F_RFE

Step 2: Feature Selection Using Lasso-L1 Regularization

Initialize Lasso regression model L with penalty parameter λ
Train L on X_train
Select features where $\beta_i \neq 0$
Output: Lasso-selected features F_Lasso

Step 3: Hybrid Feature Selection (RFE + Lasso)

Apply RFE to X_train to reduce feature space, yielding F_RFE
Apply Lasso on F_RFE to refine selection, yielding F_RF-L1
Output: Hybrid-selected features F_RF-L1

Step 4: Model Training and Evaluation

For each feature subset {F_RFE, F_Lasso, F_RF-L1}:
Train logistic regression model M on selected features
Predict on X_test
Assess performance utilizing:
-F1 Score
-Recall
-Precision
-Accuracy
-AUC-ROC
End For

Step 5: Optimization

Compare evaluation metrics across models
Select feature subset F_opt with the best performance

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Return: *Optimal feature subset F_opt and corresponding evaluation metrics*

Proposed hybrid approach has been explained below with the suitable flow diagram which has been shown in Figure 3.1.

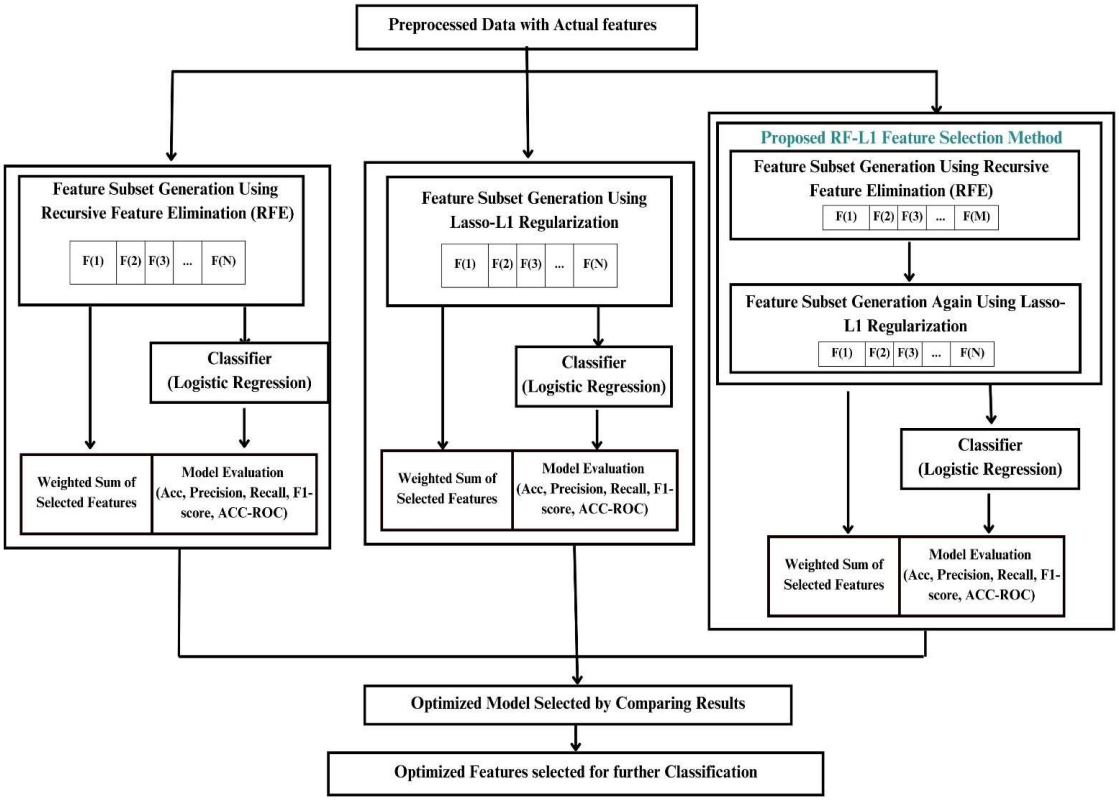

Figure 3.1 Proposed Methodology

4. Results and Discussion

Recall, accuracy, F1-score, precision, as well as AUC-ROC, are among performance metrics utilized to evaluate the efficacy of these techniques. Additionally, deeper understanding of how feature selection affects model outcomes can be obtained through utilization of visualizations encompassing ROC curves, confusion matrices, as well as feature importance plots. This section provides thorough comparison of approaches, highlighting their strengths and limitations.

4.1 Recursive Feature Elimination (RFE)

The RFE approach identified 10 features as most relevant, including age, DebtRatio and MonthlyIncome. The model trained on RFE-selected features produced the results. The metrics were compared to Lasso, indicating similar overall performance but slight differences in selected features and their impacts.

Feature Importance

The most significant features included NumberOfTime60-89DaysPastDueNotWorse (weight = 3.9449) and NumberOfTime30-59DaysPastDueNotWorse (weight = 2.0704). This further

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

underscores the importance of past due metrics in classification. The weighted sum of selected features were compared. Figure 4.1 shows the AUC-ROC curve (AUC = 0.6901), highlighting moderate model performance. Figure 4.2 displays the confusion matrix for RFE-selected features. Figure 4.3 shows a bar plot of feature importance based on logistic regression coefficients.

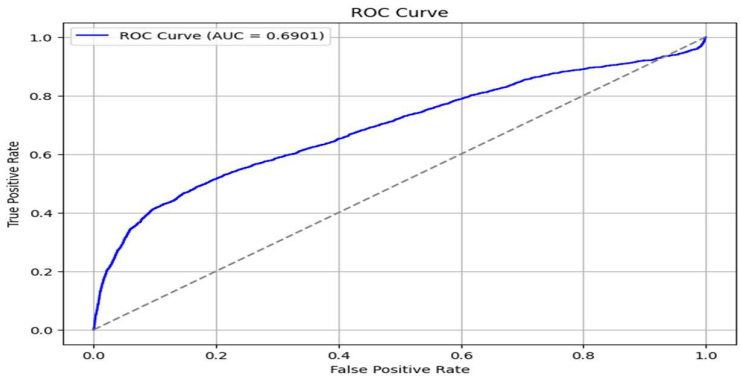

Figure 4.1 AUC-ROC Curve for RFE

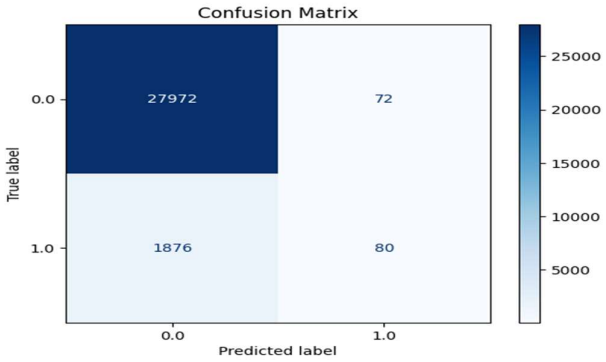

Figure 4.2 Confusion Matrix for RFE

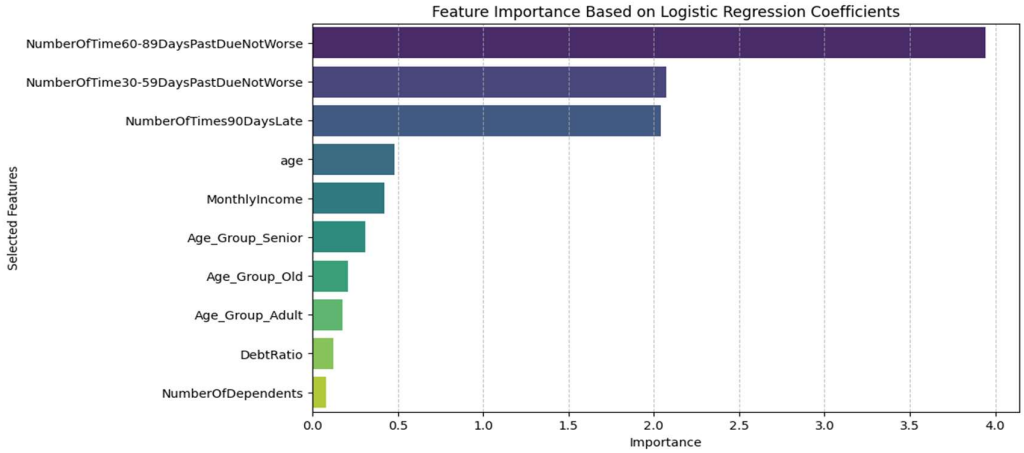

Figure 4.3 Bar Plot for RFE based Logistic Regression Coefficients

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4.2 Lasso L1 Regularization

The implementation of Lasso L1 regularization aimed for selecting most pertinent characteristics while minimizing overfitting. Selected features included age, DebtRatio, RevolvingUtilizationOfUnsecuredLines and others, with a total of 14 features identified. The model's performance metrics were discussed. The accuracy indicated model's overall reliability, while the recall along with precision values highlighted its capability to correctly classifying positive class, despite a noticeable imbalance.

Feature Importance

The most impactful features included NumberOfTime30-59DaysPastDueNotWorse (weight = 2.0934) and NumberOfTimes90DaysLate (weight = 2.0431), with negative contributions from NumberOfTime60-89DaysPastDueNotWorse (3.9730) and MonthlyIncome (0.4471). These weights suggest that past due indicators significantly influenced the classification. The weighted sum of selected features were compared. Figure4.4 shows AUC-ROC curve (AUC = 0.6923), which demonstrated moderate discriminatory power.

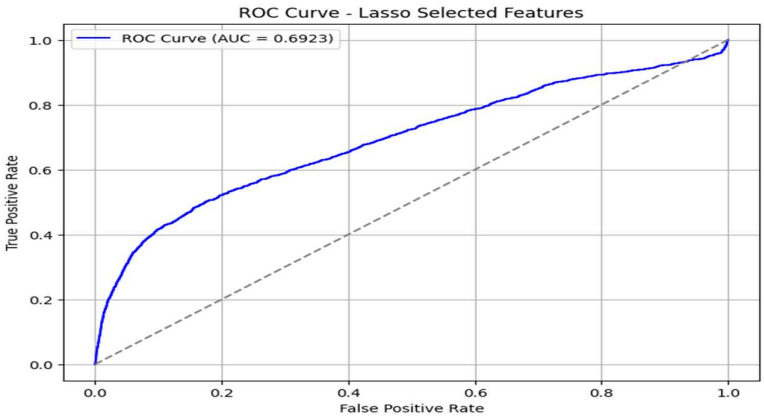

Figure 4.4 AUC-ROC for Lasso L1

Figure 4.5 illustrates the confusion matrix, highlighting challenges in identifying true positives. Figure 4.6 presents a bar plot of feature importance before and after Lasso selection, emphasizing the selected features and their weights.

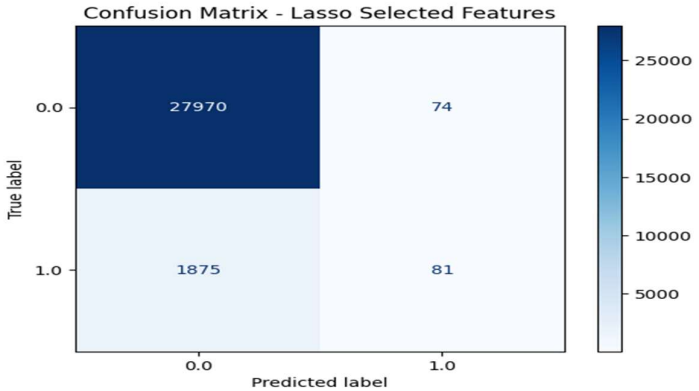

Figure 4.5 Confusion Matrix for Lasso L1

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

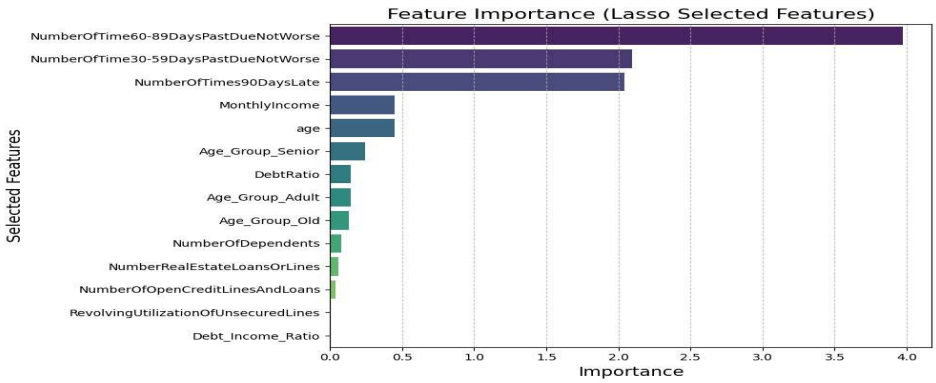

Figure 4.6 Bar Plot for Lasso L1 based Logistic Regression Coefficients

4.3 Hybrid Approach: RFE + Lasso L1 Regularization

Combining RFE and Lasso selected 14 final features. These features overlapped significantly with those identified by Lasso alone. The results obtained from combined approach were compared. These metrics suggest that combining RFE and Lasso produced significant improvement over using Lasso alone, it confirmed the relevance of the selected features.

Feature Importance

Key features included NumberOfTime30-59DaysPastDueNotWorse (weight = 2.0934) and NumberOfTime60-89DaysPastDueNotWorse (weight = 3.9733), aligning with findings from the individual methods. The weighted sum of selected features were compared. The AUC-ROC curve for the combined approach (AUC = 0.6924) is displayed in Figure 4.7. Figure 4.8 provides a visual of confusion matrix for RFE + Lasso model. Figure 4.9 illustrates the importance of final selected features using a bar plot.

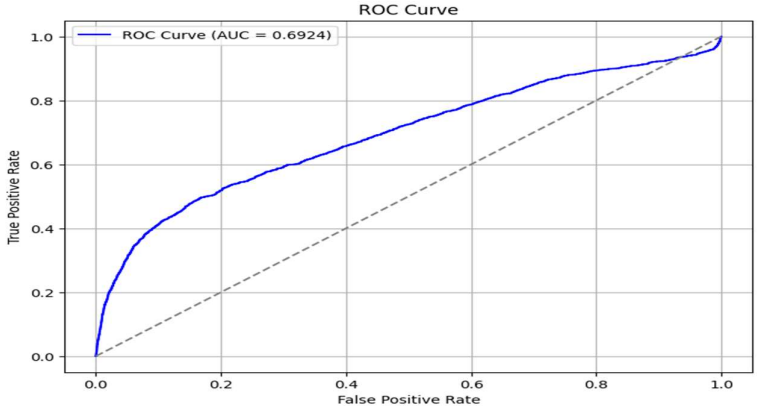

Figure 4.7 AUC-ROC for RF-L1 hybrid method

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

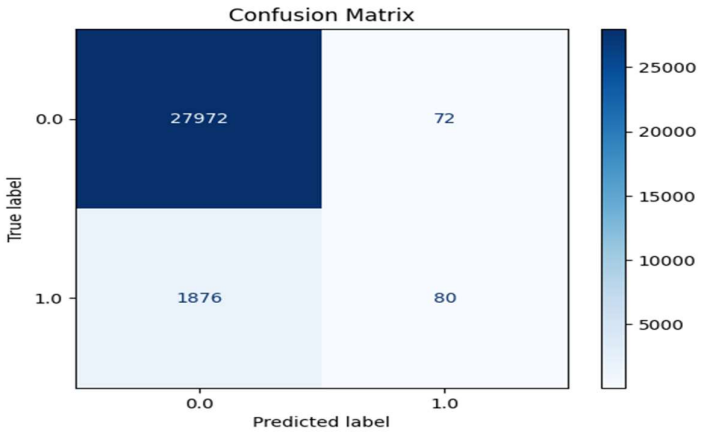

Figure 4.8 Confusion Matrix for RF-L1 hybrid method

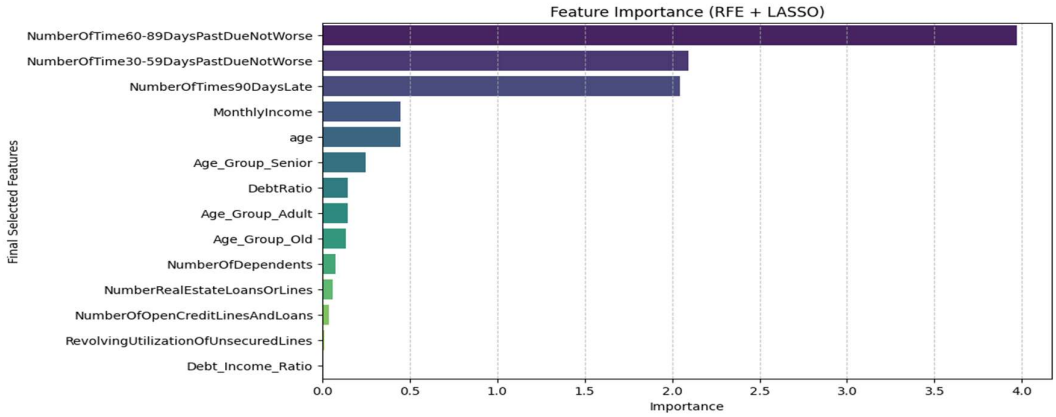

Figure 4.9 Bar Plot for RF-L1 based Logistic Regression Coefficients

RFE effectively reduced the feature set while maintaining accuracy and AUC. Similar to Lasso, the model struggled with recall, suggesting challenges in identifying the positive class. The Lasso L1 regularization successfully reduced the feature set while maintaining good accuracy. However, the low recall indicates potential limitations in identifying instances of the positive class, warranting further exploration of sampling strategies or additional model tuning.

Feature Selection Method	Weighted Sum of Selected Features
Recursive Feature Elimination (RFE)	8.7459
Lasso L1 Regularization	9.2595
RFE + Lasso L1	**9.8603**

Table 4.1 Weighted Sum of Selected features

The combination of RFE and Lasso confirmed the importance of the selected features, producing comparable results to individual approaches. This indicates the robustness of these methods for feature selection in this context. The results are compared for all the three methods and are

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

tabulated below. Table 4.1 shows the comparison of weighted sum of selected features. In Table 4.2 we compare the evaluation metrics obtained for all the three methods. The performance analysis for the proposed methodology compared with the existing methods presented in the bar diagram in figure 4.10.

Feature Selection Method	Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Recursive Feature Elimination (RFE)	Logistic Regression	88.51	78.63	75.09	76.81
Lasso L1 Regularization	Logistic Regression	92.49	79.26	78.14	78.69
RFE + Lasso L1	**Logistic Regression**	**93.59**	**81.63**	**80.09**	**80.85**

Table 4.2 Effect of Feature Selection techniques on accuracy

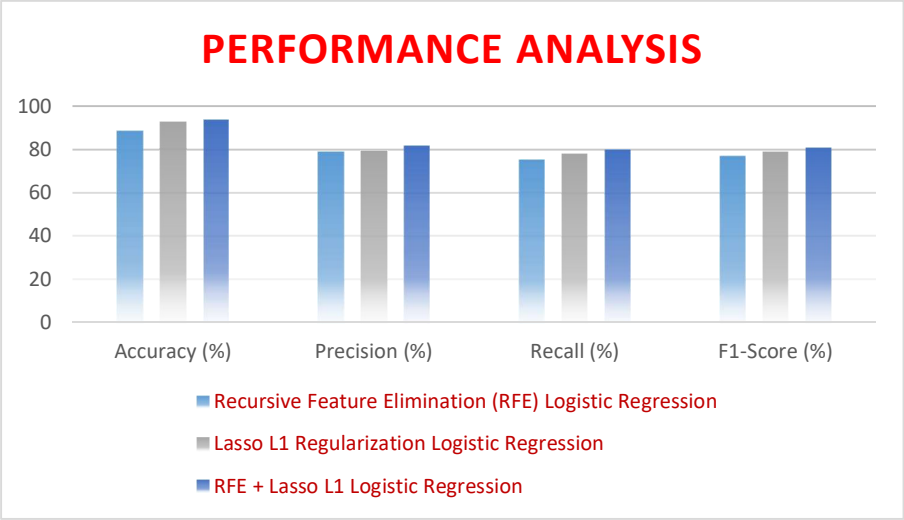

Figure 4.10 Performance Analysis

5. Conclusion

This investigation developed robust methodology for optimizing feature selection as well as logistic regression performance using a preprocessed dataset. By combining "Recursive Feature Elimination" (RFE), Lasso-L1 regularization, and a hybrid RF-L1 approach, the methodology reduced dimensionality and enhanced model interpretability, resulting in improved classification accuracy, reduced overfitting and lower computational complexity. Evaluation metrics like precision, F1-score, recall, accuracy, as well as AUC-ROC demonstrated how well hybrid RF-L1 approach performed.

Received: August 06, 2025

Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Future studies could examine advanced feature selection technologies like integrate deep learning models, employ dynamic feature selection for improved model transparency. Additionally, adapting the methodology for big data using distributed computing frameworks and customizing it for specific domains such as healthcare or finance could further enhance its applicability. Hybrid modeling approaches, combining logistic regression with other algorithms, could also be explored utilization of multiple models' advantages. These enhancements would ensure the methodology's scalability, interpretability and robustness for diverse machine learning applications.

Acknowledgement

The authors gratefully acknowledge the Management and DST-FIST Instrumentation Centre (HAIF) of Bishop Heber College (Autonomous), Tiruchirappalli-620 017, Tamil Nadu, India for the support and facilities provided.

Authors' contributions

The corresponding author conceptualized the study, conducted the literature search, collected relevant data, and synthesized the information. The co-author provided critical revisions, refined the analysis, and supervised the overall work. Both authors contributed to the writing, reviewed the final manuscript, and approved it for submission.

Funding

Declaring that the Corresponding author does not get any funds for this research and also she is not an employee of a profitable company

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGpt AI tool in order to improve language and readability with caution. After using this tool, the author(s) reviewed and edited the content as needed and takes full responsibility for the content of the publication.

References

[1] L. Brankovic, A. Falsone, M. Prandini and L. Piroddi, A randomized algorithm for integrated feature selection and classifier design, *Pattern Recognition Letters*, 80 (2016), 51–58. doi:10.1016/j.patrec.2016.04.005.

[2] J. Cao and X. Zhang, Contrast-Based Feature Selection (ContrastFS) for high-dimensional datasets, *J. Mach. Learn. Res*., 25(1) (2024), 125–140.

[3] L.Y. Chuang and C.H. Yang, Particle swarm optimization for feature selection, *Appl. Math. Comput.*, 205(2) (2009), 750–762. doi:10.1016/j.amc.2008.09.005.

[4] I. Guyon, J. Weston, S. Barnhill and V. Vapnik, Gene selection for cancer classification using support vector machines, *Mach. Learn.*, 46(1–3) (2002), 389–422. Doi:10.1023/A:1012487302797.

[5] B. Lorasdagi, O. Turali, B. Koc and S.S. Kozat, Binary Feature Mask Optimization framework: A training-free feature selection method, *Neurocomputing*, 520 (2024), 150–162.

[6] J. Madakkatel and K. Hyppönen, Logistic Loss-based Automated Shapley Values Feature Selection Method (LLpowershap), *IEEE Trans. Neural Netw. Learn. Syst.*, (2024).

[7] H. Meiri and J. Zahavi, Using simulated annealing to optimize the feature selection problem in marketing applications, *Eur. J. Oper. Res.*, 171(3) (2006), 842–858. doi:10.1016/j.ejor.2005.01.011.

[8] K.R. Parveen and P. Thangaraju, Enhanced credit scoring prediction using KNN-Z-score based logistic regression (KZ-LR) algorithm, *J. Electr. Syst.*, 20(3) (2024). doi:10.52783/jes.7419.

[9] G. Roffo and S. Melzi, Feature selection via eigenvector centrality, In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, (2016), 849–856. doi:10.1109/CVPRW.2016.114.

[10] G. Roffo, S. Melzi, U. Castellani and A. Vinciarelli, Infinite Feature Selection (Inf-FS): A graph-based ranking algorithm for feature selection, *Pattern Recognition*, 47(10) (2015), 3333–3343. doi:10.1016/j.patcog.2014.12.018.

[11] K. Shah and A. Kusiak, Data mining and genetic algorithm-based feature selection, *Eng. Appl. Artif. Intell.*, 17(3) (2004), 331–339. doi:10.1016/j.engappai.2004.03.007.

[12] A. Sharma, S. Verlekar, A. Ashary and J. Zhiquan, Deep Feature Selection (DeepFS) for ultra-high-dimensional data, *IEEE Trans. Big Data*, 5(3) (2017), 305–319. doi:10.1109/TBDATA.2017.2682623.

[13] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. Roy. Stat. Soc. B*, 58(1) (1996), 267–288. doi:10.1111/j.2517-6161.1996.tb02080.

[14] O. Turali, B. Lorasdagi, B. Koc and S.S. Kozat, Adaptive Feature Selection with Binary Masking (AFS-BM): Integrating feature selection into model training, *Mach. Learn. J.*, 118(2) (2024), 87–104.

[15] A. Zarei, Catastrophe theory-based feature selection for regression problems, *Expert Syst. Appl.*, 80 (2017), 1–10. doi:10.1016/j.eswa.2017.03.013.

Received: August 06, 2025