sciendo

BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **24**, No 4
Sofia • 2024 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2024-0034

Latest Advancements in Credit Risk Assessment with Machine Learning and Deep Learning Techniques

Umangbhai Soni[1], *Dr. Gordhan Jethava*[2], *Dr. Amit Ganatra*[3]

[1]*Department of Computer Science and Engineering, Parul Institute of Engineering & Technology, FET, Parul University, Vadodara, 391760, Gujarat, India*
[2]*Department of Information Technology, Parul Institute of Engineering & Technology, FET, Parul University, Vadodara, 391760, Gujarat, India*
[3]*Parul University, Vadodara, 391760, Gujarat, India*
E-mails: umanghsoni@gmail.com g.jethava@gmail.com provost@paruluniversity.ac.in

***Abstract*:** *A loan is vital for individuals and organizations to meet their goals. However, financial institutions face challenges like managing losses and missed opportunities in loan decisions. A key issue is the imbalanced datasets in credit risk assessment, hindering accurate predictions of defaulters. Previous research has utilized machine learning techniques, including single or multiple classifier systems, ensemble methods, and class-balancing approaches. This review summarizes various factors and machine learning methods for assessing credit risk, presented in a tabular format to provide valuable insights for researchers. It covers data complexity, minority class distribution, sampling techniques, feature selection, and meta-learning parameters. The goal is to help develop novel algorithms that outperform existing methods. Even a slight improvement in defaulter prediction rates could significantly influence society by saving millions for lenders.*

***Keywords*:** *Credit risk assessment, Single classifier system, Multiple classifier system, Dynamic selection, Class-Imbalance.*

1. Introduction

The term “credit” in the banking and finance industry refers to an agreement wherein a financial organization provides a finite amount of money to individuals or organizations, with the commitment that the borrowed funds will be repaid later either in a lump sum or through multiple installments. Lending entails a certain level of risk, as it involves making complex decisions that can potentially lead to the financial institution’s bankruptcy. Therefore, it is crucial to assess credit operations carefully. In light of this importance, banks establish credit limits, and economic organizations must align their credit operations with their risk capacity. As a result, prospective defaulters can be identified based on factors such as the level of risk, provided collateral, and the nature of the financial transactions conducted. The

decision to approve loans is made by the relevant authorities using various techniques and security measures.

Between 2014 and 2021, the gross non-performing assets of India's public sector banks have surged. When factoring in private sector banks and NonBank Financial Companies (NBFCs), the combined total poses a significant threat to the financial stability of many organizations and the country's economic framework [1]. Despite the pandemic's impact on earnings, the demand for education loans in India increased significantly, with banks and NBFCs disbursing a substantial amount over six months in 2020 [2]. A major concern for banks and NBFCs is that a notable portion of disbursed loans has turned into NPAs, leading to significant financial losses [26].

Machine learning techniques are used to classify genuine and non-genuine customers in the finance industry, where high credit demand and competition drive the need for reliable credit models. Traditional models assess borrowers' likelihood of default based on loan installments and due dates [27]. Credit risk models commonly rely on computerized techniques such as Support Vector Machines (SVM) [3, 7], Decision Trees (DT) [5], Logistic Regression (LR) [4], and BackPropagation Neural Networks (BPNN) [9]. In addition to these methods, the use of Multiple Classifier Systems (MCS) has been explored to assess credit risk operations [14]. These advancements have been made possible through the contributions of researchers in the field of Machine Learning (ML) and Artificial Intelligence (AI), leading to continual improvements in credit assessment methods. Financial institutions employ various tools such as Machine Learning (ML), Data Mining (DM), and Database Analysis to enhance decision-making efficiency in credit operations [27]. The ultimate goal is to establish an effective and efficient method for approving credit to individuals and organizations. These methods encompass ML learners, including Random Forest (RF) [8, 31], AdaBoost [17], XGBoost [15, 18], and stacking ensemble [20], among others.

However, before utilizing the aforementioned methods, several processes must be performed to enhance the accuracy of the models. These processes encompass pre-processing raw data, balancing class distributions, feature selection, and more.

1.1. Our contributions

The significant contributions of the paper are summarized as follows:

- Presenting the latest research findings in credit risk assessment using machine learning techniques.
- Discussion on prospects and trends in the field of credit risk assessment.
- Emphasizing the significance of the dynamic classifier system in improving classification accuracy, particularly for under-represented minority instances like probable defaulters.
- Demonstrating the critical role of class-balancing techniques in enhancing model performance for credit risk assessment.
- Evaluating various criteria and methods for classifier selection to identify the most suitable models.

- Exploring different approaches for defining the Region Of Competence (ROC) to handle complex credit risk scenarios better.
- Investigating novel algorithms and techniques to advance credit risk assessment.
- Highlighting the application of the most recent Meta-DES-based framework as a promising avenue for improving credit risk assessment models.

The remainder of the article is organized as follows: Section 2 provides a brief overview on what is types of risk, classifies the credit-risk models, discusses the workflow of credit-risk assessment, and discusses traditional or statistical methods and semi-parametric techniques used for risk assessment based on credit-related financial operations. Both Sections 3 and 4 review the Single Classifier System (SCS) and Multiple Classifier System (MCS) respectively and summarize a number of the latest research articles with findings and the future directions for the novel research. Section 5 emphasizes the importance of dynamic selection in multiple classifier systems which focuses on the various approaches for defining a ROC for classifiers, numerous classifier selection schemes, and the latest methods used in the mentioned approaches. Section 6 highlights the challenges faced during the classification due to imbalanced data are elaborates on probable solutions using data sampling methods, feature selection methods, and hybrid/ensemble methods. Section 7 provides research findings. Section 8 discusses the future score of work in the domain. The final section concludes the work.

2. Theoretical background

2.1. Types of risks

Fig. 1 depicts various types of risks. There are four main types of risks: systemic risk, operational risk, financial risk, and legal risk [27]. Credit risk and liquidity risk are subcategories of financial risk. Operational risk is contributed to by risks arising from instrument failure, human error, and operational system failure. Furthermore, operational risk also increases liquidity risk and credit risk. Legal risk encompasses any acts that are monitored and controlled by a regulatory body and are unsuitable for liquidation. Additionally, the risks mentioned earlier, whether individually or in combination, can be the cause of systemic risk. Systemic risk can be defined as the probability of gaps in credit and/or liquidity by borrowers in the financial system.

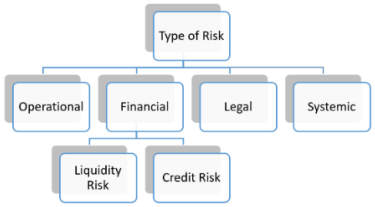

Fig. 1. Types of risks [27]

2.2. Classification of credit-risk models

The assessment and measurement of the credit risk-based model can be further categorized into three classes: (i) Portfolio risk, (ii) Stochastic risk, and

(iii) Classification risk, as shown in Fig. 2. In the case of risk-based classification models, the borrower's risk is computed based on the specific operation performed. Stochastic models, on the other hand, incorporate stochastic behavior using parameters used in the calculation of credit risk. Additionally, portfolio risk models approximate the credit portfolio value or provide risk estimation, while also calculating the distribution of the probability of loss.

Credit scoring models, also known as risk classification models, can be further divided into two categories: (1) Credit approval models, and (2) Behavioral scoring models. The credit approval model takes into account the organization's registration data, whereas the behavioral scoring model analyses data from previous financial operations conducted by financial institutions. Machine learning models play a crucial role in classifying risk-based models to determine the approval or denial of credit operations.

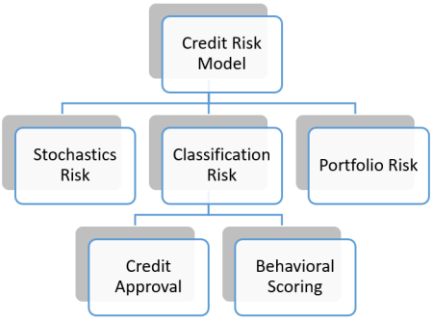

Fig. 2. Classification of models based on credit risk [27]

2.3. Statistical Methods and semi-parametric techniques for credit-risk assessment

Statistical methods were predominantly used to predict the default risk of different organizations before the widespread adoption of various intelligent approaches, such as SCS and MCS. To effectively build and analyze corporate failure predictions, it is necessary to have an optimized combination of various attributes. The steps involved in statistical-based methods include: (i) selecting attributes and sample size, (ii) choosing a method and determining the financial ratios based on specific attributes, and (iii) validating the method used for performance measurement [28]. One of the most critical tasks is selecting the appropriate method based on the available data and the analysis objective. However, statistical techniques have limitations such as assuming linear separability, multivariate normality, and independence between pre-existing forms of function and predictive attributes. These assumptions overlook boundaries, inter-relationships among financial attributes, and the complex nature [3].

In general, the goal is to strike a balance between underfitting and overfitting when developing a model. The semi-parametric approach combines the advantages of both parametric and non-parametric methods. Parametric models tend to offer good reliability but may have concerns regarding accuracy. On the contrary, non-parametric models prioritize accuracy but may lack stability. K i m and Y o o [29] proposed a semi-parametric technique for predicting the bankruptcy of financial

institutions by integrating parametric models such as Multi-variant Discrete Analysis and Logistic Regression (MDALR) with non-parametric methods like Neural Networks (NN). This approach provides a favorable trade-off between bias and variance.

2.4. Workflow of credit-risk assessment model

The major steps in the credit risk assessment procedure are illustrated in Fig. 3. The financial details or data points of the borrower are obtained from a financial institution. Before feeding the data directly into the model, it is crucial to ensure that the input data is complete and in the proper format, which is accomplished during the pre-processing step. Pre-processing is an essential task that involves filtering and filling in missing values, transforming categorical data, normalizing the data, and constructing a sample set for model training, validation, and testing [10]. Once the pre-processing of the raw data is complete, a feature selection technique is applied to reduce the number of attributes using either linear or non-linear transformation methods. Linear transformation employs PCA, while ISOMAP and LLE are used for non-linear cases. Subsequently, the data is passed to a learner, which can be a statistical method, a single classifier system, or a multiple classifier system. At this stage, additional domain-specific information is provided to the learner, including observation classes, class distribution, variable types, and so on. Finally, the assessment is conducted using various performance metrics [3].

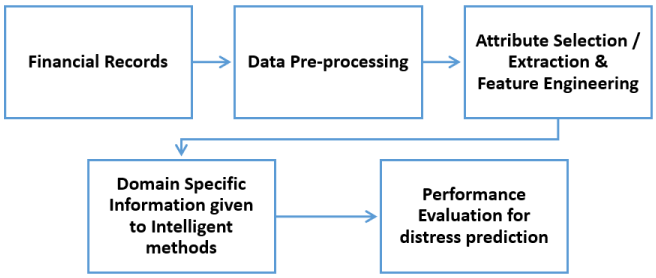

Fig. 3. Pipeline for the assessment of credit risk [3]

3. Single Classifier System (SCS)

Due to the rise of modern computer techniques, the world is rapidly moving towards digitalization. This transformation holds numerous benefits, particularly in the finance and banking sectors, especially concerning credit risk management. When assessing a new loan application, accurately evaluating the risk associated with the proposed client becomes an extremely crucial task. For instance, tasks like document submission, application verification, and making the final loan decision can be easily and swiftly accomplished. This importance grows even further when dealing with financial risk assessment for medium and small enterprises, as it requires caution due to the substantial loan amounts involved, which can significantly impact the functioning of lending institutions. In these cases, smart applications may carry the potential risk of default, leading to significant capital losses for financial institutions.

Thus, evaluating the borrower's credit risk becomes critically important in such applications. To mitigate financial losses in these situations, various algorithms, techniques, and models have been explored.

Before delving into the literature reviews on SCS, it is prudent to understand the design procedure of SCS, which comprises three steps: feature selection, algorithm design, and performance evaluation. Fig. 4 illustrates the block diagram representing this design procedure.

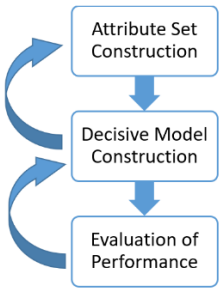

Fig 4. Design Procedure of SCS [23]

Table 1. Finding and future scope of the reviewed articles based on a SCS

Reference	Findings	Future scopes
Liu et al. [31] (2024)	Loan default prediction is enhanced by the use of network centrality factors. Three machine learning models were used: MLP, RF, and Elastic Net	Generalise to other financial data and apply further graph-theory measurements
Quan and Sun [32] (2024)	Factorization Machine (FM) model performs better than SVM and k-NN. Captures feature interactions effectively	Loss functions optimization in the FM model Adoption of online credit risk assessment
Chen, Jin and Lu [33] (2024)	Implemented a Neural Network modified by a Genetic Algorithm for credit risk assessment in Micro, Small, and Medium-sized Enterprises (MSMEs)	Refining GA-BPNN for complex data and various industry sectors
Moscato, Picariello and Sperlí [6] (2021)	Comparative analysis presented on renowned P2P lender, Lending Club data in combination with various models and sampling techniques to assess the risk of the borrower	Ensemble and deep learning-based approaches can be evaluated on other P2P lending data
Gao, Yang and Zhao [19] (2024)	Comparison of Variant of Long Short-Term Memory (LSTM) was done with Back Propagation Neural Network (BPNN), Recurrent Neural Network (RNN), LSTM for rural microcredit risk assessment	Refine data collection methods in rural areas can enhance sustainable, effective, and robust risk assessment
Bulut and Arslan [8] (2024)	FeatUre reduction and data splitting improved the performance of credit risk models., i.e., DT, LR, RF, and NB	Approach can apply to deep learning models Generalize the finding by using other credit data
Du, Liu and Lu [9] (2021)	Back-Propagation Neural Network (BPNN) was implemented for early warning of credit risk of MSME and compared with genetic algorithm-based BPNN. Target variable categorized into four categories	Adopt extensive data samples to maximize the reliability, and accuracy of other data

It functions as a feedback system, wherein modifications are required at any earlier stage if the performance metrics fail to meet the required specifications. Optimum feature selection not only simplifies the algorithmic tasks but also enhances performance. Similarly, even with poorly selected features, a well-designed algorithm can still yield accurate results [23]. Thus, to achieve a robust performance of the SCS, a perfect combination of feature selection and algorithm is necessary. Additionally, the work related to the single classifier system explored by researchers is summarized in Table 1.

4. Multiple Classifier System (MCS)

MCS, which stands for Multiple Classifier Systems, refers to the process of combining individual models [11]. In the recent era, MCS has become widely used in the credit risk assessment domain, surpassing SCS or other predictive methods. This is done to minimize the prediction error caused by different intelligent methods such as decision trees, SVM, Neural Networks (NN), Logistic Regression (LR), and more. There are several reasons why these techniques are widely opted for, including robustness, better performance in handling both linear and non-linear data, prevention of over-fitting or under-fitting, minimization of bias and/or variance, improved stability, and reduced vulnerability against noisy data points [12].

According to the reviewed literature, multiple classifier systems are known by various names such as ensemble methods, combining classifiers, a mixture of experts, committees of learners, and consensus theory [13]. The individual models or learners within these systems are referred to as ensemble members, which can be of either similar types or different types. These ensemble members can use the same training data or be trained on different subsets of datasets. Since the characteristics of credit risk datasets vary, it is not feasible to apply a single intelligent method or classifier to all of them. The ensemble members overcome the drawbacks of a particular base learner by leveraging the benefits in local regions, thus enhancing the final predicted accuracy in credit risk assessment [14].

Before discussing the literature reviews on MCS, it is important to understand the design procedure of MCS, which consists of three steps: pool generation, selection integration, and evaluation of performance, as illustrated in Fig. 5. The design procedure functions as a feedback system, wherein modifications are required at any of the earlier stages if the performance metrics fail to meet the required specifications. The pool generation involves training multiple models to obtain accurate and diverse outputs. In the subsequent phase, a subset of classifiers is selected, and their outputs are combined using various methods such as majority vote, Borda count, Bayesian theory, fuzzy integral, Dempster-Shafer rule, fuzzy rules, neural networks, Markov chains, and stacking, among others [23].

The known methods for generating the classifier's pool are bagging and boosting. Bagging, also known as bootstrap aggregation, is an example of a parallel ensemble method while boosting is an example of a sequential ensemble method. In the sequential method, ensemble members depend on the outcomes of earlier models. Each model in the sequence aims to correct the prediction errors made by the previous

learner. Consequently, the overall performance of the entire system can be improved by assigning weightage to the earlier outcomes. On the other hand, the parallel method involves ensemble members providing independent outputs simultaneously, and the final output is combined using a combination approach. The parallel ensemble method can be further classified into homogeneous and heterogeneous techniques. Homogeneous methods utilize the same base learners with different parameters, whereas heterogeneous methods involve the use of various base learners [14].

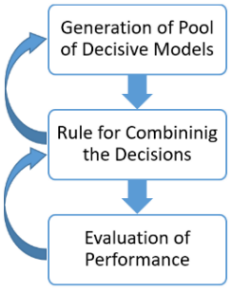

Fig. 5. Design Procedure of MCS [23].

All the classifiers generated by the classifier pool are not useful as they fail to produce the optimal output. Some of them may suffer from accuracy issues, while others produce identical results. Consequently, it becomes necessary to select the most effective learners or create a subset of models from the pool of classifiers based on both accuracy and diversity. This selection process is commonly referred to as ensemble selection [24].

The ensemble members provide complementary information for the data points, aiming to harness the strengths of each member while mitigating the weaknesses of individual learners. The objective is to enhance the overall classification performance of the system. Furthermore, ensemble selection can be categorized into two types: static classifier system and dynamic classifier system, as illustrated in Fig. 6. In a static system, the best classifiers are selected based on accuracy and diversity during the training phase, and the chosen classifier is used to classify all test instances [14]. On the other hand, in the dynamic system, either the best single classifier or an ensemble of classifiers is selected for each new borrower or data point.

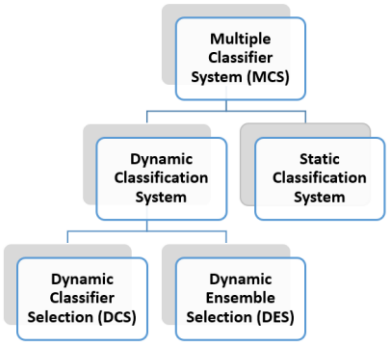

Fig. 6. Types of MCS

The block-diagram of the static classifier system is depicted in Fig. 7 for better understanding. Various well-known ensemble algorithms based on static selection have been extensively investigated for credit risk assessment. These include Random Forest, AdaBoost, Stacking, Gradient Boosting Decision Tree, XGBoost, and LightGBM [14]. If the system selects a single best classifier for each new test instance, it is referred to as Dynamic Classifier Selection (DCS). On the other hand, if the system chooses an ensemble of classifiers, it is known as Dynamic Ensemble Selection (DES) [14].

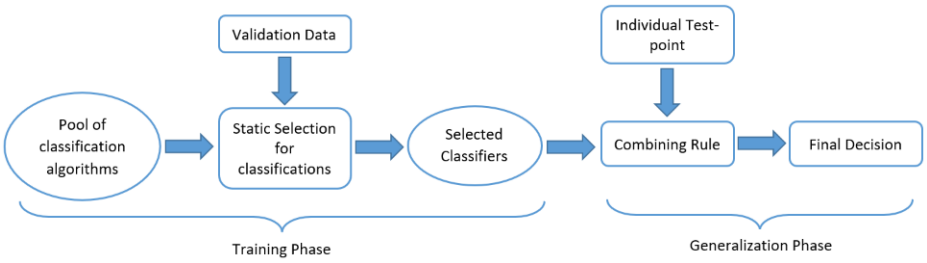

Fig. 7. Block-diagram of SCS [14]

Table 2. Finding and future scope of the reviewed articles based on static ensemble selection

Reference	Findings	Future scopes
Rao, Liu and Goh [15] (2023)	XGBoost inspired by PSO, aided by feature selection technique and Smote-Tomek Link for balancing data, markedly surpassed conventional models in assessing credit risk	Investigation of a novel feature selection technique Use the deep learning methods
Sun and Zhu [16] (2021)	Multi-class classification for risk assessment of corporate using One-Versus-One (OVO) SMOTE AdaBoost ensemble model	Addition of non-financial features along with financial parameters
Tsai and Hung [17] (2021)	AdaBoost was used to evaluate the financial performance of enterprises post-COVID-19 Initialize higher weight to the instances generated after COVID-19	Datasets other than the semiconductor and tourism sector can be explored
Yu et al. [18] (2024)	A combination of SMOTE-ENN and t-SNE was used with LightGBM, XGBoost, and TabNet to enhance credit risk prediction LightGBM achieved the best results	Applying to larger datasets and improving model optimization for more accurate financial risk predictions
Zhang and Li [20] (2021)	Stacking classifier of five ensemble members used to evaluate the risk of internet finance SVM was used in the second stage	Data balancing technique and combinations of ensemble members can be explored for comparative analysis
Ruan, Zhang and Li [21] (2021)	LightGBM was used with a class balancing technique using the k-Means algorithm for evaluating the financial distress situation The ensemble approach was used to reduce the dimensionality	Novel feature selection methods can be explored Experiment can be performed on data from other geography
Yang and Xiao [22] (2024)	Integrating external data sources, used bagging-based oversampling and stacking classifier with an optimized voting-weight technique for risk assessment of MSME	Explore methods across diverse datasets

Many researchers have explored various multiple classifier systems based on static selection in the domain of financial risk assessment. The list of models that have been explored includes Random Forest [8, 31], AdaBoost [16, 17], XGBoost [15, 18], LightGBM [21], and stacking [20]. Before reviewing the methods based on static selection, it is wise to understand the flowchart of this technique. As discussed in the previous paragraph, an ensemble of classifiers is selected during the training phase, and this group of classifiers remains the same for all unknown test data points. Furthermore, the work related to static ensemble selection explored by researchers is also summarized in Table 2.

5. Ensemble method (dynamic selection)

In the case of static ensemble selection, the average competence of classifiers is measured during the training phase using a validation dataset, and the same chosen classifier is used for all unknown instances or borrowers. However, this approach is unlikely to yield the best outcomes for all new test instances. This drawback has led to the development of an alternative ensemble-based dynamic selection approach that aims to select a competent and unique classifier for each new instance or borrower. By doing so, the probability of accuracy can be improved, thereby helping financial institutions minimize monetary losses. This alternative approach is known as dynamic selection and can be further divided into two categories: DCS and DES.

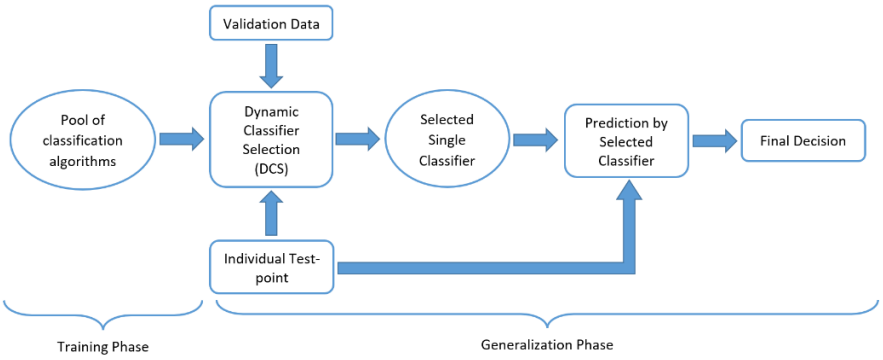

Fig. 8. Block-diagram based on DCS-based classification system [14]

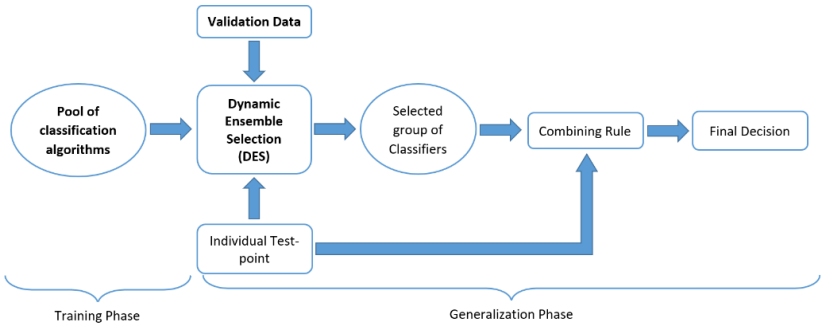

Fig. 9. Block-diagram based on DES-based classification system [14]

Figs 8 and 9 provide a visual representation of both DES and DCS-based classification systems, respectively. In the first dynamic selection technique, only a single classifier with the highest competence is selected for each new test instance. In the second case, a group or ensemble of classifiers is selected based on the most competent classifiers available [14]. To assess the competence of the classifier(s), the local region surrounding the test instance is taken into consideration.

As mentioned previously, selection approaches in the DCS can be further divided into two categories: DCS and DES selection techniques. Numerous articles have been published on these topics, exploring various methods based on classifier competency within the local region of the feature space and the selection criterion used to choose the classifier(s). The classifier competency can be assessed using techniques such as K-Nearest Neighbor (KNN), clustering, base learners' decisions, or potential functions. The selection criterion itself is categorized into two sub-categories: (1) Individual and (2) Group-based groups. These sub-categories can be further divided into branches, as illustrated in Fig. 10.

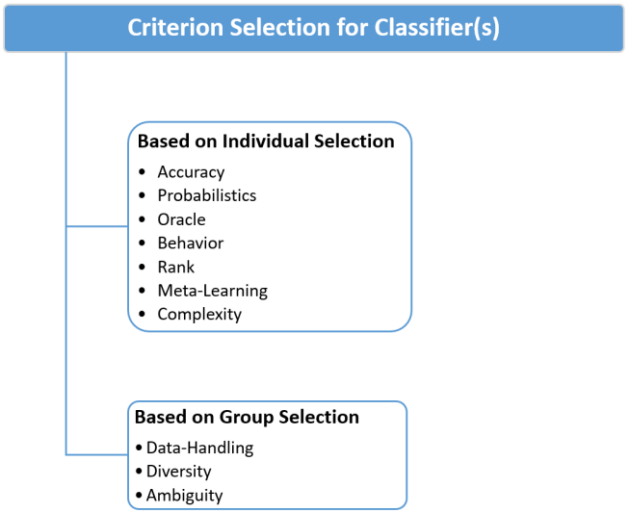

Fig. 10. Criteria for classifier(s) selection [30]

Additionally, crucial information regarding various classifier selection approaches and methods is summarized in Table 3 for DCS-based techniques and Table 4 for DES-based techniques.

The region surrounding the query, where the competency of the base learner needs to be estimated, is referred to as the region of competence. The performance of the system largely relies on how this local region is defined, as Dynamic Selection methods in these regions are sensitive to data distribution. Previous literature suggests that Dynamic Selection performance can be improved by defining these local regions using various techniques. The nearby region can be determined using methods such as k-NN, k-Means, potential function, or the decision space of the base learner. This requires a dataset known as the Dynamic SELection dataset (DSEL). Summary Table 5 provides an overview of the different approaches for defining the local region and the corresponding selection methods [30].

Table 3. Summary table of the DCS-based selection approach of the DS with selection criterion [13, 30]

DCS-based selection technique	Selection criterion	Remark on classifier selection
Modified classifier ranking	Ranking	Number of rightly classified instances in consecutive manner – in the region of competence [30]
Overall local accuracy	Accuracy	% of samples – rightly classified – in the local region (region of competence) [34]
Local classifier accuracy	Accuracy	% of samples – rightly classified with reference to a particular class – in the region of competence [13]
Modified local accuracy	Accuracy	Weight each point in feature space by the distance of the test point – aimed to solve the issue of the size of the local region [13]
Dynamic selection on complexity	Complexity, & Accuracy	Considering three parameters – two parameters are for complexity and remaining for accuracy in the local region
Multiple Classifier Behavior (MCB)	Behavior	Vector computation of test-point – local region around test-point – Similarity between MCB vectors of (1) test point, and (2) a point of local region – Calculates overall local accuracy – threshold values
A priori	Probabilistic	Probability of rightly classified instances in the local region by an individual classifier
A posteriori	Probabilistic	Probability of rightly classify test instance to particular class-label by an individual classifier

Table 4. Summary table of the DES based selection approach of the DS with selection criterion [13, 30, 35]

DES-based selection technique	Selection criterion	Remark on classifiers selection
DES-clustering	Accuracy, and Diversity	K-clusters using k-Means – Classifiers ranked as per (1) accuracy (Descending), and (2) diversity (Ascending)
DES-KNN	Accuracy, and Diversity	KNN used for region of competence – Classifiers ranked as per (1) accuracy (Descending), and (2) diversity (Ascending)
k-Nearest Oracles	Oracle	four methods – (1) KNORA-E, (2) KNORA-U (3) KNORA-E-W (4) KNORA-U-W
Randomized reference classifier	Probabilistic	Base learner which perform better to random learner – Dependency of competency of base learner are (1) the source competency, and (2) the Gaussian-potential function
DES-P	Probabilistic	Classifier competency is computed – taking difference – accuracy of base learner in local region and accuracy of classifier randomly chosen the same local region
DES-KL	Probabilistic	Competent base classifier computed as per Kullback-Leibler divergence – to weight competent source, Gaussian based potential functions used
KNOP	Behavior	Similarity between output profiles of query-instance and samples from dynamic-selection database – Most similar *K* instances from dynamic-selection database selected
META-DES	Meta-learning	five criterions are to be chosen to assess the competency of base classifier to select the learner or not to classify the test-instance. – Meta-features are used to train the meta-classifier
META-DES.Oracle	Meta-learning	Variant of META-DES – Binary PSO was used to select optimum meta-features among 15 meta-features to have the optimum performance of meta-learner

Table 5. Summary table of approaches for defining region of competence and corresponding selection methods [13, 30]

Approach for defining region of competence	Dynamic Selection Methods	Remarks
k-NN	Classifier Rank, OLA, LCA, A priori, A posteriori, MLA, KNORA, META-DES.Oracle, DSOC	• Defining of nearest neighbors of query instance in DSEL data-sets • Data-points of local regions are considered for competent base classifier • More precise estimation local region than Clustering approach • Involved higher computational cost • Different variants of KNN should be explored for better estimates local region
k-Means (clustering)	DES-Clustering	• Defining of clusters in DSEL data-sets • Distance calculation between centroid of cluster and test-query • Competent classifier is measured upon the instances belonging nearest cluster • Fast in generalization phase
Decision space	MCB, KNOP, Meta-DES	• Dependency is on the decision yields of base classifiers • Training-Testing data-points required to transform into output profiles • Selection of most similar profiles of DSEL with the profile of query
Potential function	DES-Performance, Randomized reference classifier, Kullback-Leibler	• Whole DSEL data-set used instead of local region • Each instance in the DSEL is weighted as per Euclidean distance with reference to query instance • Function based on Gaussian potential • Avoiding the need of setting the size of K in local region • Increases computational complexity

6. Class-imbalance challenges & probable solution

In the real world, many domains require classification tasks that involve classifying instances into different class labels. As the world embraces digitalization, it is unlikely for any domain to be left behind without utilizing the latest technologies such as machine learning and data mining. For example, in credit risk assessment, the classification problem may involve detecting bankruptcy and/or fraud. In such cases, there are very few instances belonging to the positive class (i.e., bankrupt and/or fraudulent detection), while the majority of data points belong to the negative class (i.e., non-bankrupt and/or genuine detection). However, accurately classifying rare instances becomes challenging when the distribution of instances across classes is imbalanced, with minority instances being rare compared to majority instances. Classifiers rely on a balanced distribution of instances across classes in the dataset to avoid biased outcomes favouring the majority class [37].

The classification task faces several difficulties, especially when dealing with imbalanced data, which refers to skewed data distributions where the minority class is underrepresented. These difficulties pose challenges for algorithms to effectively classify instances. Here are some key points to consider:

- Standard models like logistic regression and decision trees may yield sub-optimal classification performance due to the imbalanced distribution of data across classes [36].
- The objective of standard classifiers is to optimize performance measures such as recall, accuracy, and precision. However, these measures often bias towards the majority data points, resulting in the neglect of minority samples, even if the overall precision is high [36].
- The rarity of minority instances in the sample space makes it more likely for them to be misclassified as noise and vice versa [36].
- Skewed data distribution is less problematic as long as suitable class separability is maintained. However, when rare instances overlap with other classes in different regions where the probability of misclassification is almost equal, it becomes challenging [36].
- High dimensionality and extreme imbalance in data sets can lead to incorrect classification of minority samples by the classifier model [36].

To summarize, the classification of imbalanced data poses various challenges, and standard classifiers may not perform optimally. It is crucial to address these issues to improve classification accuracy, especially for minority class instances.

For datasets with skewed distributions, the imbalance ratio refers to the ratio of the number of instances belonging to the majority class to the number of instances belonging to the minority class. This parameter is widely recognized as a challenging factor for classifiers when it comes to accurately classifying the minority instances. The performance of the classifier tends to be fragmented, often resulting in the decomposition of instances from the minority classes, leading to disjointed and sub-conceptual understanding. Table 6 displays the data distribution of the minority class, encompassing examples categorized as safe, rare, borderline, and outliers. Typically, examples are classified based on whether they are considered safe or unsafe. The unsafe examples can be further categorized into noisy and borderline examples. In imbalanced data scenarios, the presence of noisy examples adversely affects the outcomes of the classifier, as they introduce attribute or class errors [40].

Table 6. Distribution of minority class [40]

Safe examples	Borderline examples	Outlier examples	Rare examples
Found in the area surrounded by one class examples Can be in homogenous area Easier for classifier to learn	Found in the area near to decision boundary of classes Might be found in the overlapping area of majority and minority class samples Might be wrongly classified from the opponent class situated on the boundary from the other side	Care should be taken care in treating such examples as noise Can be rare represented but valid sub-concepts to which there isn't any representation collected during the training phase	Found in region of majority in the group of triples or pairs Far from class decision boundary

Data complexity refers to the level of difficulty in classifying a given dataset, particularly in a 2-class classification problem. Data complexity can be categorized into three main categories: (1) Overlap measurement among classes in the feature space, (2) Class separability, and (3) Topology, Geometry, & manifolds' density.

Table 7 presents the different techniques used to measure various aspects of data complexity.

Table 7. Measure of data complexity with various methods [38]

Measure of data complexity	Name of methods
Overlap measurement among classes in the feature space	Feature efficiency, Maximum Fisher discriminant-ratio, Volume of overlap-region,
Class separability	Mixture identifiability, Linear separability
Topology, Geometry, and manifolds' density	Nonlinearity (Average number of points)/dimension

Fig. 11 illustrates the overlap between classes, with circles representing the negative class and triangles representing the positive class. Specifically, Fig. 11i showcases the overlap between two classes, with circles denoting the negative class and triangles representing the positive class. It is important to note that the presence of disjuncts, which are typically formed through different sub-concepts or sub-classes within a single class, can have a detrimental effect on the performance of the classifier [39]. The overlapping of classes and the presence of disjuncts introduce complexity, making it more challenging to distinguish between classes and necessitating the adoption of more stringent classification rules.

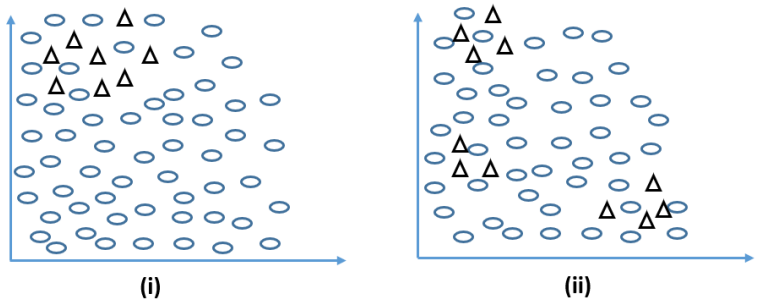

Fig. 11. Examples of (i) overlapping of classes, (ii) small disjuncts [39]

From the literature survey, it can be inferred that techniques for addressing skewed data distribution can be classified into two main categories: (1) Data-level approaches, and (2) Algorithm-level approaches. Both categories can be further divided into sub-categories. Data-level techniques primarily focus on the distribution of sample classes in datasets, which involves either adding minority class samples or removing majority class samples. Algorithm-level techniques, on the other hand, aim to modify classifier algorithms to reduce bias towards majority class samples.

As shown in Fig. 12, the data-level approach can be further divided into data sampling methods and feature selection methods. In data sampling methods, the distribution of data during the training phase of the classifier is altered to improve the correct classification of minority instances. Simple techniques like Random Under-Sampling (RUS) and Random Over-Sampling (ROS) are commonly used. Additionally, SMOTE is an intelligent over-sampling method that creates synthetic minority instances by interpolating among existing minority instances, generating new instances close to the existing ones. This approach helps balance the skewed distribution between minority and majority class instances. However, a major

drawback of over-sampling methods is the increase in the training dataset size, which can lead to overfitting. Nonetheless, SMOTE handles the overfitting issue better than simple random over-sampling [41].

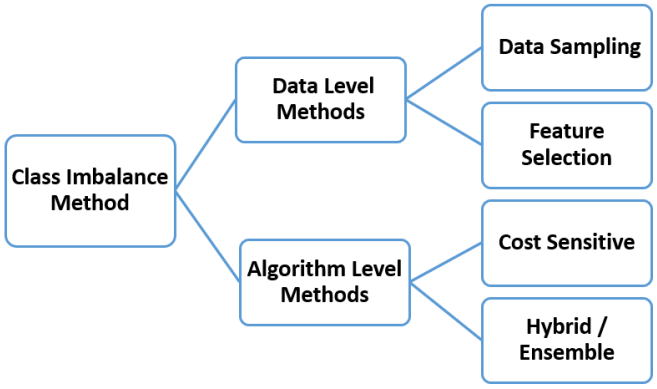

Fig. 12. Categories of techniques addressing techniques skewed data distribution [36]

The under-sampling method, aimed at balancing the instances of the minority and majority classes, presents a significant challenge. One major concern is the loss of crucial information due to the removal of majority class instances. To address this issue, researchers have delved into different data-sampling methods, which are outlined in Table 8.

Table 8. Various types of sampling and its techniques [43]

Sampling type	Name of the methods
Under-sampling	With replacement under-sampling majority randomly, Tomek-links of majority-minority, under-sampling using cluster centroid, Near miss, Condensed nearest-neighbour, One-Sided Selection, Neighbourhood-cleaning Rule, Edited Nearest-Neighbour (ENN), Threshold based on Instance-hardness, Repeated edited nearest-neighbours, Random Over-Sampling Example (ROSE)
Over-sampling	With replacement over-sampling minority randomly, SMOTE, 85 SMOTE variants, ADASYN
Over sampling + Under sampling	SMOTE + Tomek links, SMOTE + ENN
Hybrid/Ensemble methods	SMOTEBoost, RUSBoost, LIUBoost, RHSBoost, HUSBoost

In addition to the data-sampling method, another technique used to address skewed data distribution is the feature-selection method. The functional diagram of this method is depicted in Fig. 13. The objective of feature selection is to choose a subset of features from the entire feature space, enabling the algorithm to achieve optimum performance. The number of features in the subset can be either adaptively selected or chosen by the user [36]. The feature-selection method is further divided into three techniques: Wrapper, Filter, and Embedded.

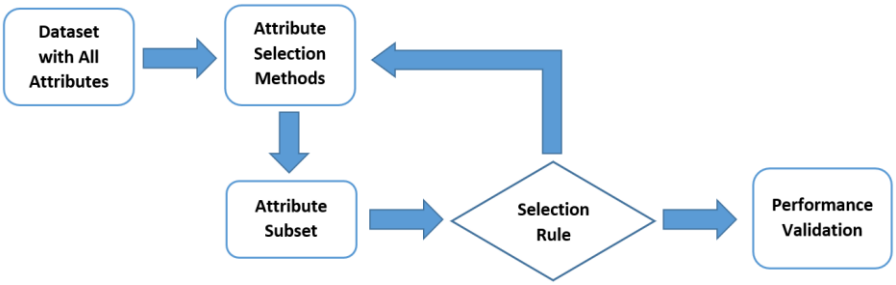

Fig. 13. Process of feature selection [42]

The filtering technique filters out features or attributes based on statistical tests. The wrapper technique is a type of search problem that requires adopting a greedy search approach to assess all possible combinations of feature subsets on the classifier's performance. As the name suggests, the embedded technique combines both the filter and wrapper techniques. The summary table for various feature selection methods and their evaluation schemes is presented in Table 9.

Table 9. Various feature selection techniques and evaluation methods [25]

Feature selection method	Evaluation methods for selected features
Filter method	Information Gain, Chi-square test, Fisher's score, correlation coefficient, variance threshold, Mean absolute difference, dispersion ratio, mutual dependence, relief
Wrapper method	Forward selection, Backward elimination, Bi-directional elimination, Exhaustive selection, recursive elimination
Embedded method	Regularization, tree-based methods (Gradient boosting, Random forest, etc.)

As shown in Fig. 12, the algorithm level approach can be further divided into two categories: cost-sensitive methods and hybrid/ensemble methods. Generally, algorithm level methods aim to learn or make decisions that enhance the significance of minority class samples or positive samples. In these methods, the base classifiers are modified to accept weights or class penalties, or the decision threshold is adjusted to increase the bias towards minority instances. In cost-sensitive methods, penalties are assigned to each class through a cost-matrix table. By increasing the cost of the positive class, the chances of incorrectly classifying minority instances can be reduced. Increasing the cost of a class implies increasing its importance. However, a major challenge for cost-sensitive learners is determining the appropriate cost values, which can be based on past experience or obtained from domain experts [36]. Another technique, apart from the cost-sensitive method, to handle skewed data distribution is the hybrid/ensemble method. As the name implies, this approach combines two or more individual methods to tackle the issue of skewed data distribution, or it combines multiple algorithms to enhance the classification rate for minority instances. Well-known algorithms like Bagging, Boosting, and their variations fall into this category.

Here, an attempt has been made to explore novel algorithms by utilizing the latest meta-learning based Meta-DES framework. The DES framework comprises three phases: over-production, meta-training, and generalization [14]. The over-

production phase, also referred to as the generation phase, focuses on generating a diverse and accurate pool of classifiers [14], as depicted in Fig. 14. The numbered circles in Fig. 14 represent the different alternative approaches that can be considered for each block, allowing for the exploration of a novel pool of classifiers.

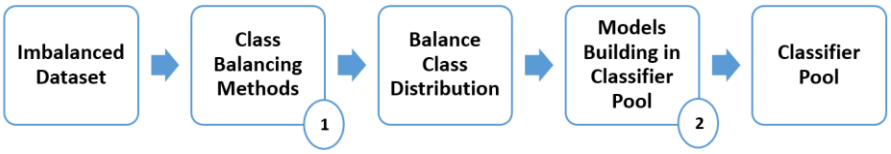

Fig. 14. Processing steps for classifier pool

1. A variety of sampling methods, as illustrated in Table 8, can be explored. Additionally, the combinations of class-sampling and feature-selection methods can be explored, as demonstrated in Table 8 and Table 9.
2. Various existing and novel classifiers can be trained, which are expected to produce accurate and diverse outputs.

The second step in Meta-DES is the meta-training phase, which is further divided into three sub-sections: sample selection, meta-feature extraction, and training of the meta-classifier. Fig. 15 depicts the block diagram of the sample selection process. During sample selection, a specific instance is chosen if the consensus among the classifiers in the pool regarding that instance is below the threshold consensus. The consensus degree is determined by the difference in votes between the winning class and the second class.

3. Novel consensus degree of classifiers can be explored
4. Various values of the preset consensus threshold can be experimented with to achieve the best final result

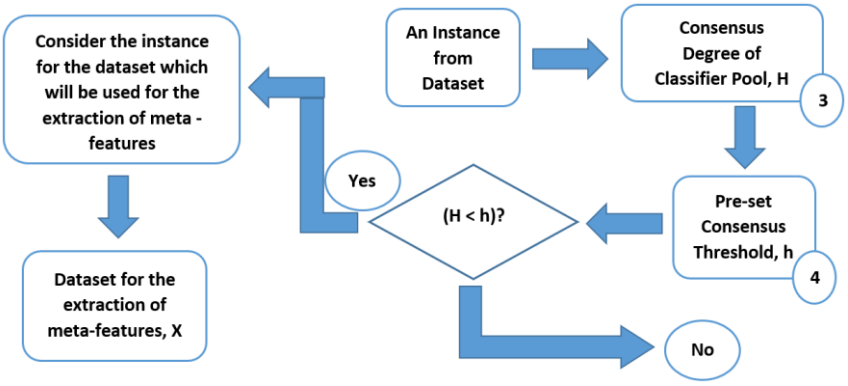

Fig. 15. Functional diagram of sample selection

The combined block diagram illustrating the process of meta-feature extraction and meta-classifier training is presented in Fig. 16. In this phase, a crucial objective is to evaluate the performance of all base classifiers in two key domains: the feature-space and decision-space. This evaluation is conducted using diverse criteria to extract the meta-features. The resulting meta-features, together with the meta-label, are then utilized to generate the dataset used for training the meta-classifier.

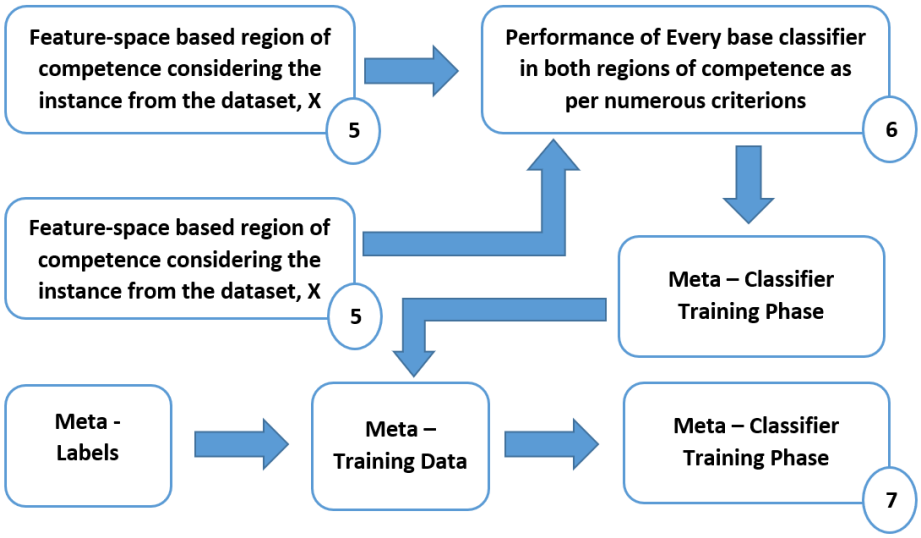

Fig. 16. Functional diagram of combination of meta-feature extraction and training of meta-classifier

5. Numerous approaches for defining the region of competence can be explored, as shown in Table 5.

6. Several criteria can be explored to assess the competency of a base classifier.

7. Various existing and novel classifiers can be trained as a meta-classifier, which is likely to yield accurate and diverse outputs.

The final step in the Meta-DES framework is generalization, where a testing instance is given as input. The third stage of Meta-DES is depicted in Fig. 17. In this stage, the region of competence in both the feature and decision space needs to be determined for the given testing instance. This information is essential to identify the relevant meta-feature vector and perform classifier competency checking. Additionally, the combination of outputs from the ensemble of classifiers also plays a crucial role in this step.

8. Numerous combining strategies can be explored in Meta-DES framework as separate combining methods are mentioned in [13, 14]

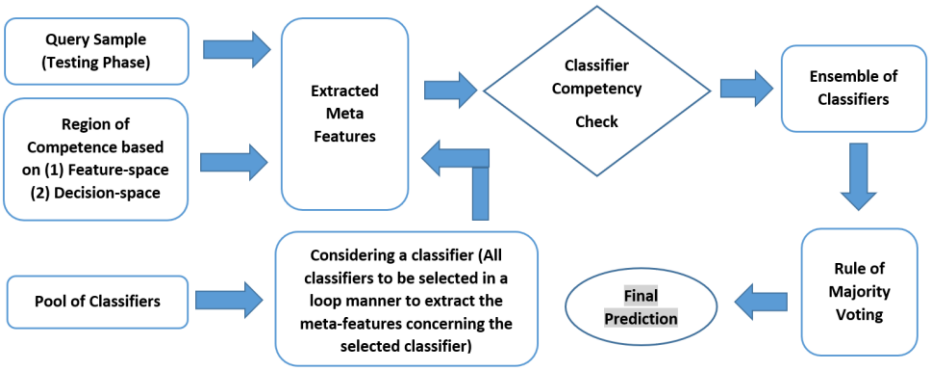

Fig. 17. Functional diagram of generalization phase of meta-classifier

7. Research findings

In this article, we aim to emphasize the importance of credit risk assessment from an economic perspective for financial institutions. In this review-based research, various statistical methods for classification tasks are mentioned, along with the design steps involved. However, these methods have been replaced with intelligent techniques due to their inability to consider parameters like overlapping class boundaries and relationships among financial features. Additionally, we explain the design procedure for single/multiple classifier systems and provide summary tables of the reviewed articles.

From the summarized tabular data, we find that the latest algorithms such as SVM with kernels, decision trees, logistic regression, random forests, multilayer perceptron, combination of SVM-PSO, back-propagation neural networks, OVO-SMOTE-AdaBoost ensemble models, extreme gradient boosting tree classifiers, and stacking-based ensemble approaches have been explored for credit risk classification. Furthermore, various class-balancing techniques like SMOTE and under-sampling-based clustering approaches have been used in conjunction with some of the aforementioned classification techniques.

We also elaborate on the different types of multiple classifier systems, namely static and dynamic systems, and discuss the difference between Dynamic Classifier Selection (DCS) and Dynamic Ensemble Selection (DES) in detail. The key finding regarding dynamic selection is the identification of competent classifier(s) based on the Region Of Competence (ROC), which can be defined using techniques such as KNN, clustering, decision space, and potential functions. Competent classifiers are selected based on criteria such as accuracy, ranking, complexity, behavior, probabilistic, oracle, meta-learning, etc.

Additionally, we delve into the challenges posed by imbalanced datasets in classification and suggest addressing them through sampling techniques, feature-selection methods, and ensemble/hybrid approaches. We propose further exploration of novel data balancing and feature selection techniques, as these two methods significantly impact the classification of minority positive instances. Furthermore, considering macroeconomic parameters could enhance the performance of the classifier.

In the context of DCS and DES-based approaches, there is room for investigating numerous permutations and combinations of classifier selection criteria and approaches for defining ROC within existing frameworks such as Meta-DES. Additionally, novel balancing and feature selection schemes could be explored.

8. Future scope

This article delves into a comprehensive exploration of various techniques, ranging from traditional statistical methods to cutting-edge ensemble techniques, for the purpose of assessing credit risk. The primary objective is to minimize both the potential for missed opportunities and financial losses in this context. However, it’s essential to note that the applicability of the Meta-DES technique extends beyond

credit risk assessment; it finds relevance in diverse sectors such as healthcare, agriculture, cybersecurity, and more.

In the pursuit of enhancing Meta-DES performance, several avenues warrant consideration:

- A novel approach to class label balancing could simplify the task for classifier pool members, potentially boosting overall effectiveness.
- To foster diversity within the classifier pool, one can explore different learning mechanisms and approaches. This might encompass training member classifiers on distinct sample sets and feature sets. Furthermore, the inclusion of both homogeneous and heterogeneous members or a combination thereof can be beneficial.
- Evaluating the competence of base classifiers is crucial, especially when dealing with imbalanced data. The exploration of various meta-features in both feature-space and decision-space can provide valuable insights into classifier competence across different dimensions, further enhancing overall performance.

9. Conclusion

In this article, we underscore the significance of credit risk assessment in the financial industry and discusses various methods and techniques employed in credit risk classification. We highlight the latest algorithms, class-balancing techniques, and the design of single/multiple classifier systems. Moreover, we emphasize the challenges posed by imbalanced datasets and suggests potential solutions through sampling, feature selection, and ensemble/hybrid approaches. Lastly, we identify opportunities for further research in the field, including the investigation of classifier selection criteria, approaches for defining the region of competence, and the exploration of novel balancing and feature selection techniques.

References

1. S u n a i n a a, C. Gross NPAs of Public Sector Banks Double in Last Seven Years, SBI Tops List. – The Times of India, 2021 (Accessed 5 April 2023). **https://timesofindia.indiatimes.com/business/india-business/gross-npas-of-public-sector-banks-double-in-last-seven-years-sbi-tops-list/articleshow/88316357.cms**
2. B h a t t a c h a r y y a, R. Education Loan Demand Hits New High Despite Pandemic. – The Economic Times, (2021) (Accessed 5 April 2023). **https://economictimes.indiatimes.com/industry/banking/finance/education-loan-demand-hits-new-high-despite-pandemic/articleshow/81090597.cms**
3. C h e n, N., B. R i b e i r o, A. C h e n. Financial Credit Risk Assessment: A Recent Review. – Artificial Intelligence Review, Vol. **45**, 2016, No 1, pp. 1-23.
4. Y u a n, Z. Research on Credit Risk Assessment of P2P Network Platform: Based on the Logistic Regression Model of Evidence Weight. – Journal of Research in Business, Economics and Management, Vol. **10**, 2018, No 2, pp. 1874-1881.
5. C h e r n, C.-C., et al. A Decision Tree Classifier for Credit Assessment Problems in Big Data Environments. – Information Systems and e-Business Management, Vol. **19**, 2021, No 1, pp. 363-386.

6. M o s c a t o, V., A. P i c a r i e l l o, G. S p e r l í. A Benchmark of Machine Learning Approaches for Credit Score Prediction. – Expert Systems with Applications, Vol. **165**, 2021, 113986.
7. P u t r i, N. H., M. F a t e k u r o h m a n, I. M. T i r t a. Credit Risk Analysis Using Support Vector Machines Algorithm. – Journal of Physics: Conference Series, Vol. **1836**. No 1, IOP Publishing, 2021.
8. B u l u t, C., E. A r s l a n. Comparison of the Impact of Dimensionality Reduction and Data Splitting on Classification Performance in Credit Risk Assessment. – Artificial Intelligence Review, Vol. **57**, 2024, No 9, 252.
9. D u, G., Z. L i u, H. L u. Application of Innovative Risk Early Warning Mode under Big Data Technology in Internet Credit Financial Risk Assessment. – Journal of Computational and Applied Mathematics, Vol. **386**, 2021, 113260.
10. M a, Z., W. H o u, D. Z h a n g. A Credit Risk Assessment Model of Borrowers in P2P Lending Based on BP Neural Network. – PLOS ONE, Vol. **16**, 2021, No 8, e0255216.
11. B r o w n l e e, J. Why Use Ensemble Learning? – In: Machine Learning Mastery, 2020.
12. M a k h i j a n i, C. Advanced Ensemble Learning Techniques. – In: Towards Data Science. 2020.
13. A b d o l i, M., M. A k b a r i, J. S h a h r a b i. Dynamic Ensemble Learning for Credit Scoring: A Comparative Study. arXiv Preprint arXiv:2010.08930, 2020.
14. H o u, W.-h., et al. A Novel Dynamic Ensemble Selection Classifier for an Imbalanced Data Set: An Application for Credit Risk Assessment. – Knowledge-Based Systems, Vol. **208**, 2020, 106462.
15. R a o, C., Y. L i u, M. G o h. Credit Risk Assessment Mechanism of Personal Auto Loan Based on PSO-XGBoost Model. – Complex & Intelligent Systems, Vol. **9**, 2023, No 2, pp. 1391-1414.
16. S u n, J., J. Z h u. Multi-Class Imbalanced Corporate Bond Default Risk Prediction Based on the OVO-SMOTE-Adaboost Ensemble Model. – In: Proceedings of CECNet 2021. IOS Press, 2021, pp. 42-53.
17. T s a i, J.-K., C.-H. H u n g. Improving AdaBoost Classifier to Predict Enterprise Performance After COVID-19. – Mathematics, Vol. **9**, No 18, 2021, 2215.
18. Y u, C., et al. Advanced User Credit Risk Prediction Model Using Lightgbm, Xgboost and Tabnet with Smoteenn. arXiv Preprint arXiv:2408.03497, 2024.
19. G a o, X., X. Y a n g, Y. Z h a o. Rural Micro-Credit Model Design and Credit Risk Assessment via Improved LSTM Algorithm. – PeerJ Computer Science, Vol. **9**, 2023, e1588.
20. Z h a n g, T., J. L i. Credit Risk Control Algorithm Based on Stacking Ensemble Learning. – In: Proc. of IEEE International Conference on Power Electronics, Computer Applications (ICPECA'21), IEEE, 2021.
21. R u a n, S., J. Z h a n g, W. L i. CUS-LightGBM-Based Financial Distress Prediction for Small-and Medium-Sized Enterprises with Imbalanced Data. Science Direct, 2021.
22. Y a n g, D., B. X i a o. Feature Enhanced Ensemble Modeling with Voting Optimization for Credit Risk Assessment. – IEEE Access, 2024.
23. Y a n g, L.-Y., Z. Q i n, R. H u a n g. Design of a Multiple Classifier System. – Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No 04EX826), Vol. **5**. IEEE, 2004.
24. M o h a m m e d, A. M., E. O n i e v a, M. W o ź n i a k. Selective Ensemble of Classifiers Trained on Selective Samples. – Neurocomputing, Vol. **482**, 2022, pp. 197-211.
25. I s l a m, M. R., et al. A Comprehensive Survey on the Process, Methods, Evaluation, and Challenges of Feature Selection. – IEEE Access, Vol. **10**, 2022, pp. 99595-99632.
26. K u m a r, C. Education Loan NPAs: Nursing, Engg Students Bigger Defaulter Than Those in MBA, Medicine. – Times of India, 2021 (Accessed 5 April 2023). **https://timesofindia.indiatimes.com/business/india-business/education-loans-nurses-engineering-students-have-more-npas-than-mbas-medicos/articleshow/81622856.cms**
27. F e n e r i c h, A., et al. Use of Machine Learning Techniques in Bank Credit Risk Analysis. – Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería, Vol. **36**, 2020, No 3.
28. B o r c h e r t, P., et al. Extending Business Failure Prediction Models with Textual Website Content Using Deep Learning. – European Journal of Operational Research, Vol. **306**, 2023, No 1, pp. 348-357.

29. K i m, M. H., P. D. Y o o. A Semiparametric Model Approach to Financial Bankruptcy Prediction. – Proc. of IEEE International Conference on Engineering of Intelligent Systems. IEEE, 2006.
30. C r u z, R. M. O., R. S a b o u r i n, G. D. C. C a v a l c a n t i. Dynamic Classifier Selection: Recent Advances and Perspectives. – Information Fusion, Vol. **41**, 2018, pp. 195-216.
31. L i u, Y., et al. Leveraging Network Topology for Credit Risk Assessment in P2P Lending: A Comparative Study under the Lens of Machine Learning. – Expert Systems with Applications, Vol. **252**, 2024, 124100.
32. Q u a n, J., X. S u n. Credit Risk Assessment Using the Factorization Machine Model with Feature Interactions. – Humanities and Social Sciences Communications, Vol. **11**, 2024, No 1, pp. 1-10.
33. C h e n, B., W. J i n, H. L u. Using a Genetic Backpropagation Neural Network Model for Credit Risk Assessment in the Micro, Small and Medium-Sized Enterprises. – Heliyon, Vol. **10**, 2024, No 14.
34. W o o d s, K., W. P. K e g e l m e y e r, K. B o w y e r. Combination of Multiple Classifiers Using Local Accuracy Estimates. – IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. **19**, 1997, No 4, pp. 405-410.
35. C r u z, R. M. O., R. S a b o u r i n, G. D. C. C a v a l c a n t i. META-DES. Oracle: Meta-Learning and Feature Selection for Dynamic Ensemble Selection. – Information Fusion, Vol. **38**, 2017, pp. 84-103.
36. M a h a d e v a n, A., M. A r o c k. A Class Imbalance-Aware Review Rating Prediction Using Hybrid Sampling and Ensemble Learning. – Multimedia Tools and Applications, Vol. **80**, 2021, No 5, pp. 6911-6938.
37. K u l k a r n i, A., D. C h o n g, F. A. B a t a r s e h. Foundations of Data Imbalance and Solutions for a Data Democracy. – Data Democracy, Academic Press, 2020, pp. 83-106.
38. E r w i n, K., A. E n g e l b r e c h t. Feature-Based Complexity Measure for Multinomial Classification Datasets. – Entropy, Vol. **25**, 2023, No 7, 1000.
39. S u n, Y., et al. A Robust Oversampling Approach for Class Imbalance Problem with Small Disjuncts. – IEEE Transactions on Knowledge and Data Engineering, Vol. **35**, 2022, No 6, pp. 5550-5562.
40. M a u l i d e v i, N. U., K. S u r e n d r o. SMOTE-LOF for Noise Identification in Imbalanced Data Classification. – Journal of King Saud University-Computer and Information Sciences, Vol. **34**, 2022, No 6, pp. 3413-3423.
41. A g u i a r, G., B. K r a w c z y k, A. C a n o. A Survey on Learning from Imbalanced Data Streams: Taxonomy, Challenges, Empirical Study, and Reproducible Experimental Framework. – Machine Learning, Vol. **113**, 2024, No 7, pp. 4165-4243.
42. A g r a w a l, P., et al. Metaheuristic Algorithms on Feature Selection: A Survey of One Decade of Research (2009-2019). – IEEE Access, Vol. **9**, 2021, pp. 26766-26791.
43. H a s i b, K. M., et al. A Survey of Methods for Managing the Classification and Solution of Data Imbalance Problem. arXiv Preprint arXiv:2012.11870, 2020.

Received: 10.06.2024; Second version: 30.09.2024; Accepted: 13.10.2024.