ASEAN
Journal of Scientific and Technological Reports
Online ISSN:2773-8752

Research article

Machine Learning Integration in Feature Selection Using Sequential Forward Floating Selection (SFFS) for Improved Credit Scoring

Dawn Iris Calibo-Senit, Clifford I. Senit[2], and Dawn Angelic Ramizo[3]

[1] College of Technology, Siquijor State College, Philippines
[2] Engine Department, Internship Crew Philippines, Philippines
[3] Velera-Credit Union Service Organization, United States of America
* Correspondence: dawniriscalibo@gmail.com

Citation:
Calibo-Senit, DI., Senit C., Ramizo, DA.; Machine learning integration in feature selection using sequential forward floating selection (SFFS) for improved credit scoring. *ASEAN J. Sci. Tech. Report.* **2026**, *29*(10), e261976. https://doi.org/10.55164/ajstr.v29i10.261976

Article history:
Received: October 20, 2025
Revised: March 2, 2026
Accepted: March 11, 2026
Available online: September 20, 2026

Publisher's Note:
This article is published and distributed under the terms of the Thaksin University.

Abstract: Credit risk evaluation serves as the backbone of any financial decision, especially in government funding programs like DOST SET-UP. However, conventional credit assessment methods often face problems such as redundant features, low prediction accuracy, and increased computational complexity, which may result in suboptimal funding decisions. This study presents an integrated machine learning framework that incorporates Sequential Forward Floating Selection (SFFS) to optimize feature selection in credit assessment. This study considered financial and operational parameters from historical data on 33 DOST SET-UP projects undertaken from 2008 to 2016. A total of twelve parameters were initially considered in the feature selection process, which was later reduced to four highly significant features using the SFFS algorithm in Python: the Debt to Asset Ratio (DAR), the Net Profit Margin (NPM), the Liquidity Ratio (LR), and the Return on Investment (ROI). These features were later incorporated into the development of the rule-based fuzzy logic system using the MATLAB platform to produce the final recommendation for the risk assessment of the proposals. The experimental results show that feature selection greatly improved the model's prediction accuracy, and the system achieved an overall accuracy of 94%. The results verified that integrating machine learning-based feature selection and fuzzy logic can develop an efficient and effective credit scoring model for the government's lending programs and improve financial sustainability for micro, small, and medium enterprises.

Keywords: Credit scoring; feature selection; sequential forward floating selection (SFFS); risk analysis

1. Introduction

Credit scoring is critical in financial decision-making, particularly in weighing the creditworthiness of loan applicants. Because all financial institutions and government lending programs aim to maximize reasonable grant distribution while minimizing risk, there is an increasing need for accurate and efficient credit scoring models. However, traditional methods of assessing credit risk often struggle with feature-selection redundancy and model efficiency. As a result, decisions may be suboptimal, leading to higher default rates. In response to these challenges, credit risk analysis has widely adopted

machine learning techniques. Among these, feature selection helps improve model accuracy and reduce computational complexity by identifying the most relevant variables for prediction. Another effective way is Sequential Forward Floating Selection. This is an iterative feature-selection technique that improves variable selection by gradually adding and removing features to boost predictive performance. By integrating SFFS into credit scoring models, this research aims to improve feature selection by ensuring that only the key variables support the risk evaluation system.

This research examines the incorporation of SFFS with machine learning models in the credit scoring process to improve efficiency and accuracy. In this way, the risk assessment strategies of various financial institutions and government lending agencies would be enhanced, along with reduced loan delinquency rates and more efficient allocation of financial resources. More specifically, the findings will help create a stronger, verifiable credit evaluation system that supports financial stability and inclusivity for the DOST SET-UP. On the verge of the global financial crisis, credit risk assessment has become a critical component of financial decision-making. Financial grants should be given to creditworthy applicants to ensure the sustainability and resilience of micro, small, and medium enterprises. The DOST-SETUP plays a critical role in supporting MSMEs through technological innovations that increase productivity and competitiveness in the market. The program provides aid through technology acquisition, technical consultancy, training, product development, and financial aid, all of which address critical challenges that restrict business growth and sustainability [1]. To fully harness the potential of financial assistance, a technology-based credit risk assessment tool must be implemented to ensure grants are provided only to those with high creditworthiness [2].

Various credit risk evaluation models have been recommended to improve decision-making [3]. One critical challenge in predicting FDP is applying appropriate FS algorithms, which significantly affect the predictive models' accuracy [4]. So far, research on predicting monetary distress has also been highly active in statistics and machine learning. Essentially, predicting insolvency starts with selecting the most relevant features, reducing complexity, and improving model performance. Researchers have identified feature selection techniques and predictive models that can enhance FDPs. Yang et. al. [5] monitored feature selection through PCA and FA in an analysis of credit rating. This analysis showed the effectiveness of FA in reducing redundancy while retaining a high predictive performance for FDP. Similarly, Ding and Yan [6] proposed a model integrating Maximum Relevance Minimum Redundancy (MRMR) and Support Vector Machine Recursive Feature Elimination (SVM-RFE), which significantly improved financial distress prediction accuracy compared to traditional models.

Wu et al. [7] proposed an ensemble feature selection method combined with an improved stacking algorithm that used mutual information, analysis of variance (ANOVA), random forests, genetic algorithms, and recursive feature elimination. Their results showed significant improvements in FDP accuracy. In addition, Zhang et. al. [8] developed an ANN model that used a two-stage feature selection process with multiple filters and a wrapper method, thus addressing the "curse of dimensionality" and enhancing predictive capabilities. Santoso and Wibowo [9] compared LDA and SVM model variants for financial distress prediction and concluded that the Stepwise-SVM hybrid model achieved a better balance among model accuracy, generalization, and stability.

Feature selection has been utilized in different domains other than financial risk evaluation. Palongpalong et al. [10] targeted the identification of key features in blended learning when developing educational assessment software; thus, they ensured it would be implemented efficiently and comprehensively. Computerized mass classification [11] applied an SFFS- and SVM-based model to optimize medical research.

This means that selecting appropriate features is important for creditworthiness prediction because it directly influences accuracy and reliability. Studies show that the latest feature selection techniques can greatly enhance the efficiency of FDP models. With these methodologies, researchers can create accurate and reliable credit risk assessment tools. In this regard, financial grants and loans would be given to the most creditworthy MSMEs.

2. Materials and Methods

2.1 Research Design

The study employed a quantitative, data-driven approach, integrating data mining, feature selection, rule-based fuzzy logic modeling, and model evaluation. This was done by analyzing historical data from funded and unfunded projects to identify the most significant predictors of creditworthiness. This methodological workflow is illustrated in Figure 1.

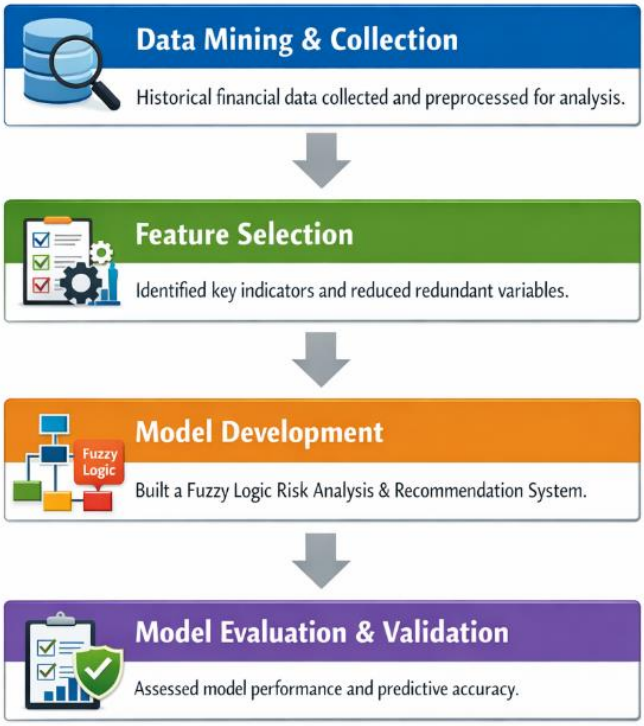


Figure 1. Structured Quantitative Data-Driven Approach

The framework adopts a data-driven pipeline that is structured to build and develop a risk analysis and recommendation model. The framework begins with data mining and collection, where historical financial data are extracted from relevant sources and undergo preprocessing techniques, including data cleaning, normalization, and transformation. This phase builds a credible dataset that captures the essential financial indicators required for predictive modeling. The next phase of the framework involves feature selection, which aims to identify the most significant financial indicators that influence risk prediction. Statistical and machine learning-driven approaches are used to eliminate redundant and irrelevant variables, enhancing model efficiency by reducing dimensionality. The framework adopts this approach to improve the model's interpretability and efficiency. During model development, the identified variables are used to design a Fuzzy Logic-based Risk Analysis and Recommendation System. Fuzzy logic is utilized to represent the uncertainties and vagueness associated with financial decision-making. Membership functions and rule-based reasoning systems are designed to represent the uncertainty and vagueness associated with financial decision-making. Finally, during model evaluation and validation, the efficiency of the developed model is assessed. Performance measures and validation processes are used to evaluate the system's accuracy, reliability, and consistency. This phase ensures that the proposed framework can provide accurate risk predictions and be used to make financial decisions.

2.2 Data Source and Description

The dataset consisted of historical financial records of DOST SET-UP Siquijor projects from 2008 to 2016, covering 33 project proposals. Each record represented a single project and included financial ratios, absolute financial values, operational indicators, and a class label.

2.2.1 Predictor Variables

Table 1. Feature Index Representation of the Dataset in Python

feature id	feature name
0	Debt-to- Asset Ratio
1	Net Profit Margin
2	Liquidity Ratio
3	Return on Investment
4	Asset
5	SET-UP Liability
6	Total Asset
7	Total Debt
8	Net Profit
9	Revenue
10	Year
11	Success rate on Objective Attainment
12	Success rate on Timely Payment

Table 1 shows the dataset used for simulation, based on the financial attributes from the data source. In the Python implementation, the attributes are converted into a formal feature list where each column is labeled using a numeric index and a feature name. The representation makes it easier for the Sequential Forward Floating Selection (SFFS) algorithm to analyze feature combinations and select the best-performing feature set for the DOST SET-UP Risk Analysis and Recommendation System. More specifically, the dataset includes financial features such as Debt-to-Asset Ratio, Net Profit Margin, Liquidity Ratio, Return on Investment, Asset, SET-UP Liability, Total Asset, Total Debt, Net Profit, Revenue, and Year, as well as operational attributes such as Success Rate on Objective Attainment and Success Rate on Timely Payment. These attributes are used as candidate features from which the algorithm selects the most significant variables for credit risk analysis. Using feature identifiers, the system allows for the systematic testing of the presence and absence of each attribute during the forward and backward phases of the SFFS process. This allows the simulation to determine the best-performing subset of financial attributes that significantly impact credit-risk prediction and to develop recommendations for the DOST SET-UP program. After determining the feature set described in Table 1, the dataset is imported into the Python simulation environment, where the SFFS algorithm analyzes different feature combinations and provides the final results for the feature selection process. This dataset is then loaded into Python to perform the simulation and generate the results for the feature selection.

2.3 Data Preprocessing

In this stage, data preprocessing was performed in Python before modeling. This includes removing non-numeric attributes, validating and correcting data types, eliminating inconsistent entries, and retaining numerical features required for machine learning. After the said process, the dataset consisted of numerical predictors and one categorical class label suitable for supervised learning.

2.4 Data Mining Tool

The research utilizes a data analysis and visualization software tool known as Tableau to extract insights from the data and display them in a graphical form. Tableau processed historical data on project proposals, whether accepted or implemented, over a period of nine (9) years from 2008 to 2016. Tableau also performed data-cleaning checks, trend analysis, and pattern identification before machine learning modeling. The application identified data distribution, anomalies, and relationships between key financial variables. The results obtained from Tableau were then processed using Python for further analysis to prepare the data for feature selection and credit risk classification. Using Tableau improved data interpretation and ensured the

financial data underlying the research was visually interpretable, consistent, and valid for more complex predictive modeling.

2.5 Feature Selection Method

To build the model, the system used Simplest Sequential Selection, which iteratively adds (or deletes) one feature at a time to maximize the intermediate criterion value until the required dimensionality is met. Figure 2 illustrates the SFFS process.

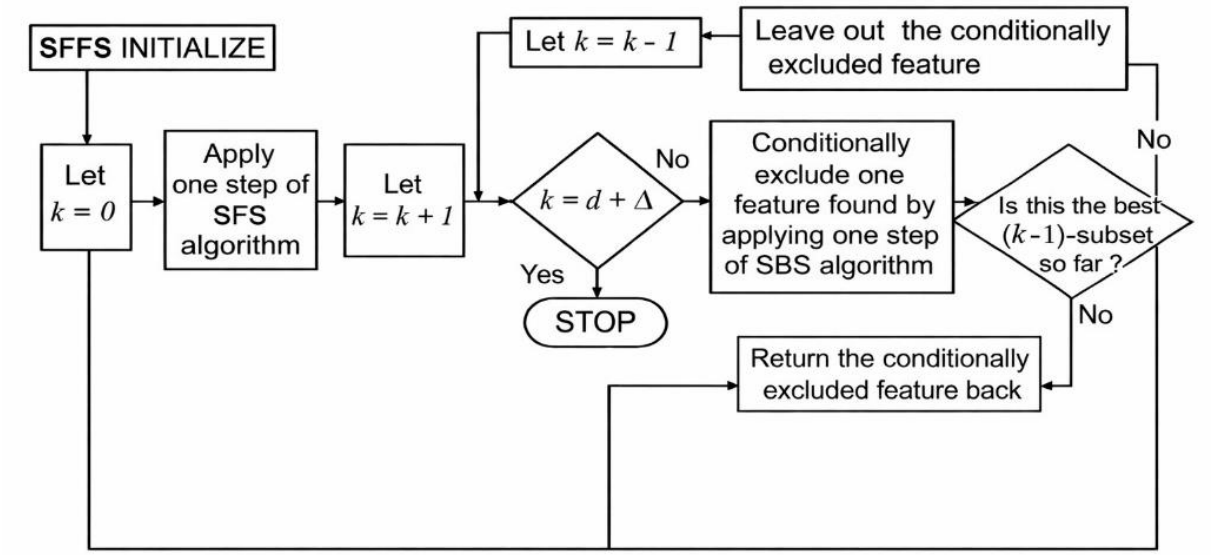


Figure 2. Sequential Forward Floating Selection (SFFS) Process Diagram. Sequential Forward Floating Selection Algorithm adopted from Somol, Novovicova, and Pudil (2010)

2.5.1 Floating Selection Algorithm Analysis

The figure shows the Sequential Forward Floating Selection (SFFS). The process consists of applying several backward steps after each forward step, as long as the resulting subsets are better than the previously evaluated ones at that level. Consequently, no backward steps occur if intermediate results at a given level cannot be improved. The same applies to the backward version of the procedure. The algorithms allow self-controlled backtracking, so they can eventually find good solutions by dynamically adjusting the trade-off between forward and backward steps. After each forward set, multiple backward steps are applied to the resulting 33 subgroups, which outperform the previously examined ones at that level. As a result, no backward steps are taken if intermediate results at a specific level cannot be improved. The same holds for the reverse-form techniques. The algorithms support self-controlled backtracking, allowing them to eventually find good solutions by dynamically altering the trade-off between forwarding steps. Setting SFFS yields a subset of d features, with the optional search-restricting parameter €[0, Dd]:

1. Start with *X* O =O*k* = 0
2. X k + 1 = ADD (*X k), k =k*+1
3. Repeat *X k-1* = REMOVE *(X k), k = k-1*, as long as it improves solutions already known for the lower *k*. (1)
4. If k < d+ Δ go to 2.

Using the Python programming language, the data were imported to start the sequential feature selection process, specifically sequential forward selection. Since the system needed numerical data for the credit risk analysis, all non-numerical data were removed from the dataset. Thus, we retained 12 constructs and one class, which is the remark stating whether the project status is completed or ongoing. This is presented in the table below.

Table 2. Numerical Financial Features Extracted from the Dataset

	DAR	NPM	LR	ROI	Asset	Liability	Total Asset	Total Debt	Net Profit	Revenue	Year	SOT	SRTLP	Remarks
0	0.55	0.45	1.20	37.88	300000.0	250000.0	550000.0	300000.0	208333.33	458333.33	2008	100	100	complete
1	0.57	0.49	1.33	40.46	600000.0	450000.0	1050000.0	600000.0	425000.00a	675000.00	2008	100	100	complete

Table 2 presents the cleaned numerical financial data used as input for the Sequential Forward Selection (SFS) method in the credit risk analysis. Before feature selection, the data was prepared in Python by removing non-numeric attributes, examining data types, and verifying numerical consistency to input the data directly into machine learning algorithms. Each entry represents a distinct borrower or project. For instance, Entry 0 has a Debt-to-Asset Ratio (DAR) of 0.55, indicating that 55% of assets are debt-financed, and a Net Profit Margin (NPM) of 0.45, indicating moderate profitability. A Liquidity Ratio (LR) of 1.20 indicates sufficient short-term assets to cover current liabilities. The Return on Investment (ROI) is 37.88%, which is a strong indicator of investment success. For Entry 0, ₱300,000 in assets and ₱250,000 in liabilities translate to ₱550,000 in total assets and ₱300,000 in total debt. Net profit is ₱208,333.33, with revenue of ₱458,333.33, further supporting a sound financial position. The performance metrics show 100% for both Success Rate on Objective Attainment (SROT) and Success Rate on Timely Loan Payment (SRTLP), indicating that the project is marked as "complete." Record 1 shows higher leverage, with a DAR of 0.57, and improved profitability, with an NPM of 0.49. The LR of 1.33 indicates improved short-term financial performance compared to Record 0. The ROI improves to 40.46%, indicating highly favorable returns. On the financial side, the borrower has ₱600,000 in assets and ₱450,000 in liabilities, with total assets of ₱1,050,000 and total debt of ₱600,000. Net profit increases to ₱425,000 with ₱675,000 in revenues. Both SROT and SRTLP remain at 100%, and the project is also marked as "complete." For the sample as a whole, the perfect SROT and SRTLP indicate that these borrowers have perfectly accomplished the project requirements and loan repayment terms. The ROI, which varies from 37.88% to 40.46%, and the liquidity ratios above 1.0 indicate that the projects are financially sound. The DAR values slightly above 0.50 indicate moderate leverage, which may still be acceptable depending on the credit policy.

3. Results and Discussion

3.1 Data Simulation Results in Python

After the DOST SET-UP Siquijor dataset from 2008-2016 was imported into Python, the SFFS algorithm was applied. To select features for the system, we adopted Simplest Sequential Selection, iteratively adding (or removing) one feature at a time to maximize the intermediate criterion value until the required dimensionality is achieved. The process starts by selecting the "best" 8 features from the *Iris dataset via Sequential Forward Selection (SFS*). Here, set forward-True and floating-false. By choosing ev=0, cross-validation is not performed; therefore, the performance (here: 'accuracy') is computed entirely on the training set. Below are the resulting iterations out of 12 features and one class:

Table 3. Data Simulation to Python during its 1st iteration out of 8 features using the Sequential Backend with 1 concurrent worker

Step	Process Description	Status	Score
1	Initialization of the feature selection process using sequential backend with a single worker	Started	–
2	Initial parallel processing completed	Completed	–
3	Completion of the evaluation of candidate feature subsets	Completed	–
4	Selection of the first optimal feature subset (1 out of 8 features)	Processing Completed	0.975

Table 3 summarizes and reformats the execution steps of the Sequential Forward Floating Selection (SFFS) algorithm implemented in Python, showing the progression of the first optimal feature subset (1 out of 8 features) and an evaluation score of 0.975.

Table 4. Data Simulation in Python during its 2nd iteration out of 8 features using the Sequential Backend with 1 concurrent worker

Step	Process Description	Status	Score
1	Initial parallel evaluation completed (1 of 1 task)	Completed	–
2	Evaluation of feature subsets completed (12 of 12 tasks)	Completed	–
3	Selection of the optimal subset with 2 out of 8 features	Processing Completed	1.000
4	Final parallel validation of the selected feature subset	Completed	–

Table 4 illustrates the Sequential Forward Floating Selection (SFFS) process. It shows that the system tested various feature sets and ultimately selected the best pair of features from the eight features, resulting in a performance score of 1.000.

Table 5. Data Simulation to Python during its 3rd iteration out of 8 features using the Sequential Backend with 1 concurrent worker

Step	Process Description	Status	Score
1	Initial parallel evaluation completed (1 of 1 task)	Completed	–
2	Evaluation of feature subsets completed (11 of 11 tasks)	Completed	–
3	Selection of the optimal subset with 3 out of 8 features	Processing Completed	0.975
4	Final parallel validation of the selected feature subset	Completed	–

Table 5 shows how the Sequential Forward Floating Selection (SFFS) algorithm works using Python code. As shown below, the system evaluates the feature set, and by the third step, the algorithm selects 3 of the 8 features. This yields a performance score of 0.975, illustrating how the algorithm identified the key feature set for the credit risk assessment process.

Table 6. Data Simulation to Python during its 4th iteration out of 8 features using Sequential Backend with 1 concurrent worker

Step	Process Description	Status	Score
1	Initial parallel evaluation completed (1 of 1 task)	Completed	–
2	Evaluation of feature subsets completed (10 of 10 tasks)	Completed	–
3	Selection of the optimal subset with 4 out of 8 features	Processing Completed	0.975
4	Final parallel validation of the selected feature subset	Completed	–

This table shows the fourth step of the Sequential Forward Floating Selection (SFFS) algorithm, implemented in Python. This shows that the system has tested different feature sets, and the best feature set, consisting of four out of the eight features, has been selected with a score of 0.975.

Table 7. Data Simulation to Python during its 5th iteration out of 8 features using the Sequential Backend with 1 concurrent worker

Step	Process Description	Status	Score
1	Initial parallel evaluation completed (1 of 1 task)	Completed	–
2	Evaluation of feature subsets completed (12 of 12 tasks)	Completed	–
3	Selection of the optimal subset with 5 out of 8 features	Processing Completed	1.000
4	Final parallel validation of the selected feature subset	Completed	–

Table 7 shows step five of the Sequential Forward Floating Selection (SFFS) algorithm, implemented in Python. In this step, the algorithm considered various feature subsets and selected the optimal feature set, which comprised five out of the eight features, resulting in a perfect score of 1.000. This shows the progressive identification of the most prominent features for credit risk evaluation.

Table 8. Data Simulation in Python during its 6th iteration out of 8 features using the Sequential Backend with 1 concurrent worker

Step	Process Description	Status	Score
1	Initial parallel evaluation completed (1 of 1 task)	Completed	–
2	Evaluation of feature subsets completed (8 of 8 tasks)	Completed	–
3	Selection of the optimal subset with 6 out of 8 features	Processing Completed	1.000
4	Final parallel validation of the selected feature subset	Completed	–

This table describes Step 6 of the Sequential Forward Floating Selection (SFFS) algorithm. It indicates that six of the eight features were selected as the best subset, resulting in a performance score of 1.000. This shows the algorithm's ability to identify the best features for credit risk assessment sequentially.

Table 9. Data Simulation in Python during its 7th iteration out of 8 features using the Sequential Backend with 1 concurrent worker

Step	Process Description	Status	Score
1	Initial parallel evaluation completed (1 of 1 task)	Completed	–
2	Evaluation of feature subsets completed (7 of 7 tasks)	Completed	–
3	Selection of the optimal subset with 7 out of 8 features	Processing Completed	1.000
4	Final parallel validation of the selected feature subset	Completed	–

Table 9 shows the seventh step of the Sequential Forward Floating Selection (SFFS) algorithm, implemented in Python, with the best features selected consisting of seven of the eight features, resulting in a performance score of 1.000.

Table 10. Data Simulation in Python during its 8th iteration out of 8 features using the Sequential Backend with 1 concurrent worker

Step	Process Description	Status	Score
1	Initial parallel evaluation completed (1 of 1 task)	Completed	–
2	Evaluation of feature subsets completed (6 of 6 tasks)	Completed	–
3	Selection of the optimal subset with all 8 features	Completed	1.000
4	Final parallel validation of the selected feature subset	Completed	–

For the DOST SET-UP program, the Sequential Forward Floating Selection (SFFS) approach was employed step by step to determine the best financial feature set that would significantly influence the credit risk evaluation process. Over eight steps, the process evaluated feature sets ranging from 1 to 8 features and assessed the performance of each selected set. The results show an increasing performance trend, culminating in a perfect score of 1.000 with the feature set containing all eight features. This demonstrates the effectiveness of the SFFS feature selection method in identifying the best feature set and supports a data-driven approach to a trustworthy credit scoring process. This means that feature selection eliminates irrelevant candidate features, thereby decreasing the complexity of the DOST SET-UP risk analysis and recommendation model. Figure 3 visualizes this.

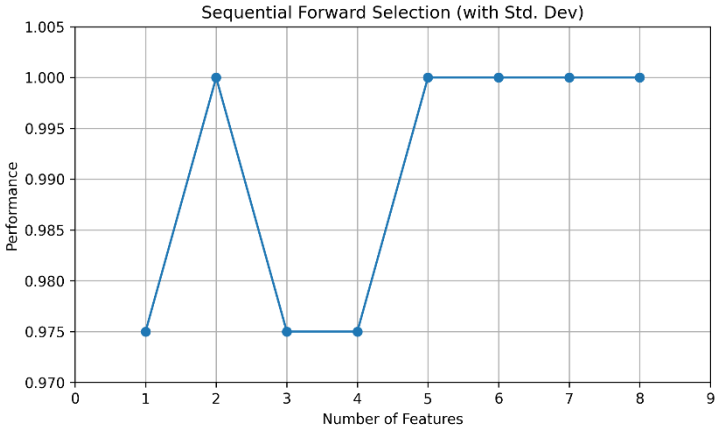


Figure 3. Selected Features using Sequential Forward Selection Result

The prompt below shows that the algorithm adds features one at a time and computes accuracy iteratively. Among the 12 features, only eight are chosen. This shows that the first feature has an accuracy rate of .975, or 97.5%, followed by .1, or 100%. However, the graph dropped twice for features 3 and 4 and repeatedly returned to an accuracy rate of .975, or 97.5%. Positively, the accuracy rate was high enough for the last four variables, reaching .975 (97.5%), followed by .1 (100%). This supports the claim that feature selection improves model performance, avoids overfitting, reduces training and testing time, and lessens measurement and storage requirements. Thus, with accuracies of 97.5% to 100%, the selected features are chosen. The result further shows that these features are the following, as displayed in Python:

```
# Which features?
feat_cols = list(sfs1.k_feature_idx_)
print(feat_cols)
```

(2)

```
[0, 1, 2, 3, 6, 9, 10, 12]
```

This shows the selected features by feature ID. Specifically, 0, 1, 2, 3, 4, 5, 6, 9, 10, and 12 in Table 3.4. The results show that the eight selected features are DAR, NPM, LR, ROI, Total Asset, Revenue, Year, and Success rate on Objective Attainment. To treat the data, only the top four features are included in the system since they are financial ratios. Specifically, these are the top four in the ranking.

Table 11. Selected features resulting from the DOST SET-UP historical data

Feature ID	Selected Features	Rank
0	Debt-to- Asset Ratio	1
1	Net Profit Margin	2
2	Liquidity Ratio	3
3	Return on Investment	4
6	Total Asset	5
9	Revenue	6
10	Year	7
11	Success rate on Objective Attainment	8

The top four selected features are all financial ratios. DAR, NPM, LR, and ROI are mathematical comparisons of financial statement accounts or categories. These relationships between financial statement accounts help financial grantors, creditors, and internal company management understand how well a business is performing and where it needs improvement. For features in ranks 5 and 6, such as total assets and revenue, they are both composed of those ratios. For rank 8, which is the year, this feature is not very relevant to include since the DOST-SET-UP is availed yearly, while the success rate on objective attainment can only

be used when the project is already completed. These ratios are calculated using the statistical formulas indicated in equations 5, 6, 7, and 8.

3.2 Mathematical Formulation of Selected Ratios

The four financial ratios were calculated using the following mathematical formulas to determine the financial performance and risk level of the selected firms in the dataset. The four selected financial ratios were computed as follows:

Computation of Net Profit Margin:

$$\boldsymbol{NMP} = \frac{\text{Net Profit}}{Total\ Revenue} \tag{3}$$

NPM is the ratio that shows the percentage of profit earned on the total revenue generated by the company. It shows how efficiently the company converts revenue into profit. If NPM is high, the company is more efficient in its operations and has a high profitability level. On the other hand, if the NPM is low, it means the company has high operating costs.

Computation of Return on Investment:

$$\boldsymbol{ROI}(\%) = \frac{\text{Net ProfitNet Profit (or Gain from Investment)}}{Cost\ of\ Investment}\boldsymbol{x}100 \tag{4}$$

ROI is a ratio that shows the profitability of the investment compared with the total cost incurred. It shows the amount of return generated on each investment made by the company. If the company has a high ROI, it means the company is more efficient in its investment and has a high level of financial performance.

Computation of Liquidity Ratio:

$$\boldsymbol{LR} = \frac{\text{Curernt Asset}}{Current\ Liability} \tag{5}$$

LR is the ratio that shows the SMSE's efficiency in meeting its financial obligations using its assets. If the liquidity ratio is more than 1, it means the business can easily pay its financial obligations using its assets.

Computation of Debt to Asset Ratio:

$$\boldsymbol{DAR} = \frac{\text{Total Liabilities}}{Total\ Assets} \tag{6}$$

DAR measures the portion of a business's assets financed through debt. It shows the amount of financial leverage employed by the establishment. The higher the DAR, the higher the risk since the proportion of debt employed is higher.

The results from the risk analysis and recommendation system are used as variables to compute the financial ratio values. The results are interpreted using a business intelligence tool to identify the enterprise's risk category based on the financial indicators.

3.3 Comparative Performance Analysis With and Without Feature Selection

To assess the contribution of Sequential Forward Floating Selection (SFFS) to the proposed credit scoring framework, a comparative analysis is conducted between two models: one developed using all features and the other developed using the optimized set of features selected through SFFS. First, the credit risk evaluation system is implemented using all the features included in the set of twelve variables. Although this approach enables the system to incorporate a broad range of information, it also introduces feature redundancy, complicating the system's evaluation process and reducing the interpretability of the fuzzy rule base. Upon applying SFFS, the system is optimized for a reduced set of four essential financial ratios: Debt-to-Asset Ratio (DAR), Net Profit Margin (NPM), Liquidity Ratio (LR), and Return on Investment (ROI). These four ratios encompass the essential statistical characteristics of a firm's leverage position, profitability, liquidity, and investment efficiency.

3.4 Model Development: Rule-Based Fuzzy Logic

3.4.1 Variable Ranges

For the study, the researcher adopted a structured business tool from Northwest Farm Credit Services (NFCS, 2015) to identify the financial ratios that measure the performance of each feature used in the algorithm simulation. This indicates the numerical benchmarks measured as low, marginal, and high. With the measurements in hand, the risk impact for each benchmark has also been stipulated as it is implemented during the simulation. This is presented in Table 12 below:

Table 12. Financial Ratios used in the Fuzzy Logic

Financial Indicator	Benchmark			Risk Impact
	Low	Marginal	High	
Net Profit Margin	≥ 25	10-25	≤ 10	The higher the margin, the lower the risk
Return on Investment	≥ 5	1-5	≤5	The higher the ratio, the lower the risk
Liquidity Ratio	≥ 1.50	1.00 - 1.50	≤ 5.0	The higher the ratio, the less risk
Debt to Asset Ratio	≤ 30	30 -55	≥ 55	The higher the ratio, the more risk

The table shows the different financial indicators and different benchmarks. For net profit margin (NPM), a higher margin indicates lower risk. For return on investment (ROI), the higher the ratio, the lower the risk. For the liquidity ratio (LR), the higher the ratio, the lower the risk, and lastly, for the debt-to-asset ratio (DAR), the higher the ratio, the more risk there will be.

3.4.2 MATLAB Model Building

MATLAB was used to create the rule-based fuzzy logic component of the system. The fuzzy system is composed of five features represented by a Fuzzy Interference System (FIS) as shown in Figure 4:

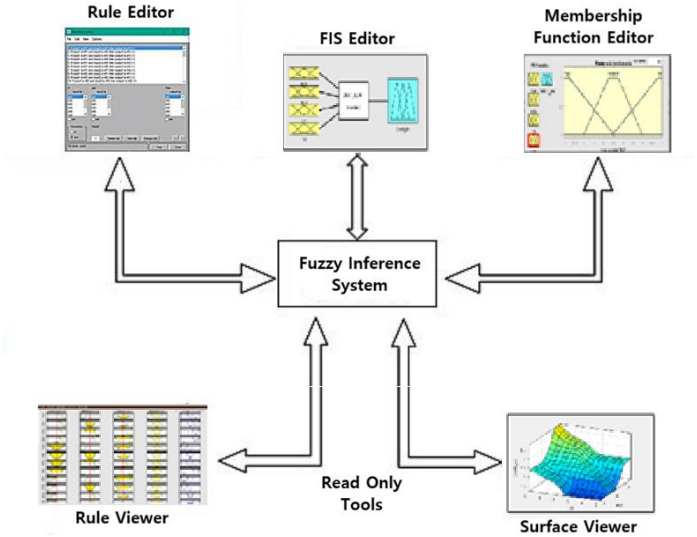


Figure 4. MATLAB Simulation of the DOST SET-UP Risk Analysis and Recommendation System

The proposed Risk Analysis and Recommendation System for the Department of Science and Technology – Small Enterprise Technology Upgrading Program (DOST SET-UP) was developed using the MATLAB Fuzzy Logic Toolbox. The fuzzy inference system (FIS) was designed to assess the credit risk of project proposals based on specific financial indicators. The system uses a rule-based fuzzy-logic technique to convert numerical financial ratios into linguistic risk terms and to produce a final credit risk score.

3.4.3 Fuzzy Inference System Structure

The fuzzy inference system has four input variables and one output variable. The input variables are the Debt-to-Asset Ratio (DAR), Net Profit Margin (NPM), Liquidity Ratio (LR), and Return on Investment

(ROI). These financial ratios were chosen as important factors in determining the financial viability and credit risk of SET-UP project proposals. The output variable is the Credit Risk Score.

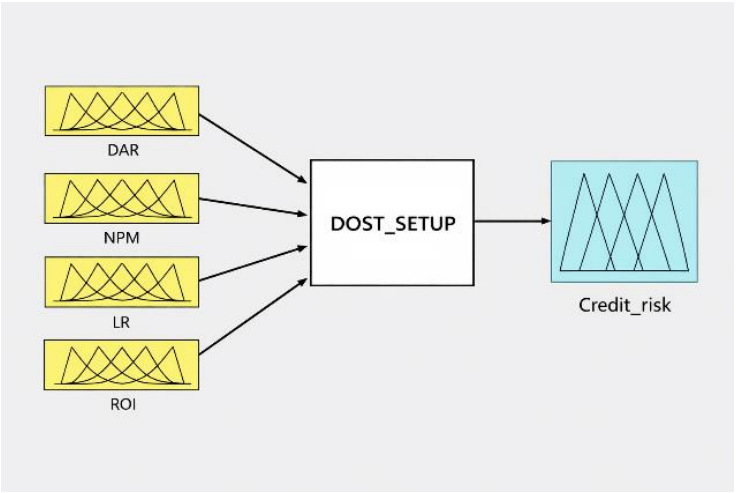


Figure 5. MATLAB Simulation of the DOST SET-UP Risk Analysis and Recommendation System

In the DOST SET-UP risk analysis system, the Debt-to-Asset Ratio (DAR) is one of the major parameters used to evaluate the firm's financial condition. The debt-to-asset ratio is evaluated based on a specific numerical range to determine the level of credit risk associated with the firm's financial structure, as shown in Table 13. The risk level is classified based on the percentage of total debt relative to total assets. In this context, credit risk is considered low if the debt-to-asset ratio is below 30%, marginal if it falls between 30% and 55%, and high if it is above 55%. This shows that the higher a firm's debt-to-asset ratio, the higher the level of credit risk. This assumes that firms with higher debt relative to their assets may find it difficult to meet their obligations. Moreover, the DOST SET-UP risk analysis system implements the list of risk assessment parameters for the Debt-to-Asset Ratio (DAR). The DAR is considered one of the main financial indicators for assessing an enterprise's credit risk by identifying the percentage of total assets financed by debt. In the study, the DAR input feature is represented by the Gaussian membership function (gaussmf) under the fuzzy logic approach provided by the MATLAB tool. The range of the risk score is between 0 and 55, while the four membership parameters, namely minimum, low, moderate, and high, are established by the system for identifying the degree of membership for every value under the fuzzy risk categories defined by the system, namely Low Risk, Marginal Risk, and High Risk. Through this fuzzy logic representation, the system can effectively model uncertainty and gradual transitions between risk levels, allowing a more flexible and realistic interpretation of the financial condition of enterprises participating in the DOST SET-UP program.

Table 13. Risk Assessment Parameters of Debt-to-Asset Ratio in the DOST SET-UP System

	Debt-to-Asset Ratio (DAR)			
	Minimum	**Low**	**Moderate**	**High**
Low Risk	-24.80	-2.75	2.75	30.19
Marginal Risk	4.73	24.8	30.19	49.8
High Risk	30.25	52.25	57.75	79.75

As shown in Table 13 above, the DOST SET-UP risk analysis system uses numerical values within specific ranges as parameters in the risk assessment process to determine the DAR. The parameters and risk-level classification are based on the standard risk-level guideline for financial risk management. As such, if the DAR level is less than 30%, the risk level is classified as Low Risk. However, if the DAR level is between 30% and 55%, the risk level is classified as Marginal Risk. Finally, if the DAR level is above 55%, the risk level is classified as High Risk. As such, the higher the debt level relative to the enterprise's total assets, the higher the enterprise's financial leverage. As such, the risk level also increases because credit risk increases. For the fuzzy membership representation, the parameters for the Low Risk level include −24.80, −2.75, 2.75, and 30.19.

The parameters for the Marginal Risk level would include the values 4.73, 24.80, 30.19, and 49.80. Finally, the parameters for the High Risk level would include the values 30.25, 52.25, 57.75, and 79.75. As such, the numerical values would form the basis for the fuzzy membership functions in the risk evaluation process.

Table 14. Risk Assessment Parameters of Net Profit Margin in the DOST SET-UP System

	Net Profit Margin (NPM)			
	Minimum	**Low**	**Moderate**	**High**
Low Risk	15	20.34	25.9	35.9
Marginal Risk	5.13	10	15	19.9
High Risk	-11.30	-1.25	4.993	10

Table 14 shows the risk assessment parameters for the Net Profit Margin (NPM), which is part of the DOST SET-UP risk analysis system. Generally, net profit margin is an essential profitability measure that defines the amount of net income generated from the enterprise's total revenue. It indicates the enterprise's ability to manage operational costs and generate net profitability. In the context of the DOST SET-UP risk analysis system, net profit margin is an assessment parameter defined by numerical ranges used to identify the credit risk level. For instance, the risk level is considered Low Risk when the net profit margin is more than 25%, Marginal Risk when the net profit margin is between 10% and 25%, and High Risk when the net profit margin is less than 10%. This analysis suggests that the higher the net profitability, the lower the credit risk, and the lower the net profitability, the higher the risk of instability in the enterprise's financial position. In the fuzzy logic analysis system using MATLAB, the input variable net profit margin (NPM), which is part of the risk analysis in the DOST SET-UP system, is defined using the Gaussian membership function (gaussmf). This is because the risk score ranges from 0 to 25, which is essential for evaluating profitability levels and defining the membership value for the respective risk category. In the fuzzy membership representation, the Low Risk category is defined by the parameter values 15, 20.34, 25.90, and 35.90. The Marginal Risk category is defined in terms of the parameter values 5.13, 10, 15, and 19.90, while the High Risk category is defined in terms of the parameter values -11.30, -1.25, 4.993, and 10. These numerical parameters are used to construct membership functions that classify the enterprise's profitability level into the appropriate credit risk category. With this fuzzy representation, the system can capture gradual transitions in profitability levels and offer a more flexible and realistic interpretation of the financial risk of enterprises participating in the DOST SET-UP program.

Table 15. Risk Assessment Parameters of Liquidity Ratio in the DOST SET-UP System

	Liquidity Ratio (LR)			
	Minimum	**Low**	**Moderate**	**High**
Low Risk	-2.18	-0.18	0.9854	2.48
Marginal Risk	0.0331	2.23	12.758	5.02
High Risk	-2.53	4.00	5.26	7.26

Table 15 shows the parameters used in the risk assessment for the Liquidity Ratio (LR) in the DOST SET-UP risk analysis system. Liquidity Ratio is a financial ratio used to evaluate an enterprise's capacity to settle its short-term financial obligations using its current assets. Liquidity Ratio is a key factor in determining an enterprise's financial stability and sustainability. In the DOST SET-UP risk analysis system, the Liquidity Ratio is used as a parameter in the risk assessment process. The risk assessment process uses a specific range to determine the enterprise's credit risk. The risk assessment process is as follows: if the liquidity ratio is less than 1, it is categorized as a low-risk enterprise. If the liquidity ratio ranges from 1 to 1.5, the risk is categorized as a Marginal Risk enterprise. If the liquidity ratio ranges from 1.5 to 5, the enterprise is categorized as a High Risk enterprise. From the description above, it is clear that the Liquidity Ratio is used to determine the enterprise's credit risk. Enterprises with a high liquidity ratio are likely to settle their financial obligations in the short term, indicating lower credit risk. In the fuzzy logic system using MATLAB, the input variable Liquidity Ratio (LR) uses the Gaussian membership function, represented as "gaussmf."The score varies from

0 to 5, enabling the fuzzy inference system to assess different levels of liquidity and classify liquidity into the relevant risk categories.

According to the generated fuzzy membership parameters, the Low Risk category has the numerical parameters of -2.18, -0.184, 0.9854, and 2.48. The Marginal Risk category has numerical parameters of 0.0331, 2.23, 2.758, and 5.02. In contrast, the High Risk category has the numerical parameters of 2.53, 4.001, 5.26, and 7.26, which serve as the basis for developing the membership functions to assess the extent of a particular liquidity ratio's affiliation with the different risk categories.

Table 16. Risk Assessment Parameters of Return on Investment in the DOST SET-UP System

	Return on Investment (ROI)			
	Minimum	**Low**	**Moderate**	**High**
Low Risk	2.48	4.76	5.26	7.26
Marginal Risk	0.985	2.25	2.75	3.975
High Risk	-2.25	-0.25	0.25	2.52

Table 16 shows a summarized version of the Return on Investment (ROI) risk classification implemented in the system. ROI is considered an inverse measure of credit risk. In more detail, credit risk is categorized as low if ROI falls within the range of 5% to 10%, marginal if ROI falls within the range of 1% to 5%, and high if ROI is less than 1%. In the proposed fuzzy inference system, ROI will utilize a Gaussian membership function represented by the function "gaussmf." This function enables a smooth transition between risk categories and helps handle uncertainty in financial performance evaluation. In more detail, the risk score will fall within the range of [0, 5], and lower risk scores indicate lower credit risk. According to the membership function, ROIs of 2.48, 4.76, 5.26, and 7.26 fall within the low-risk range. ROI values of 0.985, 2.25, 2.75, and 3.975 fall within the range of marginal risk, while ROI values of -2.25, -0.25, 0.25, and 2.52 fall within the range of high risk. It can be seen that ROI values overlap across different risk categories, indicating that risk assessment does not fall abruptly but is proportionally weighted according to membership degree. Notably, the effectiveness of the suggested approach is confirmed by comparing model performance before and after the feature selection process. The model had higher dimensionality and redundancy when all twelve variables, related to financial and operational aspects, were considered, leading to less stable and interpretable results.

3.4.3 Overall Risk Scoring

For the fuzzy logic output, CREDIT RISK uses the classical set for credit scoring (Shang, 2013). This is illustrated using the following diagram:

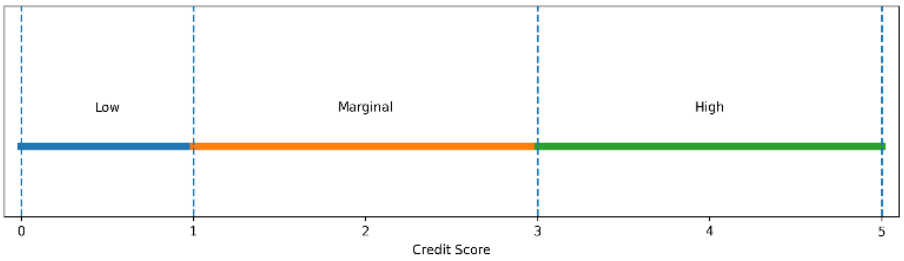


Figure 6. Classical Set for Credit Score

Figure 6 shows the classical set representation of the credit score used in the study. Each set has a membership function that defines the extent to which an element belongs to a given risk-level set. In this case, the credit risk level is considered low if the value is less than 1, marginal if the value lies between 1 and 3, and high if the value lies between 3 and 5. This shows a linear relationship between the credit risk score and the level of risk: as the credit risk score increases, the level of risk increases, and as the credit risk score decreases, the level of risk decreases. The data were simulated using the MATLAB program to calculate the actual fuzzy membership value. The system output, which is the credit risk score, has a triangular membership function, denoted as 'trimf.' The risk score ranges from [0,5]. A summary of the classification of the credit risk score is shown in the table below:

Table 17. Simulation Result for Risk Assessment in the DOST SET-UP System

Risk Category	Credit Risk Score Range	Description
Low Risk	1.00 – 2.38	Indicates minimal credit exposure; borrower demonstrates strong repayment capacity.
Marginal Risk	2.39 – 3.38	Represents moderate credit exposure; repayment capacity requires closer monitoring.
High Risk	3.39 – 5.00	Indicates significant credit exposure; higher probability of default.

Table 17 below shows the simulation results obtained from the MATLAB program used to implement the credit risk score classification. From the result obtained, low-risk classification refers to a credit risk score of 1.00 to 2.38, marginal risk refers to a credit risk score of 2.39 to 3.38, while high-risk classification refers to a credit risk score of 3.39 to 5.00. The numerical value was obtained after determining the parameters of the fuzzy inference system. After determining the membership values, the result was evaluated, and the output was defuzzified to obtain a crisp credit risk score. This was done to validate the system output and ensure its accuracy.

3.4.4 Credit Risk Output Representation

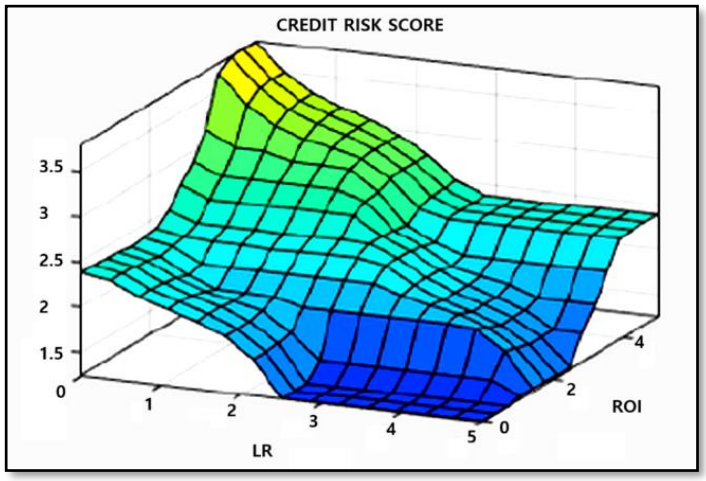


Figure 7. Surface View of the Credit Risk for Liquidity Ratio and Return on Investment

Figure 6 shows the three-dimensional surface plots of the proposed fuzzy-based credit scoring and recommendation system. The model has multiple input variables, such as Liquidity Ratio (LR), Return on Investment (ROI), and Net Profit Margin (NPM), which map to a single output variable, i.e., Credit Risk Score. Since the system has four input variables and one output variable, the relationship can be shown using two different plots. The plots show the effect of changes in various financial indicators on the overall credit risk score, which is calculated using the proposed model. The relationship between input variables such as LR and ROI and the output variable (i.e., Credit Risk), as well as the relationship between NPM and LR and the output variable (i.e., Credit Score), shows the nonlinear relationship using the fuzzy inference system. The overall credit risk, which is the center of gravity of the output, is 2.29, as shown by the computation of the proposed system. The proposed model incorporates a rule-based fuzzy inference system, in which all decision rules are embedded in the model and can be adjusted by the system administrator. The rules can be adjusted based on the numerical values obtained using the MATLAB simulation tool. Each rule has an equal effect on the output variable depending upon the extent of its activation. Ultimately, it acts as an intelligent credit scoring and recommendation system. It evaluates financial performance indicators, calculates a risk score through fuzzification, rule evaluation, and defuzzification, and produces a corresponding recommendation. The final recommendation will consolidate all input variables into a single assessment. After validation, the administrator may enable the notification system, which will send updates on the status of their application

to SET-UP applicants. This model offers a structured, data-based, and dynamic framework that helps maintain consistency and accuracy in project screening.

3.5 Model Evaluation Result using Confusion Matrix

To ensure the output of the rule-based fuzzy logic model is accurate, a confusion matrix was developed based on actual project implementation data versus the results obtained when running the newly developed Risk Analysis and Recommendation System of DOST SET-UP. The table shows the confusion matrix describing the performance of the rule-based fuzzy logic system for credit scoring using binary classifiers in MATLAB. Because the system used historical data from the actual DOST-SET UP implementation, it was limited to only 33 projects. The data included two possible predicted classes: "Accepted" and "Rejected." "Accepted" means the project was funded by DOST, while "Rejected" means that it was not funded. The population of 33 was tested, and the classifier predicted "Accepted" 30 times and "Rejected" 3 times. In reality, those who were funded were 30, and 3 were not. Thus, the following classifiers are identified as true positive (TP), true negative (TN), false positive (FP), and false negative (FN).

Table 18. Confusion Matrix for DOST SET-UP Credit Scoring

n=33	Predicted: Rejected	Predicted: Accepted	
Actual Rejected	True Negative (TR)= 2	False Positive (FP)= 1	**3**
Actual Accepted	False Negative (FN)= 1	True Positive (TP)= 29	**30**
	3	**30**	

Thus, the following classifiers have been identified as true positive (TP), true negative (TN), false positive (FP), and false negative (FN). This was computed in MATLAB, and the following result has been generated:

Computation of System Accuracy:

$$\boldsymbol{Accuracy} = \frac{(\mathrm{TP} + \mathrm{TN})}{\mathrm{n}}$$
$$\boldsymbol{Accuracy} = .93 \quad (7)$$

The result shows that the DOST SET-UP is .93 percent accurate in terms of output generation. To compute it, the false positives and true negatives are divided by the total number of the proposal population. So, one added to twenty-nine and two divided by thirty-three is equal to .93, or 93%. The accuracy rate is the proportion of correct predictions. Ultimately, an accurate risk analysis system can automatically evaluate project proposals, reducing paperwork and supporting sound decision-making [12].

Computation of System Error Rate:

$$\boldsymbol{Error\ Rate} = \frac{(\mathrm{FP} + \mathrm{FN})}{\mathrm{n}}$$
$$\boldsymbol{Error\ Rate} = .6 \quad (8)$$

In addition, the result shows that the system's miscalculation rate is 0.06. The generated error rate is deducted from 1, resulting in 0.04, which is considered the recall. To compute this, false positives and false negatives are divided by the total number of proposals. So, 1 is added to 1 and divided by 33, equaling 0.6. This result means that a lower error rate indicates higher reliability, which is advantageous when relying on the results [13].

Computation of True Positive Rate:

$$\boldsymbol{True\ Positive\ Rate} = \frac{(\mathrm{TP} + \mathrm{TN})}{\mathrm{n}}$$
$$\boldsymbol{True\ Positive\ Rate} = .92 \quad (9)$$

This data shows that the system has a sensitivity of .92, meaning it correctly predicts 92% of actual accepted cases. To compute it, divide the true positives by the total number of actual accepted proposals. So, twenty-three

divided by twenty-five is equal to .92 or 92%. This result means the actual and predicted data are highly similar, which supports their accuracy [14].

Computation of False Positive Rate:

$$\boldsymbol{False\ Postive\ Rate} = \frac{\text{FP}}{\text{Actual Rejected}}$$
$$\boldsymbol{False\ Positive\ Rate = .33} \tag{10}$$

This result shows that when it's actually rejected, it predicts acceptance 33% of the time. To compute this, divide the false positives by the total number of actually rejected proposals. So, one divided by three equals 33. This means the system has high accuracy in identifying rejection [12, 15].

Computation of True Negative Rate:

$$\boldsymbol{True\ Negative\ Rate} = \frac{\text{TN}}{\text{Actual Rejected}}$$
$$\boldsymbol{False\ Positive\ Rate = .66} \tag{11}$$

The result shows that the system predicts rejection 66% of the time when it's actually rejected. To compute this, the true negative is divided by the total actual rejected proposals. So, 2 divided by 3 equals .66. The generated rate is divided by 1, resulting in .34, which is considered the *specificity* [12, 14].

Computation of Precision:

$$\boldsymbol{Precision} = \frac{\text{TP}}{\text{Predicted Accepted}}$$
$$\boldsymbol{Precision} = .97 \tag{12}$$

This clearly shows that the output of DOST SET-UP using rule-based fuzzy logic is 97% precise. To compute this, divide the true positives by the predicted accepted proposals. So, twenty-nine divided by thirty is equal to .97 or 97%. This is the proportion of positive cases that were correctly identified. This shows that the system's output is reliable [13].

Computation of Prevalence:

$$\boldsymbol{Prevalence} = \frac{\text{Actual Accepted}}{\text{n}}$$
$$\boldsymbol{Prevalence} = .94 \tag{13}$$

Performance metrics such as recall, precision, and specificity are also widely acknowledged in the credit-scoring literature as offering a more complete evaluation than accuracy in datasets with class imbalance [12–14]. Recall, for instance, helps evaluate a model's ability to identify positive cases correctly. Similarly, precision helps ensure a model's reliability by reducing false positives. These metrics are best practices for financial risk evaluation systems [15, 16].

4. Conclusions

This study demonstrates the viability of integrating machine learning-based feature selection with a rule-based fuzzy logic approach to improve credit risk analysis under the DOST SET-UP program. From an initial set of twelve financial and operational features, the proposed Sequential Forward Floating Selection (SFFS) technique was able to effectively select the four most important financial features, namely the Debt-to-Asset Ratio (DAR), Net Profit Margin (NPM), Liquidity Ratio (LR), and Return on Investment (ROI), as the most significant factors influencing the creditworthiness of the entrepreneurs. By incorporating the proposed features into the fuzzy inference-based credit risk analysis and recommendation system, the vagueness and uncertainty in entrepreneurs' decision-making process were effectively addressed. Although the study made remarkable findings, it is important to note some limitations. To start with, the study has used a given data set in its analysis. However, the results may not be generalizable. Moreover, the use of the SFFS feature selection method and a specific classifier does not guarantee that the results would be the same with a different

classifier. The results may vary if a different classifier is used. Results may also vary if a different feature selection method is used. However, the proposed system achieved an overall classification accuracy of 94%, indicating its reliability in distinguishing between approved and rejected project proposals. The results validate the potential of integrating the SFFS technique with fuzzy inference systems for credit scoring, especially in data-scarce environments such as the DOST SET-UP program. Thus, the proposed framework presents a structured approach for evaluating credits in a data-driven and transparent manner. The approach is beneficial not only for credit evaluation but also for ensuring more objective funding decisions within DOST-SETUP. The proposed methodology is useful not only for its intended purpose, but also for establishing a platform for other government institutions and financial institutions to manage credits effectively.

Authors' Contribution: D.I., C-S.: Conceptualization, Methodology, Software, Investigation, Data curation, Formal analysis, Validation, Visualization, Writing – original draft. C.I, S.: Investigation, Resources, Writing – review and editing. D.A., R.: Investigation, Resources, Writing – review and editing. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

[1] Department of Science and Technology. *Small Enterprise Technology Upgrading Program (SETUP)—Program Description and Guidelines*; 2023. https://region1.dost.gov.ph/setup/.

[2] Beaver, W. Financial Ratios as Predictors of Failure. *J. Account. Res.* **1966**, *4*, 71–111. https://doi.org/10.2307/2490171

[3] Doumpos, M.; Zopounidis, C.; Dounias, G.; Karagiannis, N. Feature Selection for the Credit Risk Prediction Problem; 2019. https://www.researchgate.net/publication/387707974_Features_Selection_for_Credit_Risk_Prediction_Problem

[4] Fallahpour, S.; Lakvan, E. N.; Zadeh, M. H. Using an Ensemble Classifier Based on Sequential Floating Forward Selection for the Financial Distress Prediction Problem. *J. Retail. Consum. Serv.* **2017**, *34*, 159–167. https://doi.org/10.1016/j.jretconser.2016.10.010

[5] Yang, J.; Liu, Z.; Wang, Y. Feature Selection in Credit Rating Assessments Using Principal Component Analysis and Factor Analysis. *arXiv* **2020**, arXiv:2011.09137. https://arxiv.org/abs/2011.09137

[6] Ding, Y.; Yan, L. A Financial Distress Prediction Model Integrating Multi-Source Data with MRMR and SVM-RFE for Selection. *arXiv* **2024**, arXiv:2404.12610. https://arxiv.org/abs/2404.12610

[7] Wu, X.; Li, H.; Zhang, P. An Ensemble Feature Selection Method and Improved Stacking Algorithm for Financial Distress Prediction. *Kybernetes* **2024**. https://doi.org/10.1108/K-08-2023-1428

[8] Zhang, L.; Chen, M.; Huang, F. A Two-Stage Feature Selection Method for Artificial Neural Network-Based Financial Distress Prediction. *Eng. Econ.* **2024**, *35*, 45–61. https://inzeko.ktu.lt/index.php/EE/article/view/30924

[9] Santoso, B.; Wibowo, A. Comparison of Linear Discriminant Analysis and Support Vector Machine for Financial Distress Prediction. *IOP Conf. Ser.: J. Phys.* **2018**, *979*, 012089. https://doi.org/10.1088/1742-6596/979/1/012089

[10] Palongpalong, P. S.; Taroc, R. D.; Bilocora, D. D. Blended Learning with Online Assessment in Higher Education: Its Global Environmental Dimensions of Sustainability; 2018.

[11] Tan, M.; Pu, J.; Zheng, B. Optimization of Breast Mass Classification Using Sequential Forward Floating Selection (SFFS) and a Support Vector Machine (SVM) Model. *Int. J. Comput. Assist. Radiol. Surg.* **2014**, *9*, 1005–1020. https://doi.org/10.1007/s11548-014-0992-1

[12] Brown, I.; Mues, C. An Experimental Comparison of Classification Algorithms for Imbalanced Credit Scoring Data Sets. *Expert Syst. Appl.* **2012**, *39*, 3446–3453. https://doi.org/10.1016/j.eswa.2011.09.014

[13] Hand, D. J.; Henley, W. E. Statistical Classification Methods in Consumer Credit Scoring: A Review. *J. R. Stat. Soc.* **1997**, *160*, 523–541. https://doi.org/10.1111/1467-9876.00088

[14] Lessmann, S.; Baesens, B.; Seow, H.-V.; Thomas, L. C. Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring: An Update of Research. *Eur. J. Oper. Res.* **2015**, *247*, 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

[15] Oreski, S.; Oreski, G. A Novel Credit Scoring Model Based on Genetic Fuzzy Rule-Based Systems. *Expert Syst. Appl.* **2015**, *42*, 5623–5634. https://doi.org/10.1016/j.eswa.2015.04.003

[16] Crook, J.; Edelman, D.; Thomas, L. Recent Developments in Consumer Credit Risk Assessment. *Eur. J. Oper. Res.* **2007**, *183*, 1447–1465. https://doi.org/10.1016/j.ejor.2006.07.004