Journal of Mahani Mathematical Research
Print ISSN: 2251-7952 Online ISSN: 2645-4505

A REVIEW ON COST-BASED FEATURE SELECTION ALGORITHMS IN THE VARIOUS APPLICATIONS OF MACHINE LEARNING

S. Beiranvand, M.B. Dowlatshahi ✉, and A. Hashemi

Article type: Research Article
(Received: 10 September 2024, Received in revised form 21 April 2025)
(Accepted: 18 June 2025, Published Online: 25 June 2025)

Abstract. Knowledge acquisition is the most important challenge in building an expert system in any field, and one of the sources of knowledge will be the data collected in that field. Traditionally, the data collection process is assumed to have a symmetric cost. For example, this assumption will not be acceptable in the medical due to various expenses. Designing a cost-sensitive classification and a cost-sensitive feature selection method are two approaches to considering cost factors. Cost-effective feature selection improves financial return by significantly saving feature data cost as well as limiting credit losses and this can be used in different areas, for example, computer imaging and medical diagnosis which also have a large number of features that may be irrelevant or redundant. Analysis of the research reviewed in this study shows that cost-sensitive feature selection focuses on selecting a feature subset with minimum total cost while achieving a classification accuracy that is as high as possible. The review of selected studies showed a downward trend in using heuristic methods in this field, Wrapper methods are in the first rank regarding usage in evaluation criteria, and 76% of selected studies are in the single-objective category. Most of the studies were classified in the single-label category based on the number of determined labels.

Keywords: Cost-based approaches, Cost-sensitive classification, Feature selection, Single-label data.
2020 MSC: 68T20.

1. Introduction

The main challenge in building expert systems is acquiring knowledge. One source of knowledge could be data collected on the same problem. Today, many experts have focused on the problems of acquiring knowledge from data using machine learning methods (Ciupke, 2006). The input data cardinality, i.e., the number of independent features, dramatically affects the model built from the data by machine learning methods. Studies have shown that with the existence of irrelevant and redundant features in the original data, the obtained model is overfitted and negatively affects the model (Beiranvand & Chahooki, 2016;

✉ dowlatshahi.mb@lu.ac.ir, ORCID: 0000-0002-4862-0626
https://doi.org/10.22103/jmmr.2025.24028.1696

Publisher: Shahid Bahonar University of Kerman
How to cite: S. Beiranvand, M.B. Dowlatshahi, A. Hashemi, *A review on cost-based feature selection algorithms in the various applications of machine learning*, J. Mahani Math. Res. 2027; 16(1): 1-44.

Beiranvand & Zare Chahooki, 2023). Therefore, the collected data must be preprocessed before being used for learning. Preprocessing includes various methods, one of the best approaches is to choose a practical group of features. The general performance of machine learning methods can be enhanced by using just an optimal group of high-dimensional features (Ali, Khan, et al., 2020).

In general, feature selection algorithms have four main steps, as shown in Figure 1. The most important of the four phases involves subset creation and subset evaluation. The candidate feature subsets are generated in the feature generation step. A search method is used to measure the quality of candidate feature subsets using an evaluation function in the subset evaluation phase. According to the impact of the subset evaluation phase, it is expected that more practical subsets of features generated in the subset generation phase (Nguyen et al., 2020).

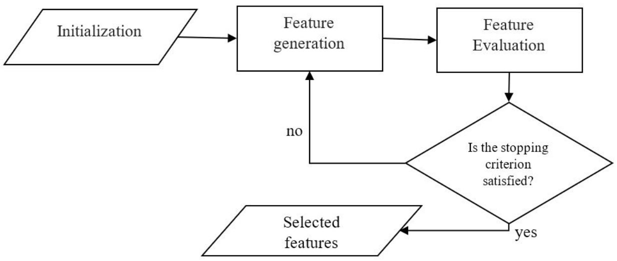

Figure 1. Four main steps in a feature selection algorithm.

Datasets can have redundant and irrelevant features that adversely affect the machine learning methods. Features that contribute to the machine learning methods should be detected using feature selection techniques. Irrelevant features provide no useful information to the machine learning process, and redundant features provide duplicate information (Büyükkeçeci & Okur, 2022).

Selecting a small set of informative features from many possibly noisy candidates is a challenging problem with many applications in machine learning and approximate Bayesian computation. In practice, the cost of computing informative features also needs to be considered (Raynal et al., 2023). Most feature selection techniques assume the data are already stored in data sets and available without incurring any cost. However, data are not free in real-world applications(Sheng et al., 2006). Traditionally, the symmetric cost of data collection is an important assumption when building learning models. In some applications like medical diagnoses, this assumption is invalid in cases where the cost of obtaining information is asymmetric; For example, the difference in the cost of various medical examinations. Costs can also be calculated in economics and the availability of certain medical tests, the risk of infection, and the computational cost of obtaining information. Therefore, the concept of cost can be considered a new aspect of evaluating a model's usefulness (Ali, Khan, et al., 2020).

The selection of the correct examination variables for diagnosing heart disease provides many benefits, including faster diagnosis and lower cost of examination (Fajri et al., 2023). Based on the volume, variety, and velocity characteristics; feature selection methods face the following threefold challenges concerning data mining: feature selection techniques usually require large amounts of learning time, so it is hard for processing speed to catch up with the change of big data; generally, datasets not only include an immense amount of irrelevant and/or redundant features, but also have possible noises of different degrees and different types, which greatly increases the difficulty of selecting features; some data are unreliable/forged, due to different means of acquisition, or even loss, which further enhances the complexity of feature selection (Rong et al., 2019).

To calculate the cost factor, two methods can be used: 1. Cost-sensitive classification design and 2. Cost-sensitive feature selection method design. The first method is highly dependent on a classification algorithm like the decision tree, and the decision maker must perform all the classification steps. In the second method, the decision maker obtains a feature subset with the most predictive power and then can use any classifier(Ali, Khan, et al., 2020). Researchers in different fields have used various methods to choose the most informative groups of the primary features to build higher accuracy, speed, and lower-cost models. In many review articles, these methods have been reviewed based on different perspectives, including Evolutionary Computation (Xue et al., 2015), clustering approaches (Hancer et al., 2020), semi-supervised feature selection methods (Sheikhpour et al., 2017), ensemble feature selection methods (Bolón-Canedo & Alonso-Betanzos, 2019), swarm intelligence approaches (Nguyen et al., 2020), and similar approaches. It is discussed that each contains a collection of articles related to that perspective being considered at different times.

Real-world applications often involve significant costs in data acquisition, including time, financial and computational resources. Most existing feature selection methods overlook the associated costs (Mohanrasu et al., 2025). Researchers (Liang, 2024) unlike most of the existing methods simply adding or deleting features one by one, the proposed method uses an adaptive swarm intelligence algorithm to search the optimal subset. This algorithm achieves a more reasonable balance between the exploration and exploitation utilizing a cosine congestion factor, and is employed in cost-sensitive feature selection problem. The application of cost-based feature selection in various fields, including medicine (Fajri et al., 2023), image processing (Smits & Annoni, 2000b), phishing (Zangooei et al., 2019), computer networks (Tahir, 2016), and other fields, show the positive effect of these methods on increasing the speed and accuracy of machine learning methods in various fields.

This paper for the first time provides a comprehensive review of the research presented in the cost-based feature selection. Search techniques, evaluation criteria, number of objective functions, and data used in cost-based feature selection methods have been investigated. The widely used areas of cost-based feature selection and the datasets used in these applications are discussed.

In many real-world applications (e.g., healthcare, IoT, finance), the cost of acquiring features can drastically impact decision-making systems. However, most traditional FS methods ignore this factor. Although various cost-sensitive FS methods exist, there is no comprehensive taxonomy that critically compares their effectiveness in different cost settings or real-time environments. We argue that the current CBFS methods often fail to model practical, domain-specific costs, and that future research must explicitly consider interpretability, adaptability, and deployment constraints. This paper categorizes CBFS algorithms from a multi-dimensional perspective including evaluation strategy, cost modeling depth, and scalability.

The rest of the article is presented in the following order. Part 2 will discuss the review and definition of cost-based learning. In the third part, we will review the works on cost-based feature selection in terms of search techniques, evaluation criteria, the number of objectives, and data used. In the fourth part, we will introduce the areas where cost-based feature selection has been used. We introduce the datasets used in works done in the fifth section. At last, in section six, summarization, conclusions, and suggestions for the future have been discussed.

2. **Cost-based Learning**

As a branch of machine learning, cost-sensitive learning focuses on cost-related problems. In machine learning, two essential types of cost can be considered: feature cost and misclassification cost (Ali, Khan, et al., 2020). Misclassification costs assume unequal costs for different types of errors, and the goal is to minimize prediction costs. Minimizing the cost of misclassification has also attracted the attention of many machine learning researchers in past years (De Bock et al., 2020; Ji & Carin, 2007; Lu et al., 2019; Pendharkar, 2006; Xiong & Zuo, 2017).

Choosing suitable features from the primary feature set is an important task in machine learning. Researchers have proposed various methods to solve this challenge. However, most of them consider an equal cost for all features. This assumption may be wrong if achieving the attribute values is expensive. For example, the diagnostic value obtained in laboratory tests is costly in medical diagnostics. Feature costs have been less studied compared to misclassification costs. In these cases, the best solution is to select a model that provides acceptable classification performance, but at a much lower cost (Teisseyre Pawełand Zufferey & Słomka, 2019).

Feature cost can be interpreted as the cost of acquiring a feature set. This cost might include various factors like difficulty, time, and money. However, the cost of different features may vary in many real-world applications, and this difference can be significant to affect the feature selection outcome (Zhou et al., 2016).

We can summarize a basic feature-cost-sensitive system as follows:

$$S = (N,\ F, L, V_a \in F \bigcup L,\ I_a\ |a \in F \bigcup L,\ C) \tag{1}$$

where N refers to the test set of data instances, F and L indicate the features and class variables, respectively. V_a shows the possible values for every $a \in F \bigcup L$ that the information function for every a is shown by $I_a :\ N \rightarrow V_a$. Finally, the feature cost function is presented by $C :\ F \rightarrow R^{+} \bigcup 0$. Let's consider the data of some patients $N = \{n_1, n_2, ..., n_r\}$ where each data sample contains some features such as gender, age, and gene type $F = \{gender,\ age,\ and\ gene\ type\}$. If we consider the decision variable as $F = \{sick,\ healthy\}$ and C is empty; Then, all the features are assumed to have equal costs (Zhou et al., 2016).

3. Existing research on Cost-based Feature Selection

In building a learning model, the purpose of feature selection is to choose the most informative, relevant, and practical group of features from a large group of redundant and possibly irrelevant features. Instead of increasing the model's accuracy, the presence of irrelevant features also affects the accuracy. Most existing methods and approaches for feature selection consider an equal cost for all features. However, each feature can have a different cost in many real-world applications. Ignoring the cost of a feature makes it a purely theoretical approach to create a good subset of features, but in practice, we cannot use them due to the cost imposed (Zhou et al., 2016).

There were fewer studies on the cost of feature selection compared to the cost of classification. However, today these methods are attracting the attention of many researchers. With a focus on the second category of cost-sensitive learning methods (cost-sensitive feature selection), This paper for the first time provides a comprehensive review of the research presented in the cost-based feature selection. search techniques, evaluation criteria, number of objective functions, and data used in cost-based feature selection methods have been investigated. The widely used areas of cost-based feature selection and the datasets used in these applications are discussed.

The articles in this area use different terminology and titles to express the cost-based feature selection. To cover as many relevant articles as possible, we have tried to include all terms equivalent to cost-based feature selection in this study. For this purpose, in the initial phase, we created a CSV file of titles, abstracts, and keywords related to 25 selected articles by preliminary review of several articles related to different years in this field. In the next step, using the

Power BI tool and the collection of words in the CSV file, some pre-processing works such as removing stop words, converting words into words with lowercase letters, and removing punctuation marks and others have been done.

Power Query, part of the Microsoft Power BI suite, is a tool that automates the process of getting data into Excel and will save you hours of dull, repetitive, and error-prone work. Power Query makes it easy to extract data from many different data sources, filter that data, aggregate it, clean it, and perform calculations on it, finally loading that data into either your worksheet or directly into the new Excel 2013 Data Model used by Power Pivot. (Machiraju & Gaurav, 2018). As a result, a clean textual dataset of frequent terms is provided to make search terms. Figure 2 shows the word cloud of the 40 most frequent words.

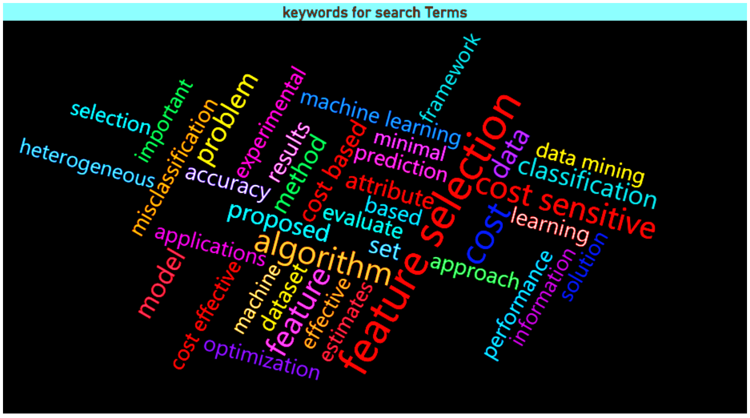

FIGURE 2. The word cloud of the 40 most frequent words.

The words that are consistent with our purpose in this research are highlighted in red. As is clear from the word cloud, cost-based feature selection is mentioned with a variety of synonyms in the reviewed articles, and accordingly; to create keywords, sentences were made from the possible combinations of these words. The following phrases are used in the search phase:

"Cost-sensitive feature selection"—— "Cost-based feature selection" ——"Cost-effective feature selection"——" Cost-based attribute selection"——" Cost-sensitive attribute selection"——" Cost-effective selection"

In the first step of the search, nearly 1506 articles were found in all periods based on the mentioned search phrases. After reviewing the found articles, irrelevant articles (articles that do not cover the selected topic, articles that were referenced only by the search term, articles that were not fully available, or articles written in the local language) are filtered. A total of 129 articles were selected to answer the questions of this study. The number of articles retrieved by year is shown in Figure 3.

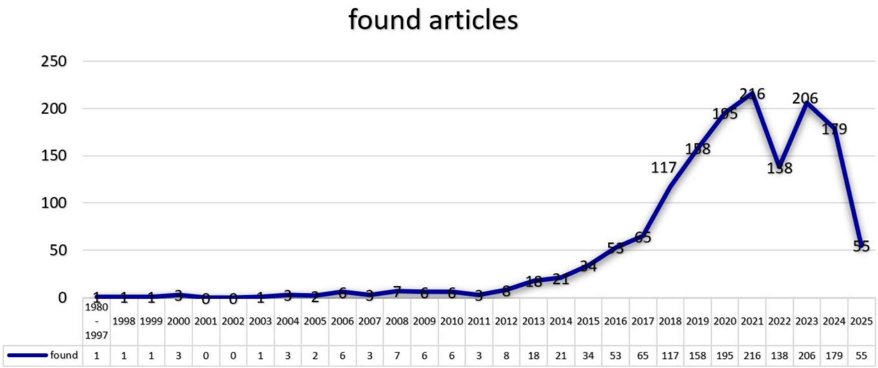

FIGURE 3. Retrieved Article Statistics by Year.

As is evident in Figure 3 graph, the concentration of researchers in this field has been growing in recent years. Figure 4 compares the number of articles found by title searched and the number of articles selected for review based on publication year.

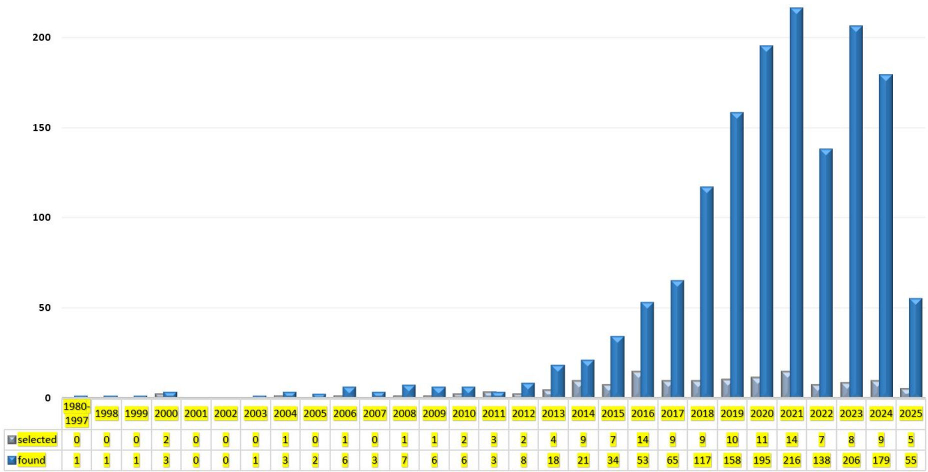

FIGURE 4. Comparing the number of found articles with the number of selected articles.

Feature selection is a challenging and complex issue, not only because we are supposed to select a small subset of the best features from a large data space but also based on the interactions and effects of the features on each other. A single feature may not be significantly related to the target class, but if it is considered together with other features, it will increase the accuracy of the machine learning model. On the other hand, there may be features that are known to be highly related to the target class alone, but when they are considered together with other features, not only they are not effective in increasing the accuracy of the model, but they may even decrease the accuracy of the learning model and be known as a redundant feature. Therefore, removing or selecting such features may be an obstacle to reaching an optimal subset of the set of primary features.

The traditional feature selection methods that go through features individually to arrive at the best subset do not help us achieve our goal well. The selected feature subset should be assessed together as a whole. Thus, two main factors, including evaluation criteria and search techniques, should be involved in the feature selection process.

The evaluation criteria measure and evaluate the feature subset's quality that guides the search, and the search techniques investigate the search space to reach the optimal subset. Feature selection performs with two goals: enhancing the performance of machine learning models and reducing the number of features to speed up learning and prediction in various domains. Often these two goals conflict with each other. Therefore, we can consider the feature selection process as a multi-objective optimization problem where a set of trade-off features should be found considering these two objectives (Zhou et al., 2016).

The step-by-step method used for this study is illustrated in the Figure 5.

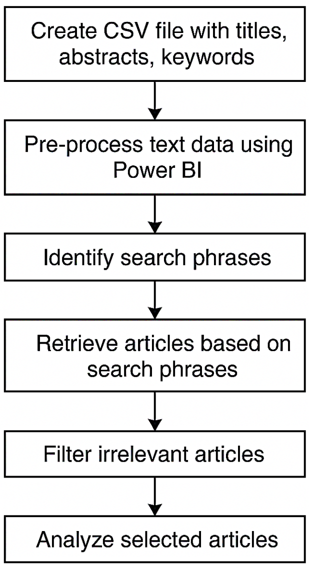

Figure 5. The step-by-step method used for this study.

Therefore, in this study, we attempted to consider the cost of the two main factors mentioned in the feature selection approaches and the number of targets these methods pursue in feature selection in reviewing the papers presented on the problem of cost-based feature selection. For each area we want to explore, we should assess and analyze our level of accuracy using data specific to that topic under the same conditions. The data used in the article is diverse and under different conditions, as this article intends to discuss how to select cost-based features used in various real-world applications. Therefore, when considering the work done in this area, the data should be considered based on the various features of other applications in addition to the previous three factors. The results of these investigations are given in the following section.

3.1. **Search Techniques.** Determining an appropriate search technique is very important for feature selection methods. Several search algorithms based on different techniques have been presented and used to find the best feature subset from the main feature set. In this research, three main categories of meta-heuristic, innovative, and comprehensive, were considered to separate the search methods used in cost-based feature selection. Meta-heuristic search techniques such as Ant colony optimization and genetic algorithm, etc. in the category of meta-heuristic techniques, forward, backward, and greedy feature selection algorithms in the category of heuristic techniques and branch and bound algorithms, backtracking, and dynamic programming in the category of brute-force techniques, were categorized.

Figure 6 shows the usage of the three categories identified by the search technique among the articles reviewed. As is evident in the diagram, most of the search algorithms used in the reviewed research are in the category of heuristic methods. Figure 6 shows a statistical comparison of search methods, and Figure 7 shows the procedure of focusing on each category of search techniques. To date, the number of peer-reviewed articles has been measured over three periods: ...-2010, 2011-2015, and 2016-present. It can be inferred from this graph that the focus of the research investigated was on using the brute-force method. In recent years, meta-heuristic methods have been used more than brute-force methods in this classification. The use of the brute force method in the graph shows the downward growth of this category of methods in recent studies.

Researchers focus on each method category because of that category's positive and influential role in achieving the researcher's goals. Therefore, it can be concluded that the category of heuristic methods achieved its goals more effectively than other methods in the achievement of researchers in cost-based feature selection. The graph shows a decreasing trend in the use of brute force techniques in this area, but the difference is not so significant that it can be asserted that these techniques do not have a positive and definite effect on achieving goals. Given the data growth rate in recent years, the high computational complexity of brute-force methods can be an obstacle to researchers turning to these methods. Heuristic methods, because they search specific parts of the data, reach results faster. Metaheuristic algorithms can solve a problem with reasonable speed and accuracy by providing a general solution without knowing the problem. In any case, the classification of search methods by frequency of use is generally classified in the following order: heuristics, metaheuristics, and brute-force.

Table 1 presents the categories of peer-reviewed articles in terms of search algorithms. Machine learning methods inspired by nature have generally yielded promising results in feature selection. In the article (Ciupke, 2006), researchers used the ant colony optimization algorithm to select cost-sensitive features. Researchers (Cui et al., 2024) and (Y. Zhang et al., 2015) have used the particle swarm optimization algorithm. Other researchers (Botes et al., 2017) have used the ant colony optimization algorithm in their research.

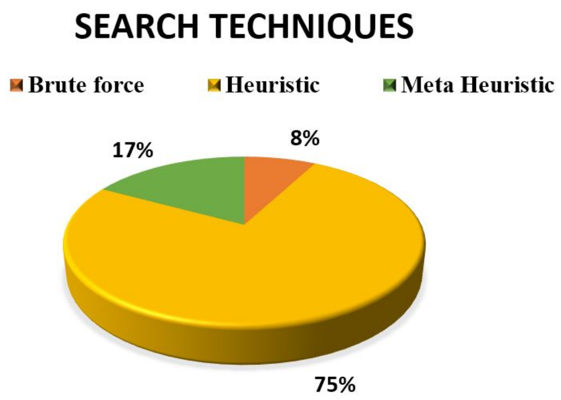

FIGURE 6. Statistics of methods based on search techniques.

Although heuristic methods are computationally efficient and easy to implement, they often rely on simplistic assumptions and lack flexibility in handling complex or multi-objective cost settings. In contrast, meta-heuristic approaches like PSO, ACO, and GA offer stronger global search capabilities and can effectively explore the trade-off between cost and accuracy. However, their performance is highly sensitive to parameter tuning and they incur higher computational overhead. Brute-force methods ensure optimality but are impractical for large datasets. Hybrid approaches attempt to combine the strengths of multiple techniques, but often at the expense of increased complexity and reduced interpretability. Thus, the choice of search strategy should consider the nature of the problem, data size, and available computational resources.

3.2. **Evaluation criteria.** Criteria for evaluating features are crucially important in constructing feature selection algorithms. Criteria define the details of

Table 1. Classification of reviewed articles based on the search technique.

Search Technique	Reference
Brute-force	(Liao et al., 2019), (Zhao et al., 2013a), (Zhao et al., 2013b), (Liao et al., 2014), (Zhao & Zhu, 2014), (Asharaf & Vijayan, 2015), (J. Li et al., 2015), (Zhao et al., 2016), (Fang et al., 2017), (S. Yu & Zhao, 2018)
Heuristic	(Smits & Annoni, 2000b), (Smits & Annoni, 2000a), (Mej-Lavalle, 2008), (Levering & Cutler, 2009), (Santos-Rodr& Garc-Garc, 2010), (D. Zhang & Shen, 2011), (Niu et al., 2014), (Bolón-Canedo, Remeseiro, et al., 2014), (Liu et al., 2014), (W. Qian et al., 2015), (Y. Zhang, Gong, et al., 2016), (Early et al., 2016), (Liu et al., 2017b), (Bach & Werner, 2018), (C.-W. Huang et al., 2018), (Y. Chen et al., 2018), (Lira et al., 2018), (Q. Huang et al., 2018), (An & Zhou, 2019), (J. Huang et al., 2019), (Das et al., 2020), (Das et al., 2021), (G. Qian et al., 2004), (Chang et al., 2012), (Joshua, 2013), (Bolón-Canedo, Porto-D, et al., 2014), (Bolón-Canedo et al., 2015), (X. Li et al., 2016), (J.-K. Li et al., 2016), (Vu et al., 2016), (Teisseyre Pawełand Zufferey & Słomka, 2019), (Zhou et al., 2016), (Tan et al., 2017), (Botes et al., 2017), (Maldonado et al., 2017), (Liu et al., 2017a), (Secerbegovic et al., 2018), (Kachuee et al., 2018), (le Roux et al., 2018), (Liao et al., 2018), (Zangooei et al., 2019), (Ben-Peña et al., 2019), (Zhao & Yu, 2019), (Jiang et al., 2019), (Lee et al., 2020), (Imran Ali et al., 2020), (Long et al., 2021), (Abdulla & Khasawneh, 2020), (Barushka & Hajek, 2020), (Chakraborty et al., 2021), (Jagdhuber, Lang,& Rahnenführer, 2020), (Raynal et al., 2023), (Javanmardi, 2011), (Pocock, 2012), (López, 2014), (Bolón-Canedo, 2014), (Porto D, 2015), (Tahir, 2016), (Saeedi, 2018), (Y. Li et al., 2022), (Momeni et al., 2021), (Teisseyre Pawełand Klonecki, 2021), (Jagdhuber & Rahnenführer, 2021), (Sun et al., 2021), (Yan et al., 2021), (BenPeña, 2021), (Gresser et al., 2021), (Gan et al., 2022), (Tao et al., 2021), (Bhuyan & Chakraborty, 2022), (Taylor et al., 2022), (Klonecki & Teisseyre, 2023), (McCombe et al., 2022), (M. Huang et al., 2022), (Mccombe et al., 2022), (Casella et al., 2022),(Raynal & Onnela, 2021) ,(Valancius et al., 2023) , (Knauer & Rodner, 2023), (Casella et al., 2023), (Yue et al., 2023), (Casella, 2023), (Janisch et al., 2024), (C. M. Chen et al., 2024), (J.-R. Yu et al., n.d.), (Yang et al., n.d.), (Ogawa et al., 2024), (Klonecki & Teisseyre Pawełand Lee, 2024), (Shi et al., n.d.), (C. M. Chen et al., 2024), (Seethalakshmi et al., 2024),(K. Huang et al., 2025), (Mohanrasu et al., 2025), (Al-Ahmari & Nadeem, 2025), (Z. Li et al., 2025), (Ahajjam et al., 2025)
Meta Heuristic	(Ciupke, 2006), (Ali, Khan, et al., 2020), (Y. Zhang et al., 2019), (Weiss et al., 2012), (Suryani et al., 2022), (Cui et al., 2024), (Zahirnia et al., 2015), (Akyon & Kalfaoglu, 2019), (Srimani & Koti, 2011), (Weiss et al., 2013), (Min et al., 2014), (Y. Zhang et al., 2015), (Bian et al., 2016), (Aydogan et al., 2016), (Niu et al., 2016), (Y. Zhang, Zhang, et al., 2016), (Min & Xu, 2016), (Y. Zhang et al., 2017), (Jagdhuber, Lang, Stenzl, et al., 2020), (Ali, Bilal, et al., 2020), (Feng et al., 2020), (Namakin et al., 2022), (Dharmalingam & Kumar, 2022), (Fajri et al., 2023)

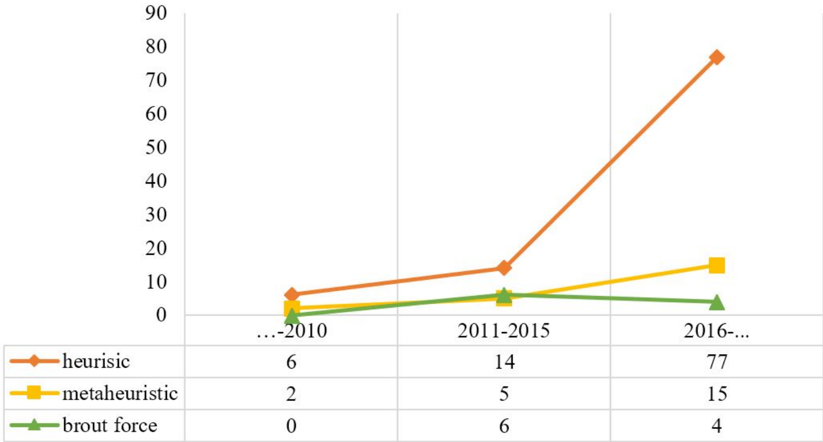

Figure 7. The growing trend of using search techniques.

feature selection evaluation. The criteria should be chosen according to the purpose of the feature selection. For example, the optimal set may be the small set that can give the best estimate based on the accuracy of the prediction. In general, feature selection aims to identify more critical, efficient, and cost-effective features in the data. The remaining features are known as redundant or unrelated features.

In the cost-based feature selection domain, as in other feature selection domains, researchers used a variety of evaluation criteria. Individual studies in cost-based feature selection were statistically analyzed based on the evaluation criteria. Based on the general classification of these methods, the articles under consideration were classified into four categories: wrappers, filters, embedded, and hybrids.

The name of the filter is obtained by filtering out unwanted features before training. Extensive research was needed to improve the efficiency of measurement accuracy. The filter method is also used independently of machine learning methods. This class of feature selection methods uses insights based on common features of data to evaluate the effectiveness of features. These methods work with high-dimensional data and provide a subset of available features that may be useful for some learning processes.

A filter feature selection model consists of two steps. In the first stage, feature selection is performed independently of the learning algorithm using criteria such as predictive power, distance, dependency, and stability. The second stage is training and testing. The second step is to perform a training and testing process to achieve test data prediction accuracy(Bayati et al., 2022; Hashemi, Joodaki, et al., 2022; Miri, Dowlatshahi, Hashemi, et al., 2022). The working procedure of the filter methods can be seen in Figure 8.

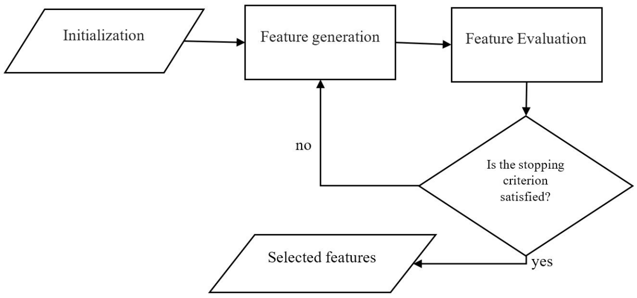

Figure 8. Working procedure of filter methods.

The simplest feature selection method can involve a learning algorithm as an evaluation method to decide whether to exclude or include a specific feature in a subset of useful features. The explicit goal of this type of method is to find the best subset of features with the highest possible accuracy. This category of feature selection methods is called wrapper methods. A wrapper model consists of two steps. The first step is to select a subset of the most valuable features based on the accuracy provided by the machine learning algorithm as a baseline. The second stage of these methods is similar to the second stage of filter methods. Since the first step only stores a subset of the best features, the second step uses this subset to test the accuracy of the learning algorithm. The first step is equivalent to the second step of reducing the dimension of the data (Hashemi et al., 2023; Hashemi, Dowlatshahi, & Nezamabadi-pour, 2021b; Hashemi, Dowlatshahi, et al., 2022; Hashemi, Pajoohan, et al., 2022) . Figure 9 shows the working routine of wrapper methods.

Filter and embedded methods introduced as feature selection methods are independently used to increase accuracy. The embedded approach is the third category of feature selection methods. These methods are similar to wrapper methods, which are based on the concept that the features selected by these methods are specifically selected for a particular learning algorithm. In addition, the features are selected during the learning process in these methods. Embedded methods that combine feature selection with used learning methods as part of the learning process can be more efficient for several reasons. They can take advantage of data availability in the learning process due to integration with feature selection and do not need to separate training data into two parts: training and evaluation data. Therefore, it can be said that this type of method can reach the most efficient subset of features faster because there is no predictor of retraining (Hashemi et al., 2020a; Hashemi, Dowlatshahi, & Nezamabadi-pour, 2021c, 2021a; Hashemi, Dowlatshahi, & Nezamabadi-Pour, 2021).

The three introductory categories of filters, wrappers, and embedded are the main categories of evaluation criteria presented in feature selection studies.

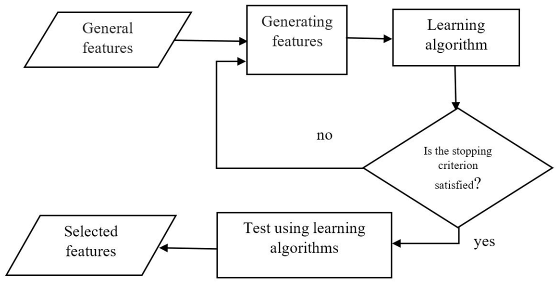

Figure 9. Working procedure of wrapper methods.

Since each category has its advantages and disadvantages, it is necessary to combine these methods to remove the disadvantages and use the advantages simultaneously. Therefore, a combination of these methods has been used in some studies as hybrid methods.

Hybrid methods represent the latest developments in feature selection. A hybrid method can be formed by combining two different methods, two methods of the same criterion, or two feature selection approaches. The hybrid method attempts to inherit the advantages of both methods by combining their complementary strengths (Ang et al., 2015). It uses different evaluation measures in different search stages to improve the efficiency and prediction performance with better computational performance. In Table 2 a comparison of filtering, wrapper, and embedded methods is presented from various aspects (Rong et al., 2019). Table 3 describes the advantages and disadvantages of each evaluation method. (Ang et al., 2015).

Table 2. Comparison of Commonly Used Feature Selection Methods (Rong et al., 2019).

Criteria	**Filter**	**Wrapper**	**Embedded**
Interact with classifiers	No	Yes	Yes
Computational cost	Comparatively low	Comparatively high	Depends
Accuracy	Comparatively low	Comparatively high	Comparatively high
Model feature dependence	Depends	Yes	Yes
Robustness	Yes	Yes	Yes
Risk of overfitting	No	Yes	Yes

Figure 10 shows the use of each of the four categories defined as evaluation criteria in the studies reviewed in this article. As can be seen in Figure 10, the majority of articles used the wrapper strategy to assess features. Filters and embedded methods are in the following order: Interestingly, the percentage of hybrid processes in this area is very low.

TABLE 3. Advantages and Disadvantages of feature evaluation(Ang et al., 2015).

Method	**Advantage**	**Disadvantage**
Filter	• Faster than wrapper • Scalable • Classifier independent • Better computational complexity than wrapper • Better generalizable property	• Ignores interaction between classifiers • Ignores dependency among features
Wrapper	• Interacts with classifier • Considers dependence among features • Higher performance accuracy than filter	• More prone to overfitting • Classifier specific • Requires expensive computation
Embedded	• Interacts with classifier • Better computational complexity than wrapper • Higher performance accuracy than filter • Less prone to overfitting than wrapper • Considers dependence among features	• Classifier specific
Hybrid	• Higher performance accuracy than filter • Better computational complexity than wrapper • Less prone to overfitting than wrapper	• Classifier specific

Figure 11 shows the process of using cost-based feature selection methods in each of the determined categories of evaluation criteria in different years. As can be seen in this graph, routines using various methods have been on the rise in recent years, indicating the willingness of researchers to apply them.

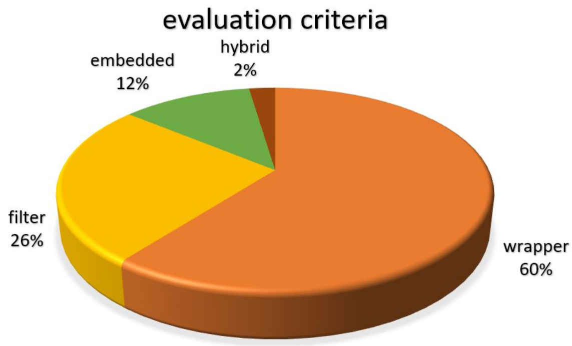

FIGURE 10. Statistics of articles based on evaluation criteria.

Among methods, the growth rate of using wrapper methods is higher than other methods. Filter, embedded, and hybrid techniques have been in the following positions over the last five years. The interesting thing to note is that the slope of the embedded methods graph is steeper than the filter methods, which can be concluded that although the number of filter methods is still more used, the tendency to use this category of methods has grown more compared to previous years. Hybrid methods are less used than other categories. In Table 4, the

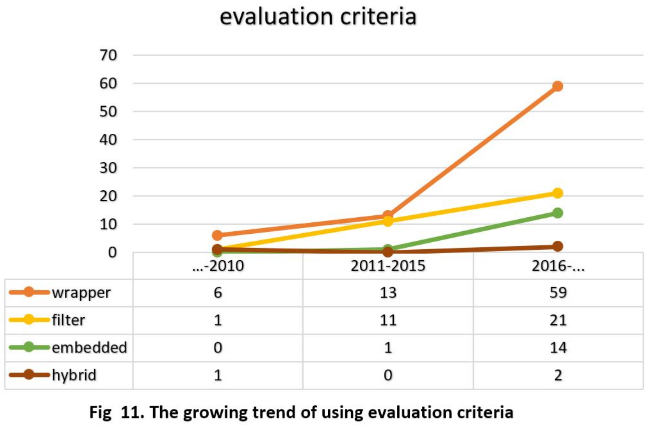

FIGURE 11. The growing trend of using evaluation criteria.

articles reviewed in this research have been categorized in terms of evaluation criteria.

Filter methods are appealing due to their independence from specific classifiers and low computational cost. However, their inability to model feature

Table 4. Classification of reviewed articles based on evaluation criteria.

Evaluation criteria	Reference
Wrapper	(Ciupke, 2006), (Ali, Khan, et al., 2020), (Fang et al., 2016), (Teisseyre Paweland Zufferey & Słomka, 2019), (Liao et al., 2019), (Zhao et al., 2013a) , (Zhao et al., 2013b), (Liao et al., 2014), (Zhao & Zhu, 2014), (Asharaf & Vijayan, 2015), (J. Li et al., 2015), (Zhao et al., 2016), (Y. Zhang, Gong, et al., 2016), (Smits & Annoni, 2000b), (Smits & Annoni, 2000a), (Mej-Lavalle, 2008), (Levering & Cutler, 2009), (Santos-Rodr& Garc-Garc, 2010), (Niu et al., 2014), (Akyon & Kalfaoglu, 2019), (W. Qian et al., 2015), (Early et al., 2016), (Liu et al., 2017b), (Q. Huang et al., 2018), (J. Huang et al., 2019), (Chang et al., 2012), (Joshua, 2013), (Botes et al., 2017), (Kachuee et al., 2018), (le Roux et al., 2018), (Liao et al., 2018), (Jiang et al., 2019), (Barushka & Hajek, 2020), (Chakraborty et al., 2021), (Jagdhuber, Lang, & Rahnenführer, 2020), (Saeedi, 2018), (Y. Li et al., 2022), (Y. Zhang et al., 2019), (Srimani & Koti, 2011), (Min et al., 2014), (Y. Zhang et al., 2015), (Aydogan et al., 2016), (Niu et al., 2016), (Y. Zhang, Zhang, et al., 2016), (Min & Xu, 2016), (Y. Zhang et al., 2017), (Jagdhuber, Lang, Stenzl, et al., 2020), (Feng et al., 2020), (Jagdhuber & Rahnenführer, 2021), (Gresser et al., 2021), (Cui et al., 2024), (BenPeña, 2021), (Gan et al., 2022) (Mccombe et al., 2022), (Bian et al., 2016) , (Bhuyan & Chakraborty, 2022), (Taylor et al., 2022), (M. Huang et al., 2022), (Weiss et al., 2012), (Suryani et al., 2022), (Fang et al., 2017), (Casella et al., 2022), (Dharmalingam & Kumar, 2022), (Valancius et al., 2023), (Knauer & Rodner, 2023), (Casella et al., 2023), (Yue et al., 2023), (Fajri et al., 2023), (Casella, 2023), (Janisch et al., 2024), (C. M. Chen et al., 2024), (J.-R. Yu et al., n.d.), (Yang et al., n.d.), (Ogawa et al., 2024), (Klonecki & Teisseyre Paweland Lee, 2024), (Shi et al., n.d.), (C. M. Chen et al., 2024), (Seethalakshmi et al., 2024), (K. Huang et al., 2025), (Mohanrasu et al., 2025), (Ahajjam et al., 2025)
Filter	(D. Zhang & Shen, 2011), (Bolón-Canedo, Remeseiro, et al., 2014), (Liu et al., 2014), (Zahirnia et al., 2015), (Bach & Werner, 2018), (Y. Chen et al., 2018), (Das et al., 2020), (Das et al., 2021), (G. Qian et al., 2004), (Weiss et al., 2013), (Bolón-Canedo, Porto-D, et al., 2014), (Bolón-Canedo et al., 2015), (X. Li et al., 2016), (J.-K. Li et al., 2016), (Vu et al., 2016), (Maldonado et al., 2017), (Liu et al., 2017a), (Zangooei et al., 2019), (Imran Ali et al., 2020), (Long et al., 2021), (Pocock, 2012), (López, 2014), (Bolón-Canedo, 2014), (Porto D, 2015), (Tahir, 2016), (Teisseyre Paweland Klonecki, 2021), (Sun et al., 2021), (Tao et al., 2021) , (Al-Ahmari & Nadeem, 2025), (Z. Li et al., 2025)
Embedded	(Zhou et al., 2016), (C.-W. Huang et al., 2018), (Lira et al., 2018), (McCombe et al., 2022), (An & Zhou, 2019), (Tan et al., 2017), (BenPeña et al., 2019), (Zhao & Yu, 2019), (Lee et al., 2020), (Abdulla & Khasawneh, 2020), (Javanmardi, 2011), (Momeni et al., 2021), (Yan et al., 2021), (Klonecki & Teisseyre, 2023)
Hybrid	(Namakin et al., 2022), (Ali, Bilal, et al., 2020), (Raynal et al., 2023)

interactions and cost-aware dynamics limits their effectiveness in real-world applications. Wrapper methods often yield better performance since they evaluate feature subsets using a learning algorithm, but they are resource-intensive and may suffer from overfitting. Embedded methods strike a balance between accuracy and efficiency but are highly model-dependent. Hybrid methods attempt to leverage the advantages of both filter and wrapper techniques, yet their design and validation can be nontrivial. Overall, the selection of an evaluation strategy must consider not only performance metrics but also practical deployment constraints like interpretability and speed.

3.3. **Number of objectives.** Identifying key features is necessary for formulating the process of selecting more efficient features as an optimization solution for machine learning algorithms. There are various criteria to determine the degree of influence of a feature on accuracy improvement, each evaluated using an objective function. The degree of importance of features can be determined by simultaneously optimizing several criteria and considering different aspects. At the same time, the criteria under consideration may contradict each other, further complicating the task. Therefore, multi-objective optimization strategies can be used to overcome this challenge. The multi-objective type arises when you make a trade-off between two opposing goals and decide to pick the best feature(Al-Tashi et al., 2020).

In this study, the reviewed articles are divided into three categories, Single-objectives, Multi-objectives, and Many-objectives, depending on the number of objectives set for optimization. Figure 12 shows the number of articles published in each category. As can be seen from this graph, 76% of the selected articles are in the Single-objectives category, and the rest are in the multi-objectives category. However, none of the peer-reviewed papers fall into the Many-objectives category.

The Many-objectives category is intended to distinguish between papers that consider more than one target or criterion when determining the effect of a function on method optimization. The lack of research work in this category can lead to two different results. First, good results can be expected from using this approach because this type of method has never been used, and second, it may have been used, but it did not lead to the submission of a research paper because it did not yield acceptable results.

More use of single-objective methods compared to multi-objective methods can indicate that these methods are more effective than multi-objective methods. However, some research areas are multi-objective in nature, and research work, including feature selection in these areas, requires the consideration of more than one criterion.

Figure 13 shows the procedure of applying the three mentioned categories in cost-based feature selection, determined based on the number of objectives in studies reviewed in different years. The graph shows that during the last five years, the rate of using single-objective and multi-objective methods has

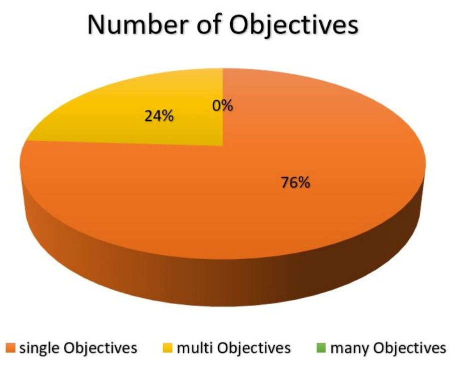

FIGURE 12. Categorization of papers based on the number of used objectives.

grown upward, and there is still more tendency towards using single-objective methods than multi-objective methods.

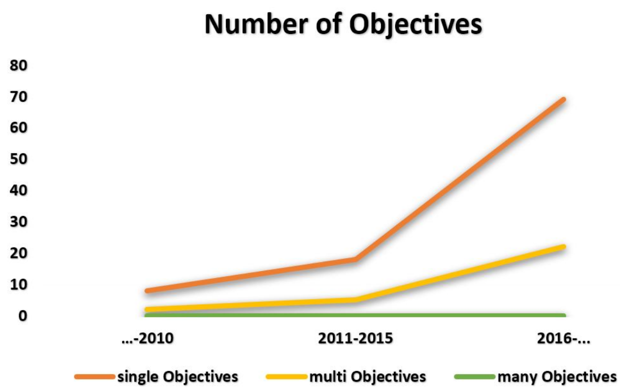

FIGURE 13. Growth routine using the number of considered objectives.

In Table 5. Classification of reviewed articles based on the number of objectives, the reviewed researches are separated based on the three categories in this section.

Single-objective methods, while simpler and more interpretable, often fail to reflect the real-world trade-offs between performance and cost. Multi-objective approaches provide a more realistic framework by optimizing for both accuracy and cost simultaneously. However, they introduce complexity in both algorithm design and solution interpretation. Pareto-optimal solutions are theoretically

Table 5. Classification of reviewed articles based on the number of objectives.

Number of objectives	Reference
Single-objectives	(Ciupke, 2006), (Ali, Khan, et al., 2020), (Teisseyre Pawełand Zuffery & Słomka, 2019), (Liao et al., 2019) , (Zhao et al., 2013a) , (Zhao et al., 2013b), (Liao et al., 2014), (Zhao & Zhu, 2014), (Asharaf & Vijayan, 2015), (J. Li et al., 2015), (Smits & Annoni, 2000a), (Mej-Lavalle, 2008), (Levering & Cutler, 2009), (Santos-Rodr& Garc-Garc, 2010), (Niu et al., 2014), (Akyon & Kalfaoglu, 2019), (Early et al., 2016), (Liu et al., 2017b), (Q. Huang et al., 2018), (J. Huang et al., 2019), (Chang et al., 2012), (Joshua, 2013), (Botes et al., 2017), [61], (Kachuee et al., 2018), (le Roux et al., 2018), (Liao et al., 2018), (Jiang et al., 2019), (Chakraborty et al., 2021), (Jagdhuber, Lang, & Rahnenführer, 2020), (Saeedi, 2018), (Y. Li et al., 2022), (Srimani & Koti, 2011), (Min et al., 2014), (Bian et al., 2016), (Aydogan et al., 2016), (Niu et al., 2016), (Jagdhuber, Lang, Stenzl, et al., 2020), (Feng et al., 2020), (Jagdhuber & Rahnenführer, 2021), (Gresser et al., 2021), (BenPeña, 2021), (Gan et al., 2022), (Mccombe et al., 2022), (Taylor et al., 2022), (M. Huang et al., 2022), (Weiss et al., 2012), (Suryani et al., 2022), (D. Zhang & Shen, 2011), (Liu et al., 2014), (Zahirnia et al., 2015), (Bach & Werner, 2018), (Y. Chen et al., 2018), (Das et al., 2020), (Das et al., 2021), (G. Qian et al., 2004), (Weiss et al., 2013), (Bolón-Canedo, Porto-D, et al., 2014), (Vu et al., 2016), (Maldonado et al., 2017), (Liu et al., 2017a), (Zangooei et al., 2019), (Imran Ali et al., 2020), (Raynal et al., 2023), (Porto D, 2015), (Tahir, 2016), (Teisseyre Pawełand Klonecki, 2021), (Sun et al., 2021), (Tao et al., 2021), (Zhou et al., 2016), (C.-W. Huang et al., 2018), (Lira et al., 2018), (S. Yu & Zhao, 2018), (McCombe et al., 2022), (An & Zhou, 2019), (Tan et al., 2017), (Ben-Peña et al., 2019), (Zhao & Yu, 2019), (Lee et al., 2020), (Abdulla & Khasawneh, 2020), (Javanmardi, 2011), (Momeni et al., 2021), (Klonecki & Teisseyre, 2023), (Smits & Annoni, 2000b), (Casella et al., 2022), (Raynal & Onnela, 2021), (Valancius et al., 2023), (Knauer & Rodner, 2023), (Casella et al., 2023), (Casella, 2023), (C. M. Chen et al., 2024), (J.-R. Yu et al., n.d.), (Ogawa et al., 2024), (Klonecki & Teisseyre Pawełand Lee, 2024), (C. M. Chen et al., 2024), (K. Huang et al., 2025), (Z. Li et al., 2025), (Ahajjam et al., 2025)
Multi-objectives	(Zhao et al., 2016), (Fang et al., 2017), (Bolón-Canedo, Remeseiro, et al., 2014), (W. Qian et al., 2015), (Y. Zhang et al., 2015), (Fang et al., 2016), (Bolón-Canedo et al., 2015), (X. Li et al., 2016), (J.-K. Li et al., 2016), (Y. Zhang, Zhang, et al., 2016), (Min & Xu, 2016), (Y. Zhang et al., 2017) , (Y. Zhang et al., 2019), (Long et al., 2021), (Yan et al., 2021), (Bhuyan & Chakraborty, 2022), (Ali, Bilal, et al., 2020), (Barushka & Hajek, 2020), (Pocock, 2012), (López, 2014), (Bolón-Canedo, 2014), (Cui et al., 2024), (Y. Zhang, Gong, et al., 2016), (Yue et al., 2023), (Fajri et al., 2023), (Janisch et al., 2024), (Yang et al., n.d.), (Shi et al., n.d.), (C. M. Chen et al., 2024), (Mohanrasu et al., 2025), (Al-Ahmari & Nadeem, 2025)
Many-objectives	-

valuable, but selecting the final subset from the Pareto front often requires additional criteria or domain knowledge. Despite these challenges, multi-objective optimization represents a more holistic approach and is essential for applications with strict cost constraints.

3.4. **Data.** To evaluate a machine learning method, we need to use data. The method is first supposed to extract a pattern from the training data and then use the test data to measure the accuracy of the learning method. This data may be the result of gathering information from previous samples and may be used offline or online. In this study, the data used in the research study was first divided into offline and online. However, as shown in Table 6, none of the reviewed articles used online data.

Supervised, unsupervised, and semi-supervised are known as three categories of feature selection methods according to class label information. Supervised methods use labeled data to measure the predictive power of features considering class labels. Unsupervised feature selection methods evaluate the relationship between the features that preserve specific properties of the data, such as those that preserve locality or variance. Due to the use of labeled information, supervised feature selection methods usually achieve better learning performance than unsupervised feature selection methods. However, sufficient labeled data is required for supervised feature selection monitoring methods, which are expensive and require extensive expertise.

There are many unlabeled and small labeled data in many real-world applications. Semi-supervised feature selection methods have been proposed to deal with the problem of small labeled data that use labeled and unlabeled data for feature selection. Semi-supervised feature selection methods use data distribution or local structure of labeled and unlabeled data and label information of labeled data to evaluate feature relevance (Dalvand et al., 2022; Hashemi et al., 2020b; Miri, Dowlatshahi, & Hashemi, 2022).

In this research, we decided to recategorize the selected articles with this view of the data. As it can be understood in Table 6, none of the data used in the reviewed research was identified as unsupervised or semi-supervised. Continuing our work, we considered types of observations according to the number of deterministic labels from two perspectives. As we can see in Figure 14, most of the studies reviewed were classified as single-label. The significant difference between using single-label types and multi-label types shows that more field researchers tend to choose cost-based features in this data category than in the others.

The sequence of data for the two categories identified in studies from different years is shown in Figure 15. The rapid growth in the use of supervised single-label methods in recent years shows a trend of cost-based feature selection researchers for this data category over the years. As shown in the figures, semi-supervised and unsupervised methods were not used in the reviewed articles.

Table 6. Classification of Studies by Supervision Type and Data Source

Data Type	Supervision Type	References
Online	Unsupervised	–
Online	Semi-supervised	–
Online	Supervised-single label	–
Online	Supervised-multi-label	–
Offline	Unsupervised	–
Offline	Semi-supervised	–
Offline	Supervised-single label	(Ciupke, 2006), (Ali, Khan, et al., 2020), (Liao et al., 2019),(Zhao et al., 2013a) , (Zhao et al., 2013b), (Liao et al., 2014), (Zhao & Zhu, 2014), (Asharaf & Vijayan, 2015), (J. Li et al., 2015), (Smits & Annoni, 2000a),(Mej-Lavalle, 2008), (Levering & Cutler, 2009), (Santos-Rodr& Garc-Garc, 2010), (Niu et al., 2014), (Akyon & Kalfaoglu, 2019), (Early et al., 2016),(Chang et al., 2012), (Joshua, 2013), (Botes et al., 2017), [61], (Kachuee et al., 2018), (le Roux et al., 2018), (Liao et al., 2018), (Jiang et al., 2019), (Chakraborty et al., 2021), (Jagdhuber, Lang, & Rahnenführer, 2020), (Saeedi, 2018), (Srimani & Koti, 2011), (Min et al., 2014), (Aydogan et al., 2016), (Niu et al., 2016), (Jagdhuber, Lang, Stenzl, et al., 2020), (Feng et al., 2020), (Jagdhuber & Rahnenführer, 2021), (Gresser et al., 2021), (BenPeña, 2021), (Gan et al., 2022), (Mccombe et al., 2022), (Weiss et al., 2012), (Suryani et al., 2022), (D. Zhang & Shen, 2011), (Liu et al., 2014), (Zahirnia et al., 2015), (Bach & Werner, 2018), (Y. Chen et al., 2018), (Das et al., 2020), (Das et al., 2021), (G. Qian et al., 2004), (Weiss et al., 2013), (Bolón-Canedo, Porto-D, et al., 2014), (Vu et al., 2016), (Maldonado et al., 2017), (Zangooei et al., 2019), (Imran Ali et al., 2020), (Raynal et al., 2023), (Porto D, 2015), (Tahir, 2016), (Teisseyre Pawełand Klonecki, 2021), (Sun et al., 2021), (Tao et al., 2021), (Zhou et al., 2016), (C.-W. Huang et al., 2018), (Lira et al., 2018), (S. Yu & Zhao, 2018), (McCombe et al., 2022), (An & Zhou, 2019), (Tan et al., 2017), (Ben-Peña et al., 2019), (Zhao & Yu, 2019), (Lee et al., 2020), (Abdulla & Khasawneh, 2020), (Klonecki & Teisseyre, 2023), (Ali, Bilal, et al., 2020), (Smits & Annoni, 2000b), (Zhao et al., 2016), (Fang et al., 2017), (W. Qian et al., 2015), (Y. Zhang et al., 2015), (Fang et al., 2016), (Bolón-Canedo et al., 2015), (X. Li et al., 2016), (J.-K. Li et al., 2016), (Y. Zhang, Zhang, et al., 2016), (Min & Xu, 2016), (Y. Zhang et al., 2017), (Y. Zhang et al., 2019), (Barushka & Hajek, 2020), (López, 2014), (Bolón-Canedo, 2014), (Cui et al., 2024), (Y. Zhang, Gong, et al., 2016), (Casella et al., 2022), (Raynal & Onnela, 2021), (Knauer & Rodner, 2023), (Casella et al., 2023), (Fajri et al., 2023), (Casella, 2023), (Janisch et al., 2024), (C. M. Chen et al., 2024), (J.-R. Yu et al., n.d.), (K. Huang et al., 2025), (Al-Ahmari & Nadeem, 2025), (Z. Li et al., 2025), (Ahajjam et al., 2025)
Offline	Supervised-multi-label	(Teisseyre Pawełand Zufferey & Słomka, 2019), (Javanmardi, 2011), (Long et al., 2021), (Yan et al., 2021), (Bhuyan & Chakraborty, 2022), (Liu et al., 2017b), (Q. Huang et al., 2018), (Bolón-Canedo, Remeseiro, et al., 2014), (J. Huang et al., 2019), (Liu et al., 2017a), (Momeni et al., 2021), (Y. Li et al., 2022), (Taylor et al., 2022), (M. Huang et al., 2022), (Pocock, 2012), (Namakin et al., 2022), (Dharmalingam & Kumar, 2022), (Valancius et al., 2023), (Yue et al., 2023), (Yang et al., n.d.), (Ogawa et al., 2024), (Klonecki & Teisseyre Pawełand Lee, 2024), (Shi et al., n.d.), (C. M. Chen et al., 2024), (Seethalakshmi et al., 2024), (Mohanrasu et al., 2025)

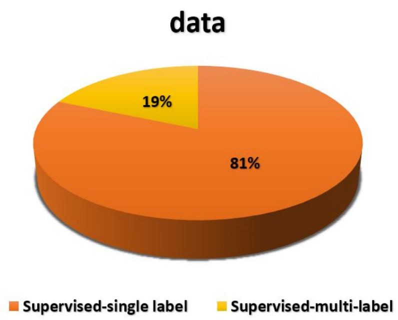

FIGURE 14. Statistics of used data.

This result could have occurred for various reasons. For example, cost-based feature selection methods did not provide suitable results in this type of data, or because due to the importance of the cost issue, these methods were not tested at all, which can be examined further. It should be noted that the authors do not claim to have reviewed all the research and some articles may not be available and have not been reviewed for various reasons.

Most current CBFS methods assume static or uniform cost values, which oversimplifies the cost structure encountered in real-world problems. A few methods incorporate instance-specific or dynamic costs, but such techniques are still underexplored. Moreover, many studies lack practical cost modeling (e.g., acquisition time, lab test cost, or energy consumption). As a result, the field would benefit from more robust and context-aware cost models that align better with operational realities and support adaptive behavior during model training or deployment.

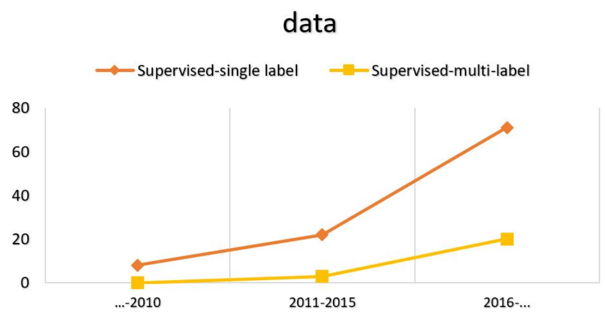

FIGURE 15. The growing trend of using Data.

4. **Applications**

Due to the positive impact of machine learning methods in increasing accuracy, these methods are used in various fields. In this article, we looked at many studies using the impact of cost-sensitive feature selection methods to improve accuracy in a variety of applications. To classify these studies, we looked at the number of works performed in various fields. For a more accurate diagnosis, medical staff must check various tests and parameters, and it is expensive to determine each parameter, so the diagnosis of various diseases is one of the widely used cost-based feature selection methods.

There are many articles reviewed in this study in this field. Most of these articles have used data from different fields of medicine, such as diabetes, heart disease, cancer, etc., in their experiments. For this reason, articles in this category have been considered the main category called medicine.

Image processing is one of the most used fields in machine learning methods, as many studies on different aspects of machine learning for working with data extracted from images have been presented in various fields. Of course, the cost-based feature selection method is no exception. The articles reviewed in this study introduce new methods in this area. For this reason, we identified images as the second category for applied research. In particular, this data model can also be used in the medical field, so studies that have used this type of data with data from other medical fields are placed in the same first category, namely the medical category.

Some studies are conducted to provide new methods and algorithms for cost-based feature selection. This group of studies generally considers the proposed methods in various areas. Therefore, we assigned a third research group to this model of articles studied in several fields.

In addition to these three categories of medicine, image, and multi-domain, some articles have used cost-based feature selection methods in other domains. Since the number of articles in each field is small, a category named OTHER was considered for this category of articles. Figure 16 shows the percentage of reviewed articles in each field. As it is known, most of the reviewed articles are in SEVERAL categories, which indicates that most of the works presented a general algorithm in the field of cost-based feature selection, and no specific application field was considered. The OTHER category is second in terms of article proportions. Figure 17 shows different application areas in this category. This diagram shows that DEFECT PROTECTION is the most frequent application area in this category.

Articles in this category are separated because the Image application category can contain articles from various fields that use experimental image data. Figure 18 shows the application areas of the articles in this category. The field of geography is the most used field for these studies. Some papers only used image data from the medical field, and in the main category, since the approach

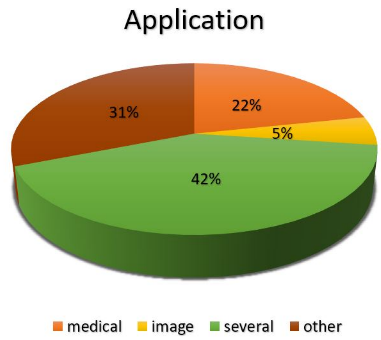

FIGURE 16. The percentage of reviewed articles in each field.

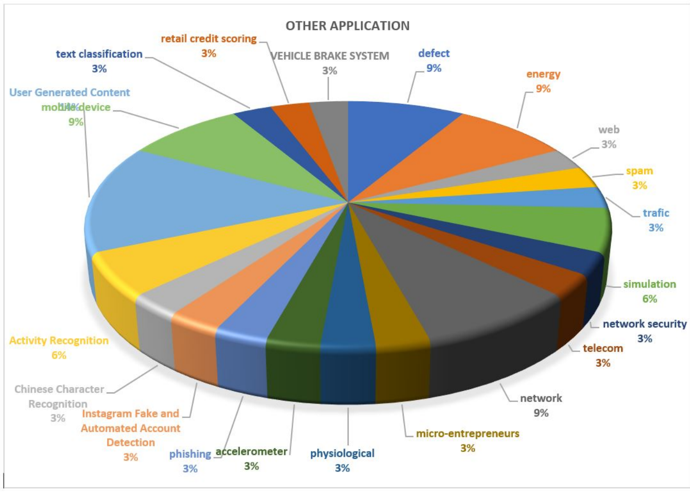

FIGURE 17. Different fields of application are placed in the OTHER category.

is based only on image data, we placed these studies in the image category and also considered a branch in the image category to the clinical images.

5. Benchmark Datasets

The data used by each application can be accessed from the various repositories that exist for it. Since the research reviewed in this article is about machine learning, we naturally used a repository of these fields to prepare data to evaluate the proposed methods. Most of the studies in the various studies in this area have used data available in the UCI repository. This repository contains many data for working in different areas of machine learning, so this was worth

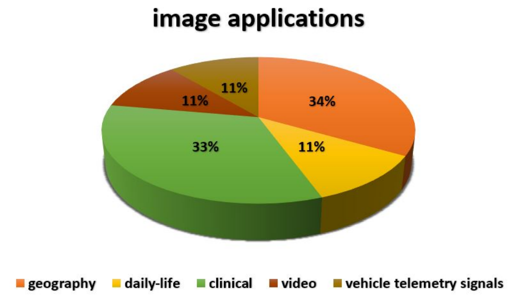

FIGURE 18. Different fields of application are placed in the OTHER category.

considering. Researchers created some data, and the rest used other repositories to access the data. Therefore, in this research, the articles were classified in terms of Benchmark Datasets into three categories: UCI Repository, Self-made Datasets, and Other Repositories. Figure 19 shows the percentage of articles placed in all three categories. Table 5 shows the data used in each research.

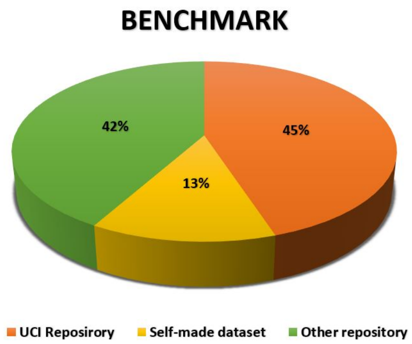

FIGURE 19. Statistics of used benchmark data in papers.

—p3cm—p8cm—

References	**Dataset**
(Smits & Annoni, 2000a)	CORINE land-cover/land-use database
(Smits & Annoni, 2000b)	CORINE land-cover/land-use database

(Ciupke, 2006)RotorKit model data (Mej-Lavalle, 2008) VLDB

(Levering & Cutler, 2009) KI-04 Dataset, Retail Store Dataset

(Weiss et al., 2012) self-made

(Santos-Rodr& Garc-Garc, 2010) UCI medical datasets (Pima Indian Diabetes, BUPA Liver Disorders, Heart Disease, Hepatitis Prognosis or Thyroid Disease)

(D. Zhang & Shen, 2011) ADNI dataset

(Niu et al., 2014) Mushroom, Tic-tac-toe, Voting and Zoo

(Bolón-Canedo, Remeseiro, et al., 2014) UCI (Letter, Magic04, Sat, Waveform, Hepatitis, Liver, Pima, Thyroid), broadinstitute (CNS, Colon, DLBCL, Leukemia)

(Liu et al., 2014) UCI (nasa)

(W. Qian et al., 2015) real UCI data sets (Adult, Annealing, Arrhythmia, Dermatology, Hepatitis, Stat Credit)

(Zahirnia et al., 2015) UCI (pima), Tabriz

(Early et al., 2016) RECS

(Y. Zhang, Gong, et al., 2016) Vowel, Ionosphere, Wisconsin Diagnosis Breast Cancer (WDBC) and Sonar

(Fang et al., 2016) UCI (Diab, Iris, Glass, Liver, Wdbc, Wine, Tic-tac-toe)

(Liu et al., 2017b) USPS and YaleB (multi-class), MSRC and TRECVID (multi-label)

(Bian et al., 2016) KDDCUP'99datasets

(C.-W. Huang et al., 2018) user0, user1, det124, det177

(Y. Chen et al., 2018) UCI (lung cancer)

(Lira et al., 2018) ACT-R

(Q. Huang et al., 2018) mulan library (Yeast, Emotions, Birds)

(Akyon & Kalfaoglu, 2019) instafake-dataset (generated)

(An & Zhou, 2019) UCI (hill, urban, arrhythmia, LSVT voice rehabilitation) and 11 genes (colon, SRBCT, Leukemia1, DLBCL, 9 Tumors, Brain Tumor1,

ALLAML, Brain Tumor2, Leukemia2, Lung Cancer, SMK-CAN-187)
(J. Huang et al., 2019) Mulan Library (Flags, Emotions, Birds, and Water-quality)
(Das et al., 2020) UCI (Parkinson's disease, Alzheimer's Disease, rare)
(Das et al., 2021) UCI (Parkinson's disease, Alzheimer's Disease, rare)
(G. Qian et al., 2004) Handwritten Chinese Character Recognition
(Srimani & Koti, 2011) UCI (pima)
(Chang et al., 2012) HRCT
(Zhao et al., 2013a) UCI (Liver, Credit, Iono, Diab)
(Zhao et al., 2013b) UCI (Liver, Wdbc, Wpbc, Diab, Iono, and Credit.)
(Joshua, 2013) UCI (nasa)
(Weiss et al., 2013) UCI (Pima, Bupa, Thyroid, Hepatitis, Breast, SPECT, Kr-Vs-Kp, Cars, Voting, Tic-Tac-Toc, Ecoli)
(Bolón-Canedo, Porto-D, et al., 2014) UCI (Hepatitis, Liver, Pima, Thyroid, Letter, Magic04, Optdigits, Pendigits, Sat, Segmentation, Waveform, Yeast, Brain, CNS, Colon, DLBCL, Leukemia)
(Liao et al., 2014) UCI (Tic-tac-toe, Voting, Zoo, Mushroom)
(Min et al., 2014) UCI (Tic-tac-toe, Voting, Zoo, Mushroom, Connect-4)
(Zhao & Zhu, 2014) UCI (Wisconsin Diagnostic Breast Cancer (Wdbc), Wisconsin Prognostic Breast Cancer (Wpbc), Diabetes (Diab), and Ionosphere (Iono))
(Asharaf & Vijayan, 2015) NASA Metrics Data Program (MDP) repository
(J. Li et al., 2015) UCI (Voting, tic-toc-toe)
(Y. Zhang et al., 2015) UCI (Wine, Vehicle, WDBC, Ionosphere, Sonar)
(Bolón-Canedo et al., 2015) VOPTICAL I1
(X. Li et al., 2016) UCI (House-votes-84, Kr-vs-kp, Mushroom, Promoters, Tic-tac-toe, Voting)
(Zhao et al., 2016) UCI (Liver, Wdbc, Wpbc, Diab, Breast, Promoters, Heart, Hepatitis, Sonar, Iono, Credit-a, Credit-g)

(Zhou et al., 2016)	UCI (Wine, Cancer, Heart, House, Ionosphere, Sonar, DLBCL, Leukemia, Colon, SRBCT)
(Aydogan et al., 2016)	Turken
(J.-K. Li et al., 2016)	UCI (Kr-vs-kp, Mushroom,Promoters, Tic-tac-toe, Voting)
(Niu et al., 2016)	UCI (Iris, Zoo, House-votes-84, Voting, Tic-tac-toe, Mushroom, Promoter, Kr-vs-kp)
(Vu et al., 2016)	self-made
(Y. Zhang, Zhang, et al., 2016)	UCI (Glass, Vehicle, WDBC, Ionosphere)
(Min & Xu, 2016)	UCI (Zoo, Tic-tac-toe, Voting, and Mushroom)
(Fang et al., 2017)	UCI (Diab, Iris, Glass, Liver, Wdbc, Wine, Tic-tac-toe)
(Y. Zhang et al., 2017)	UCI (including Vowel, Wine, Vehicle, Segmentation, WDBC, Ionosphere, Satellite, Sonar, LSVT, and CNAE-9)
(Tan et al., 2017)	UCI (Wine, Parkinsons, Ionosphere, Breast Cancer, Soner, Clean, Colon, Lymph)
(Botes et al., 2017)	NSL-KDD data set
(Maldonado et al., 2017)	Chilean bank
(Liu et al., 2017a)	Barcelona, MSVCv2 and TRECVID2005, LUNG, COIL20, Isolet1, USPS, YaleB, UMIST
(Kachuee et al., 2018)	UCI (HAPT, MNIST, Reuters R8, Yahoo TRC, Mushroom, Landsat, CTG, Synthesized, thyroid)
(le Roux et al., 2018)	sheep and rhinoceros
(Liao et al., 2018)	UCI (Diab, German, Heart, Image, Iono, Liver, Sonar, Wdbc, Wpbc)
(S. Yu & Zhao, 2018)	UCI (Liver, Wpbc, Promoters, Voting, Ionosphere, Credit, Prostate-GE, SMK-CAN-187, and Waveform)
(Zangooei et al., 2019)	Real-world data are collected from 5000 live English phishing and legitimate pages from November 2015 to January 2016.
(Teisseyre Pawełand Zufferey Słomka, 2019)	music, yeast, scene, flags, birds, media mill, cal500, nuswide, medical, genbase, bookmarks, bibtex, real dataset MIMIC-II and UCI (hepatit)

(Ben-Peña et al., 2019) UCI (wisconsin, votes, nursery, Australian, careval, leukemia, gastrointestinal)

(Y. Zhang et al., 2019) UCI (Zoo, WPBC, Ionosphere, Promoters, Sonar, Urban land cover (ULC), MUSK1, LSVT)

(Zhao & Yu, 2019) AMLALLML, LEUML, UCI (Biodeg, Clean1, Credit-a, Credit-g, DNA, EEG-EveState, German, Ionosphere, LEUML, Prostate-GE, Sonar, Spam, Wdbc, Wpbc)

(Liao et al., 2019) UCI (Bridges, CAD-diagnosis, Credit, Cylinder-bands, Diabetes, Diabetic-retinopathy, German, Heart, Hepatitis, Image, Ionosphere, Mice-protein, Sonar, Wdbc, Wine, Wpbc)

(Jiang et al., 2019) UCI (diabetes, heart-c, hepatitis (36 datasets of UCI in several applications)

(Lee et al., 2020) UCI (Chronic Kidney, Heart, Thyroid, and Breast Cancer, Colon, Leukemia, and Prostate, Synthetic 1, Synthetic 2, Synthetic 3)

(Jagdhuber, Lang, Stenzl, et al., 2020) self-made

(Imran Ali et al., 2020) CKD

(Long et al., 2021) Scene, Flags, Emotions, Gnegative, Plant, Birds, CAL500, Virus, Gpositive, Yeast

(Ali, Khan, et al., 2020) CKD

(Ali, Bilal, et al., 2020) CDK

(Abdulla & Khasawneh, 2020) Leukemia and DLBCL datasets

(Barushka & Hajek, 2020) Hyves, Twitter dataset

(Feng et al., 2020) (uciBreast-tissue, glass016vs5, newthyroid1 and shuttlec2vsc4, bupa, cleveland)
(Chakraborty et al., 2021) self-made

(Jagdhuber, Lang, Rahnenführer, 2020) Ada, Author, Qsar, Spam, Tokyo, Wdbc

(Raynal et al., 2023) self-made

(Javanmardi, 2011) PAN 2010 Wikipedia dataset

(Pocock, 2012) fbis, la12, ohscal, re0, re1

(López, 2014) VOPTICAL R dataset

(Bolón-Canedo, 2014) UCI (Hepatitis, Liver, Madelon, Magic04, Mushrooms, Mushrooms, Pima, Spambase, Splice)
(Porto D, 2015) UCI (Hepatitis, Liver, Pima, Thyroid, Letter, Magic04, Optdigits, Pendigits, Sat, Segmentation, Waveform, Yeast, Brain, CNS, Colon, DLBCL, Leukemia)
(Tahir, 2016) WARD, Torch, VIPeR, iLIDS
(Saeedi, 2018) UCI (Heterogeneity HAR, REALDISP HAR, Opportunity, Real World, Daily, and Sports Activities, 30-Movements-18-Activity)
(Y. Li et al., 2022) isolate, musk, Madelon, Inonosphere, Sonar, German, Hill-with and Hill-without
(Momeni et al., 2021) self-made
(Teisseyre Pawełand Klonecki, 2021) mimic ii+ self-made
(Jagdhuber & Rahnenführer, 2021) spambase+self-made
(Sun et al., 2021) HyperPlane, SEA, LED, Rotating spiral, Spam, Sensor, Electricity, Airlines
(Yan et al., 2021) Naval propulsion plants, Steel plate faults, Spam filter, Concrete strength, Remote sensing, Landsat, Thyroid disease, Vehicle silhouette, Bank Marketing, US censors income
(Cui et al., 2024) Australian, german, Thomas, hmeq, cashbus, lendingclub, pakdd2010, paipaidai, gmsc
(BenPeña, 2021) wisconsin, votes, nursery, Australian, careval, leukemia, leukemia
(Gresser et al., 2021) self-made
(Gan et al., 2022) Qsar, Madelon, Secom, Relathe, Pcmac, Basehock, German, Isolet, Chess, Hillwith, Hillwithout and Pima
(Tao et al., 2021) Madelon, SECOM, chess, isolate, Hill-with, Hill-without, musk, and sonar
(Mccombe et al., 2022) ADNI (Alzheimer's Disease Neuroimaging Initiative)
(Bhuyan & Chakraborty, 2022) Anneal, audiology, breast cancer, diabetes, glass, hepatitis, Ionosphere, iris, mushroom, sonar, vehicle, vowel
(Taylor et al., 2022) LED

(Klonecki & Teisseyre, 2023)	MIMIC (diabetes), MIMIC (hypertension), MIMIC (liver), Heart dataset, Thyroid dataset, Alzheimer dataset, Artificial dataset
(McCombe et al., 2022)	ADNI (Alzheimer's Disease Neuroimaging Initiative)
(M. Huang et al., 2022)	RCC
(Suryani et al., 2022)	Z-Alizadeh sani
(Namakin et al., 2022)	Glass, Breast Cancer, Heart, Wine, German, Ionosphere, Sonar, Hill-valley, Musk1, Arrhythmia, LSTV, Isolet5
(Casella et al., 2022)	self-made
(Raynal & Onnela, 2021)	self-made
(Dharmalingam & Kumar, 2022)	The data set comprises 2000 CT images.
(Valancius et al., 2023)	CUBE-
(Knauer & Rodner, 2023)	self-made
(Casella et al., 2023)	self-made
(Yue et al., 2023)	Bands, Hcvdat, Heart, Lung Cancer, Lymphography, Voting, Waveform
(Fajri et al., 2023)	Z-Alizadeh Sani, Cleveland, and Statlog
(Casella, 2023)	Honda Smart Home US
(Janisch et al., 2024)	Synthetic, Threatcrowd, Hepatitis, Mutagenesis, Ingredients, SAP, Stats
(C. M. Chen et al., 2024)	Online lending, offline lending
(J.-R. Yu et al., n.d.)	Heart disease, Heart failure, Housing, Ionosphere, Parkinson's, SPECT heart, Wisconsin, diagnostic
(Yang et al., n.d.)	Seeds, Algerian Forest Fires, Heart Failure Clinical Records, Pima, Raisin, Vowel, Sonar, Cardiotocography, Segment, Fetal Health Classification, MEU-Mobile KSD, Wine Quality, Electrical Grid Stability Simulated Data, Magic, Default of Credit Card Clients, Shuttle, Census, Susy
(Ogawa et al., 2024)	self-made
(Shi et al., n.d.)	self-made

(Seethalakshmi et al., 2024)	self-made
(K. Huang et al., 2025)	Overruling, AGNews, SciQ Hellasw,ag, Banking
(Mohanrasu et al., 2025)	Arts, Education, Entertainment, Genbase, GpositivePseAAC, Health, Recreation, Reference, Social, Yeast
(Al-Ahmari & Nadeem, 2025)	routine SSI surveillance data
(Z. Li et al., 2025)	Sonar2,Sonar4,Sonar8,Sonar16, Iono2,Iono4, Iono8, Iono16, ZOO, Wine2,Wine4, Wine8
(Ahajjam et al., 2025)	FIRMS

The majority of cost-based FS methods are evaluated on generic benchmark datasets, which may not reflect the challenges present in real domains such as healthcare, IoT, or edge computing. In domains like medicine, cost is not just numeric—it can represent delay, risk, or ethical constraints. Thus, methods need to be designed or adapted with domain-specific cost semantics in mind. Additionally, domain constraints may impose limitations on model complexity or explainability, which are rarely considered in existing CBFS frameworks. Future studies should focus more on real-world deployment.

Trend prediction is a critical aspect of understanding how research topics might evolve, allowing researchers and industry professionals to anticipate future directions. Topics and their associated qualifiers in machine learning research are Kernel Learning, Generalized Linear Models, Learning Theory, Nearest Neighbors and CBR, Regression Analysis, Regularization Methods, Reinforcement Learning (Devers Cantero, 2024).

6. Conclusions AND FUTURE WORK

This survey presents a comprehensive taxonomy and comparative analysis of cost-based feature selection (CBFS) methods. Beyond the descriptive aspect, our critical examination revealed several novel insights:

- Despite their popularity, many existing CBFS methods fail to capture real-world cost dynamics, often relying on synthetic or simplified assumptions.
- Meta-heuristic and hybrid strategies outperform traditional heuristics in multi-objective and high-dimensional scenarios, albeit at the expense of higher computational complexity.
- There is a lack of domain-adaptive and interpretable CBFS techniques that align well with application-specific cost constraints.

Contributions of this work include:

(1) A structured categorization of over 120 CBFS methods based on their search strategies, evaluation techniques, and application domains.

(2) A detailed analysis of each method's advantages, limitations, and computational trade-offs.
(3) Identification of underexplored directions such as online cost-aware learning, deep neural cost modeling, and interpretable selection frameworks.

Rather than merely describing prior work, this survey aims to offer critical assessments and act as a foundation for future research and innovation in cost-sensitive learning systems.

To further advance the field of cost-based feature selection, we outline key open issues and potential directions:

- **Online and Streaming CBFS:** Design algorithms capable of real-time, incremental feature selection under cost constraints.
- **Deep Learning Integration:** Explore cost-sensitive attention mechanisms and graph neural networks for automatic cost encoding.
- **Interpretable CBFS:** Develop transparent methods that balance cost efficiency with human interpretability, crucial for healthcare and finance.
- **Domain-Specific Modeling:** Incorporate realistic cost factors (e.g., medical test prices, sensor energy) into the selection process.
- **Benchmarking and Reproducibility:** Establish standardized datasets and metrics that reflect actual costs in real-world scenarios.
- **Multi-objective Optimization:** Advance CBFS methods that dynamically trade off between cost, accuracy, and complexity using evolutionary strategies.

7. Author Contributions

"Conceptualization, MB.Dowlatshahi and S.Beiranvand.; methodology, MB.Dowlatshahi and S.Beiranvand.; software, S.Beiranvand.; validation, MB.Dowlatshahi.; formal analysis, MB.Dowlatshahi.; investigation, S.Beiranvand.; resources, S.Beiranvand.; data curation, S.Beiranvand.; writing—original draft preparation, S.Beiranvand and A.Hashemi.; writing—review and editing, MB.Dowlatshahi and A.Hashemi.; visualization, S.Beiranvand.; supervision, MB.Dowlatshahi.; project administration, MB.Dowlatshahi. All authors have read and agreed to the published version of the manuscript."

8. Data Availability Statement

"Not applicable" here.

9. Aknowledgement

We would like to thank the reviewers for their thoughtful comments and efforts towards improving our manuscript.

10. Ethical considerations

The study was approved by the Ethics Committee of the University of ABCD (Ethical code: FR.AMU.REC.2022.500). The authors avoided from data fabrication and falsification.

11. Funding

The study has no fund.

12. Conflict of interest

The authors declare no conflict of interest.

References

[1] Abdulla, M., & Khasawneh, M. T. (2020). G-Forest: An ensemble method for cost-sensitive feature selection in gene expression microarrays. Artificial Intelligence in Medicine, 108, 101941.

[2] Ahajjam, A., Allgaier, M., Chance, R., Chukwuemeka, E., Putkonen, J., & Pasch, T. (2025). Enhancing prediction of wildfire occurrence and behavior in Alaska using spatio-temporal clustering and ensemble machine learning. Ecological Informatics, 85, 102963.

[3] Akyon, F. C.,& Kalfaoglu, M. E. (2019). Instagram fake and automated account detection. 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), 1–7.

[4] Al-Ahmari, S., & Nadeem, F. (2025). Improving Surgical Site Infection Prediction Using Machine Learning: Addressing Challenges of Highly Imbalanced Data. Diagnostics, 15(4), 501.

[5] Al-Tashi, Q., Abdulkadir, S. J., Rais, H. M., Mirjalili, S.,& Alhussian, H. (2020). Approaches to multi-objective feature selection: a systematic literature review. IEEE Access, 8, 125076–125096.

[6] Ali, S. I., Bilal, H. S. M., Hussain, M., Hussain, J., Satti, F. A., Hussain, M., Park, G. H., Chung, T., & Lee, S. (2020). Ensemble feature ranking for cost-based non-overlapping groups: A case study of chronic kidney disease diagnosis in developing countries. IEEE Access, 8, 215623–215648.

[7] Ali, S. I., Khan, W. A., Lee, S., & Lee, S.-H. (2020). Cost-Sensitive Feature Selection using Particle Swarm Optimization: A Post-Processing Approach. 2020 International Conference on Information Networking (ICOIN), 97–101.

[8] An, C., & Zhou, Q. (2019). A cost-sensitive feature selection method for high-dimensional data. 2019 14th International Conference on Computer Science & Education (ICCSE), 1089–1094.

[8] Ang, J. C., Mirzal, A., Haron, H., & Hamed, H. N. A. (2015). Supervised, unsupervised, and semi-supervised feature selection: a review on gene selection. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 13(5), 971–989.

[9] Asharaf, L. A., & Vijayan, V. (2015). Cost-Sensitive Boosting Networks for Data Defect Prediction. International Journal of Science and Research (IJSR), 4.

[10] Aydogan, E. K., Ozmen, M., & Delice, Y. (2016). Cost Sensitive Feature Selection in Decision-Theoretic Rough Set Models for Customer Churn Prediction: The Case of Telecommunication Sector Customers. International Journal of Economics and Management Engineering, 10(4), 1317–1321.

[11] Bach, M., & Werner, A. (2018). Cost-sensitive feature selection for class imbalance problem. Information Systems Architecture and Technology: Proceedings of 38th International Conference on Information Systems Architecture and Technology–ISAT 2017: Part I, 182–194.

[12] Barushka, A., & Hajek, P. (2020). Spam detection on social networks using cost-sensitive feature selection and ensemble-based regularized deep neural networks. Neural Computing and Applications, 32(9), 4239–4257.

[13] Bayati, H., Dowlatshahi, M. B., & Hashemi, A. (2022). MSSL: a memetic-based sparse subspace learning algorithm for multi-label classification. International Journal of Machine Learning and Cybernetics, 13(11), 3607–3624.

[14] Beiranvand, S., & Chahooki, Z. (2016). Bridging the semantic gap for software effort estimation by hierarchical feature selection techniques. Journal of AI and Data Mining, 4(2), 157–168.

[15] Beiranvand, S., & Zare Chahooki, M. A. (2023). Accuracy Improvement in Software Cost Estimation based on Selection of Relevant Features of Homogeneous Clusters. Journal of AI and Data Mining, 11(3), 453–476.

[16] Ben-Peña, S., Blanquero, R., Carrizosa, E., & Ram-Cobo, P. (2019). Cost-sensitive feature selection for support vector machines. Computers & Operations Research, 106, 169–178.

[17] BenPeña, S. (2021). New models and methods for classification and feature selection. a mathematical optimization perspective.

[18] Bhuyan, H. K., & Chakraborty, C. (2022). Explainable machine learning for data extraction across computational social system. IEEE Transactions on Computational Social Systems.

[19] Bian, J., Peng, X., Wang, Y., & Zhang, H. (2016). An Efficient Cost-Sensitive Feature Selection Using Chaos Genetic Algorithm for Class Imbalance Problem. Mathematical Problems in Engineering, 2016(1), 8752181.

[20] Bolón-Canedo, V. (2014). Novel feature selection methods for high dimensional data.

[21] Bolón-Canedo, V., & Alonso-Betanzos, A. (2019). Ensembles for feature selection: A review and future trends. Information Fusion, 52, 1–12.

[22] Bolón-Canedo, V., Porto-D, I., Sánchez-Maroño, N., & Alonso-Betanzos, A. (2014). A framework for cost-based feature selection. Pattern Recognition, 47(7), 2481–2489.

[23] Bolón-Canedo, V., Remeseiro, B., Sánchez-Marono, N., & Alonso-Betanzos, A. (2014). mC-ReliefF-An Extension of ReliefF for Cost-based Feature Selection. International Conference on Agents and Artificial Intelligence, 2, 42–51.

[24] Bolón-Canedo, V., Remeseiro, B., Sánchez-Maroño, N., & Alonso-Betanzos, A. (2015). Real-time tear film classification through cost-based feature selection. Transactions on Computational Collective Intelligence XX, 78–98.

[25] Botes, F. H., Leenen, L., & de La Harpe, R. (2017). Ant tree miner amyntas: Automatic, cost-based feature selection for intrusion detection. Journal of Information Warfare, 16(4), 73–92.

[26] Büyükkeçeci, M., & Okur, M. C. (2022). A comprehensive review of feature selection and feature selection stability in machine learning. Gazi University Journal of Science, 36(4), 1506–1520.

[27] Casella, E. (2023). Machine-Learning-Powered Cyber-Physical Systems.

[28] Casella, E., Cantor, M. C., Setser, M. M. W., Silvestri, S., & Costa, J. H. C. (2023). A machine learning and optimization framework for the early diagnosis of bovine respiratory disease. IEEE Access.

[29] Casella, E., Cantor, M. C., Silvestri, S., Renaud, D. L., & Costa, J. H. C. (2022). Cost-aware inference of bovine respiratory disease in calves using precision livestock technology. 2022 18th International Conference on Distributed Computing in Sensor Systems (DCOSS), 109–116.

[30] Chakraborty, B., Divakaran, D. M., Nevat, I., Peters, G. W., & Gurusamy, M. (2021). Cost-aware feature selection for IoT device classification. IEEE Internet of Things Journal, 8(14), 11052–11064.

[31] Chang, Y., Kim, N., Lee, Y., Lim, J., Seo, J. B., & Lee, Y. K. (2012). Fast and efficient lung disease classification using hierarchical one-against-all support vector machine and cost-sensitive feature selection. Computers in Biology and Medicine, 42(12), 1157–1164.

[32] Chen, C. M., Tso, G. K. F., & He, K. (2024). Quantum Optimized Cost Based Feature Selection and Credit Scoring for Mobile Micro-financing. Computational Economics, 63(2), 919–950.

[33] Chen, Y., Wang, Y., Cao, L., & Jin, Q. (2018). An effective feature selection scheme for healthcare data classification using binary particle swarm optimization. 2018 9th International Conference on Information Technology in Medicine and Education (ITME), 703–707.

[34] Ciupke, K. (2006). Cost-sensitive feature selection. Diagnostyka, 2 (38), 45–48.

[35] Cui, H., Zhang, L., Yang, H., Wang, J., & Liu, Z. (2024). Maximizing the lender's profit: profit-oriented loan default prediction based on a weighting model. Annals of Operations Research, 1–34.

[36] Dalvand, A., Dowlatshahi, M. B., & Hashemi, A. (2022). SGFS: a semi-supervised graph-based feature selection algorithm based on the PageRank algorithm. 2022 27th International Computer Conference, Computer Society of Iran (CSICC), 1–6.

[37] Das, S., Iyer, R., & Natarajan, S. (2020). Cost Aware Feature Elicitation. KiML 2020, 3.

[38] Das, S., Iyer, R., & Natarajan, S. (2021). A clustering based selection framework for cost aware and test-time feature elicitation. Proceedings of the 3rd ACM India Joint International Conference on Data Science & Management of Data (8th ACM IKDD CODS & 26th COMAD), 20–28.

[39] De Bock, K. W., Coussement, K., & Lessmann, S. (2020). Cost-sensitive business failure prediction when misclassification costs are uncertain: A heterogeneous ensemble selection approach. European Journal of Operational Research, 285(2), 612–630.

[40] Devers Cantero, J. (2024). Trend analysis in machine learning. Universitat Politècnica de Catalunya.

[41] Dharmalingam, V., & Kumar, D. (2022). Hybrid feature selection model for classification of lung disorders. Journal of Ambient Intelligence and Humanized Computing, 13(12), 5609–5625.

[42] Early, K., Fienberg, S. E., & Mankoff, J. (2016). Cost-Effective Feature Selection and Ordering for Personalized Energy Estimates. AAAI Workshop: AI for Smart Grids and Smart Buildings.

[43] Fajri, Y. A. Z. A., Suryani, E., & Setyawan, S. (2023). Use of Artificial Bee Swarm Optimization (ABSO) for Feature Selection in System Diagnosis for Coronary Heart Disease.

[44] Fang, Y., Liu, Z.-H., & Min, F. (2016). Multi-objective cost-sensitive attribute reduction on data with error ranges. International Journal of Machine Learning and Cybernetics, 7, 783–793.

[45] Fang, Y., Liu, Z.-H., & Min, F. (2017). A PSO algorithm for multi-objective cost-sensitive attribute reduction on numeric data with error ranges. Soft Computing, 21(23), 7173–7189.

[46] Feng, F., Li, K.-C., Shen, J., Zhou, Q., & Yang, X. (2020). Using cost-sensitive learning and feature selection algorithms to improve the performance of imbalanced classification. IEEE Access, 8, 69979–69996.

[47] Gan, J., Li, J., & Xie, Y. (2022). Robust SVM for cost-sensitive learning. Neural Processing Letters, 1–22.

[48] Gresser, M., Brehl, D., Rabe, M. D., & Bäker, B. (2021). Optimization of the Design of Diagnostic Systems through requirement-based Feature Selection. International Journal of Electronics and Electrical Engineering, 9(2), 26–35.

[49] Hancer, E., Xue, B., & Zhang, M. (2020). A survey on feature selection approaches for clustering. Artificial Intelligence Review, 53(6), 4519–4545.

[50] Hashemi, A., Dowlatshahi, M. B., & Nazamabadi-pour, H. (2023). Minimum redundancy maximum relevance ensemble feature selection: A bi-objective Pareto-based approach. Journal of Soft Computing and Information Technology (JSCIT) Vol, 12(1).

[51] Hashemi, A., Dowlatshahi, M. B., & Nezamabadi-pour, H. (2021a). A pareto-based ensemble of feature selection algorithms. Expert Systems with Applications, 180, 115130.

[52] Hashemi, A., Dowlatshahi, M. B., & Nezamabadi-pour, H. (2021b). An efficient Pareto-based feature selection algorithm for multi-label classification. Information Sciences, 581, 428–447.

[53] Hashemi, A., Dowlatshahi, M. B., & Nezamabadi-pour, H. (2021c). VMFS: A VIKOR-based multi-target feature selection. Expert Systems with Applications, 182, 115224.

[54] Hashemi, A., Dowlatshahi, M. B., & Nezamabadi-pour, H. (2022). Ensemble of feature selection algorithms: a multi-criteria decision-making approach. International Journal of Machine Learning and Cybernetics, 13(1), 49–69.

[55] Hashemi, A., Dowlatshahi, M. B., & Nezamabadi-Pour, H. (2020a). MFS-MCDM: Multi-label feature selection using multi-criteria decision making. Knowledge-Based Systems, 206, 106365.

[56] Hashemi, A., Dowlatshahi, M. B., & Nezamabadi-Pour, H. (2020b). MGFS: A multi-label graph-based feature selection algorithm via PageRank centrality. Expert Systems with Applications, 142, 113024.

[57] Hashemi, A., Dowlatshahi, M. B., & Nezamabadi-Pour, H. (2021). A bipartite matching-based feature selection for multi-label learning. International Journal of Machine Learning and Cybernetics, 12, 459–475.

[58] Hashemi, A., Joodaki, M., Joodaki, N. Z., & Dowlatshahi, M. B. (2022). Ant colony optimization equipped with an ensemble of heuristics through multi-criteria decision making: A case study in ensemble feature selection. Applied Soft Computing, 124, 109046.

[59] Hashemi, A., Pajoohan, M.-R., & Dowlatshahi, M. B. (2022). Online streaming feature selection based on Sugeno fuzzy integral. 2022 9th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), 1–6.

[60] Huang, C.-W., Chou, C.-K., & Chen, M.-S. (2018). A salient ensemble of trees using cascaded linear classifiers with feature-cost constraints. Proceedings of the 2018 SIAM International Conference on Data Mining, 486–494.

[61] Huang, J., Qian, W., Wu, B., & Wang, Y. (2019). Cost-Sensitive Feature Selection Based on Label Significance and Positive Region. 2019 International Conference on Machine Learning and Cybernetics (ICMLC), 1–7.

[62] Huang, K., Shi, Y., Ding, D., Li, Y., Fei, Y., Lakshmanan, L., Xiao, X. (2025). ThriftLLM: On Cost-Effective Selection of Large Language Models for Classification Queries. ArXiv Preprint ArXiv:2501.04901.

[63] Huang, M., Ye, X., Imakura, A., & Sakurai, T. (2022). Sequential reinforcement active feature learning for gene signature identification in renal cell carcinoma. Journal of Biomedical Informatics, 128, 104049.

[64] Huang, Q., Qian, W., Shu, W., Wu, B., & Feng, S. (2018). Multi-label cost-sensitive feature selection algorithm in incomplete data. 2018 International Conference on Machine Learning and Cybernetics (ICMLC), 1, 56–62.

[65] Imran Ali, S., Ali, B., Hussain, J., Hussain, M., Satti, F. A., Park, G. H., Lee, S. (2020). Cost-sensitive ensemble feature ranking and automatic threshold selection for chronic kidney disease diagnosis. Applied Sciences, 10(16), 5663.

[66] Jagdhuber, R., Lang, M., & Rahnenführer, J. (2020). Feature Selection Methods for Cost-Constrained Classification in Random Forests. ArXiv Preprint ArXiv:2008.06298.

[67] Jagdhuber, R., Lang, M., Stenzl, A., Neuhaus, J., & Rahnenführer, J. (2020). Cost-Constrained feature selection in binary classification: adaptations for greedy forward selection and genetic algorithms. BMC Bioinformatics, 21, 1–21.

[68] Jagdhuber, R., & Rahnenführer, J. (2021). Implications on Feature Detection When Using the Benefit–Cost Ratio. SN Computer Science, 2(4), 316.

[69] Janisch, J., Pevnỳ, T., & Lisỳ, V. (2024). Classification with costly features in hierarchical deep sets. Machine Learning, 1–36.

[70] Javanmardi, S. (2011). Measuring content quality in user generated content systems: a machine learning approach. University of California, Irvine.

[71] Ji, S., & Carin, L. (2007). Cost-sensitive feature acquisition and classification. Pattern Recognition, 40(5), 1474–1485.

[72] Jiang, L., Kong, G., & Li, C. (2019). Wrapper framework for test-cost-sensitive feature selection. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 51(3), 1747–1756.

[73] Joshua, A. (2013). Improving Software Quality Using Two Stage Cost Sensitive Learning. Proc. of International Conf. on Science and Research, 1726–1728.

[74] Kachuee, M., Darabi, S., Moatamed, B., & Sarrafzadeh, M. (2018). Dynamic feature acquisition using denoising autoencoders. IEEE Transactions on Neural Networks and Learning Systems, 30(8), 2252–2262.

[75] Klonecki, T., & Teisseyre, P. (2023). Feature selection under budget constraint in medical applications: analysis of penalized empirical risk minimization methods. Applied Intelligence, 53(24), 29943–29973.

[76] Klonecki, T., & Teisseyre Pawełand Lee, J. (2024). Cost-constrained multi-label group feature selection using shadow features. ArXiv Preprint ArXiv:2408.01851.

[77] Knauer, R., & Rodner, E. (2023). Cost-Sensitive Best Subset Selection for Logistic Regression: A Mixed-Integer Conic Optimization Perspective. German Conference on Artificial Intelligence (Künstliche Intelligenz), 114–129.

[78] le Roux, S. P., Wolhuter, R., & Niesler, T. (2018). Energy-aware feature and model selection for onboard behavior classification in low-power animal borne sensor applications. IEEE Sensors Journal, 19(7), 2722–2734.

[79] Lee, I. G., Zhang, Q., Yoon, S. W., & Won, D. (2020). A mixed integer linear programming support vector machine for cost-effective feature selection. Knowledge-Based Systems, 203, 106145.

[80] Levering, R., & Cutler, M. (2009). Cost-sensitive feature extraction and selection in genre classification. Journal for Language Technology and Computational Linguistics, 24(1), 57–72.

[81] Li, J.-K., Zhao, H., & Zhu, W. (2016). Feature Selection with Multi-Cost Constraint. , 17(5), 981–991.

[82] Li, J., Zhao, H., & Zhu, W. (2015). Fast randomized algorithm with restart strategy for minimal test cost feature selection. International Journal of Machine Learning and Cybernetics, 6, 435–442.

[83] Li, X., Zhao, H., & Zhu, W. (2016). An exponent weighted algorithm for minimal cost feature selection. International Journal of Machine Learning and Cybernetics, 7, 689–698.

[84] Li, Y., Ma, C., Tao, Y., Hu, Z., Su, Z., & Liu, M. (2022). A robust cost-sensitive feature selection via self-paced learning regularization. Neural Processing Letters, 54(4), 2571–2588.

[85] Li, Z., Cheng, P., Yin, L., & Guan, Y. (2025). A Recursive Attribute Reduction Algorithm and Its Application in Predicting the Hot Metal Silicon Content in Blast Furnaces. Big Data and Cognitive Computing, 9(1), 6.

[86] Liang, Y. (2024). Cost-Sensitive Feature Selection Based on Adaptive Hunting Optimization. 2024 4th International Conference on Computer Communication and Artificial Intelligence (CCAI), 546–551.

[87] Liao, S., Zhu, Q., & Min, F. (2014). Cost-Sensitive Attribute Reduction in Decision-Theoretic Rough Set Models. Mathematical Problems in Engineering, 2014(1), 875918.

[88] Liao, S., Zhu, Q., & Qian, Y. (2019). Feature–granularity selection with variable costs for hybrid data. Soft Computing, 23, 13105–13126.

[89] Liao, S., Zhu, Q., Qian, Y., & Lin, G. (2018). Multi-granularity feature selection on cost-sensitive data with measurement errors and variable costs. Knowledge-Based Systems, 158, 25–42.

[90] Lira, H., Ko, I.-Y., & Jimenez-Molina, A. (2018). Mental workload assessment in smartphone multitasking users: a feature selection approach using physiological and simulated data. 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), 639–642.

[91] Liu, M., Miao, L., & Zhang, D. (2014). Two-stage cost-sensitive learning for software defect prediction. IEEE Transactions on Reliability, 63(2), 676–686.

[92] Liu, M., Xu, C., Luo, Y., Xu, C., Wen, Y., & Tao, D. (2017a). Cost-sensitive feature selection by optimizing F-measures. IEEE Transactions on Image Processing, 27(3), 1323–1335.

[93] Liu, M., Xu, C., Luo, Y., Xu, C., Wen, Y., & Tao, D. (2017b). Cost-sensitive feature selection via f-measure optimization reduction. Proceedings of the AAAI Conference on Artificial Intelligence, 31(1).

[94] Long, X., Qian, W., Wang, Y., & Shu, W. (2021). Cost-sensitive feature selection on multi-label data via neighborhood granularity and label enhancement. Applied Intelligence, 51, 2210–2232.

[95] López, B. R. (2014). Advancing the diagnosis of dry eye syndrome: development of automated assessments of tear film lipid layer patterns. Universidade da Coruña.

[96] Lu, H., Xu, Y., Ye, M., Yan, K., Gao, Z., & Jin, Q. (2019). Learning misclassification costs for imbalanced classification on gene expression data. BMC Bioinformatics, 20, 1–10.

[97] Machiraju, S., & Gaurav, S. (2018). Power BI Data Analysis and Visualization. De-G Press.

[98] Maldonado, S., Pérez, J., & Bravo, C. (2017). Cost-based feature selection for support vector machines: An application in credit scoring. European Journal of Operational Research, 261(2), 656–665.

[99] McCombe, N., Ding, X., Prasad, G., Finn, D. P., Todd, S., McClean, P. L., Wong-Lin, K., Initiative, N., & others. (2022). Multiple Cost Optimisation for Alzheimer's Disease Diagnosis. 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 1098–1104.

[100] Mccombe, N., Ding, X., Prasad, G., Gillespie, P., Finn, D. P., Todd, S., Mcclean, P. L., & Wong-Lin, K. (2022). Alzheimer's disease assessments optimized for diagnostic accuracy and administration time. IEEE Journal of Translational Engineering in Health and Medicine, 10, 1–9.

[101] Mej-Lavalle, M. (2008). Applying cost sensitive feature selection in an electric database. International Symposium on Methodologies for Intelligent Systems, 644–649.

[102] Min, F., Hu, Q., Zhu, W. (2014). Feature selection with test cost constraint. International Journal of Approximate Reasoning, 55(1), 167–179.

[103] Min, F., & Xu, J. (2016). Semi-greedy heuristics for feature selection with test cost constraints. Granular Computing, 1(3), 199–211.

[104] Miri, M., Dowlatshahi, M. B., & Hashemi, A. (2022). Evaluation multi label feature selection for text classification using weighted borda count approach. 2022 9th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), 1–6.

[105] Miri, M., Dowlatshahi, M. B., Hashemi, A., Rafsanjani, M. K., Gupta, B. B., Alhalabi, W. (2022). Ensemble feature selection for multi-label text classification: An intelligent order statistics approach. International Journal of Intelligent Systems, 37(12), 11319–11341.

[106] Mohanrasu, S. S., Phan, L. T., Rajan, R., & Manavalan, B. (2025). Cost-sensitive feature selection for multi-label classification: multi-criteria decision-making approach. Applied Computing and Informatics.

[107] Momeni, N., Valdés, A. A., Rodrigues, J., Sandi, C., & Atienza, D. (2021). CAFS: cost-aware features selection method for multimodal stress monitoring on wearable devices. IEEE Transactions on Biomedical Engineering, 69(3), 1072–1084.

[108] Namakin, M., Rouhani, M., & Sabzekar, M. (2022). An evolutionary correlation-aware feature selection method for classification problems. Swarm and Evolutionary Computation, 75, 101165.

[109] Nguyen, B. H., Xue, B., & Zhang, M. (2020). A survey on swarm intelligence approaches to feature selection in data mining. Swarm and Evolutionary Computation, 54, 100663.

[110] Niu, J., Zhao, H., & Zhu, W. (2014). A Logarithmic Weighted Algorithm for Minimal Test Cost Attribute Reduction. Rough Sets and Knowledge Technology: 9th International Conference, RSKT 2014, Shanghai, China, October 24-26, 2014, Proceedings 9, 129–138.

[111] Niu, J., Zhao, H., & Zhu, W. (2016). Feature selection with test cost constraint through a simulated annealing algorithm. Journal of Internet Technology, 17(6), 1133–1140.

[112] Ogawa, H., Koga, T., Pham, N.-A., Bernards, N., Gregor, A., Sata, Y., Kitazawa, S., Hiraishi, Y., Ishiwata, T., Aragaki, M., & others. (2024). Clinical and pathological predictors of engraftment for patient-derived xenografts in lung adenocarcinoma. Lung Cancer, 194, 107863.

[113] Pendharkar, P. C. (2006). A data mining-constraint satisfaction optimization problem for cost effective classification. Computers & Operations Research, 33(11), 3124–3135.

[114] Pocock, A. C. (2012). Feature selection via joint likelihood. University of Manchester.

[115] Porto Diaz, I. (2015). Novel machine learning methods based on information theory.

[116] Qian, G., Yeung, D., Tsang, E. C. C., & Shu, W. (2004). Feature selection for Chinese character recognition based on inductive learning. International Journal of Pattern Recognition and Artificial Intelligence, 18(08), 1453–1471.

[117] Qian, W., Shu, W., Yang, J., & Wang, Y. (2015). Cost-sensitive feature selection on heterogeneous data. Advances in Knowledge Discovery and Data Mining: 19th Pacific-Asia Conference, PAKDD 2015, Ho Chi Minh City, Vietnam, May 19-22, 2015, Proceedings, Part II 19, 397–408.

[118] Raynal, L., Hoffmann, T., & Onnela, J.-P. (2023). Cost-based feature selection for network model choice. Journal of Computational and Graphical Statistics, 32(3), 1109–1118.

[119] Raynal, L., & Onnela, J. P. (2021). Selection of Summary Statistics for Network Model Choice with Approximate Bayesian Computation. ArXiv Preprint ArXiv:2101.07766, 1–34.

[120] Rong, M., Gong, D., & Gao, X. (2019). Feature selection and its use in big data: challenges, methods, and trends. Ieee Access, 7, 19709–19725.

[121] Saeedi, R. (2018). EFFICIENT MACHINE LEARNING ALGORITHMS FOR AUTOMATIC RECONFIGURATION OF MOBILE. Washington State University.

[122] Santos-Rodriguez, R., & Garcia-Garcia, D. (2010). Cost-sensitive feature selection based on the set covering machine. 2010 IEEE International Conference on Data Mining Workshops, 740–746.

[123] Secerbegovic, A., Gogic, A., Suljanovic, N., Zajc, M., & Mujcic, A. (2018). Computational balancing between wearable sensor and smartphone towards energy-efficient remote healthcare monitoring. Advances in Electrical and Computer Engineering, 18(4), 3–10.

[124] Seethalakshmi, T., Murugan, M., & Vimalan, P. M. A. S. (2024). Smart material selection strategies for sustainable and cost-effective high-performance concrete production using deep learning. Automatika, 65(4), 1533–1544.

[125] Sheikhpour, R., Sarram, M. A., Gharaghani, S., & Chahooki, M. A. Z. (2017). A survey on semi-supervised feature selection methods. Pattern Recognition, 64, 141–158.

[126] Sheng, V. S., Ling, C. X., Ni, A., & Zhang, S. (2006). Cost-sensitive test strategies. Proceedings of the National Conference on Artificial Intelligence, 21(1), 482.

[127] Shi, S., Mahzoon, V., Lindner, G., Vucetic, S., & Miskovic, S. (n.d.). Optimizing Detector Placement, Number, and Size in Radioactive Particle Tracking Using Feature Selection. Number, and Size in Radioactive Particle Tracking Using Feature Selection.

[128] Smits, P. C., & Annoni, A. (2000a). Cost-based feature selection for GIS-embedded data fusion. IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No. 00CH37120), 6, 2614–2616.

[129] Smits, P. C., & Annoni, A. (2000b). Cost-based feature subset selection for interactive image analysis. Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, 2, 386–389.

[130] Srimani, P. K., & Koti, M. S. (2011). The impact of rough set approach on medical diagnosis for cost effective feature selection. IJCR, 3(12), 175–178.

[131] Sun, Y., Li, M., Li, L., Shao, H., & Sun, Y. (2021). Cost-Sensitive Classification for Evolving Data Streams with Concept Drift and Class Imbalance. Computational Intelligence and Neuroscience, 2021(1), 8813806.

[132] Suryani, E., Setyawan, S., Putra, B. P., & others. (2022). The cost-based feature selection model for coronary heart disease diagnosis system using deep neural network. IEEE Access, 10, 29687–29697.

[133] Tahir, S. F. (2016). Resource-constrained re-identification in camera networks. Queen Mary University of London.

[134] Tan, A., Wu, W., & Tao, Y. (2017). A set-cover-based approach for the test-cost-sensitive attribute reduction problem. Soft Computing, 21, 6159–6173.

[135] Tao, Y., Lu, G., Ma, C., Su, Z., & Hu, Z. (2021). Semi-supervised Feature Selection Based on Cost-Sensitive and Structural Information. Databases Theory and Applications: 32nd Australasian Database Conference, ADC 2021, Dunedin, New Zealand, January 29–February 5, 2021, Proceedings 32, 23–36.

[136] Taylor, P., Griffiths, N., Hall, V., Xu, Z., & Mouzakitis, A. (2022). Feature selection for supervised learning and compression. Applied Artificial Intelligence, 36(1), 2034293.

[137] Teisseyre Pawełand Klonecki, T. (2021). Controlling costs in feature selection: information theoretic approach. Computational Science–ICCS 2021: 21st International Conference, Krakow, Poland, June 16–18, 2021, Proceedings, Part II 21, 483–496.

[138] Teisseyre Pawełand Zufferey, D., & Słomka, M. (2019). Cost-sensitive classifier chains: Selecting low-cost features in multi-label classification. Pattern Recognition, 86, 290–319.

[139] Valancius, M., Lennon, M., & Oliva, J. (2023). Acquisition Conditioned Oracle for Nongreedy Active Feature Acquisition. ArXiv Preprint ArXiv:2302.13960.

[140] Vu, L., Nguyen, P., & Turaga, D. (2016). Firstfilter: a cost-sensitive approach to malicious URL detection in large-scale enterprise networks. IBM Journal of Research and Development, 60(4), 1–4.

[141] Weiss, Y., Elovici, Y., & Rokach, L. (2013). The cash algorithm-cost-sensitive attribute selection using histograms. Information Sciences, 222, 247–268.

[142] Weiss, Y., Fledel, Y., Elovici, Y., & Rokach, L. (2012). Cost-sensitive detection of malicious applications in mobile devices. Mobile Computing, Applications, and Services: Second International ICST Conference, MobiCASE 2010, Santa Clara, CA, USA, October 25-28, 2010, Revised Selected Papers 2, 382–395.

[143] Xiong, Y., & Zuo, R. (2017). Effects of misclassification costs on mapping mineral prospectivity. Ore Geology Reviews, 82, 1–9.

[144] Xue, B., Zhang, M., Browne, W. N., & Yao, X. (2015). A survey on evolutionary computation approaches to feature selection. IEEE Transactions on Evolutionary Computation, 20(4), 606–626.

[145] Yan, D., Qin, Z., Gu, S., Xu, H., & Shao, M. (2021). Cost-sensitive selection of variables by ensemble of model sequences. Knowledge and Information Systems, 63, 1069–1092.

[146] Yang, T., Wang, X., Zheng, K., Guo, Z., Qian, Y., & Qiao, F. (n.d.). An Efficient Minimum Cost Attribute Reduction Algorithm for Large Scale Data. Available at SSRN 4927688.

[147] Yu, J.-R., Lin, C.-Y., & Lien, D. D. (n.d.). A Cost-Effective and Accuracy-Sensitive L1-Norm Support Vector Machine Model for Feature Selection. Available at SSRN 4904961.

[148] Yu, S., & Zhao, H. (2018). Rough sets and Laplacian score based cost-sensitive feature selection. PloS One, 13(6), e0197564.

[149] Yue, L., Hu, P., Chu, S.-C., & Pan, J.-S. (2023). Multi-objective gray wolf optimizer with cost-sensitive feature selection for predicting students' academic performance in college english. Mathematics, 11(15), 3396.

[150] Zahirnia, K., Teimouri, M., Rahmani, R., & Salaq, A. (2015). Diagnosis of type 2 diabetes using cost-sensitive learning. 2015 5th International Conference on Computer and Knowledge Engineering (ICCKE), 158–163.

[151] Zangooei, A., Derhami, V., & Jamshidi, F. (2019). A Novel Architecture for Detecting Phishing Webpages using Cost-based Feature Selection. Journal of AI and Data Mining, 7(4), 607–616.

[152] Zhang, D., & Shen, D. (2011). Multicost: multi-stage cost-sensitive classification of alzheimer's disease. Machine Learning in Medical Imaging: Second International Workshop, MLMI 2011, Held in Conjunction with MICCAI 2011, Toronto, Canada, September 18, 2011. Proceedings 2, 344–351.

[153] Zhang, Y., Cheng, S., Shi, Y., Gong, D., & Zhao, X. (2019). Cost-sensitive feature selection using two-archive multi-objective artificial bee colony algorithm. Expert Systems with Applications, 137, 46–58.

[154] Zhang, Y., Gong, D., & Cheng, J. (2015). Multi-objective particle swarm optimization approach for cost-based feature selection in classification. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 14(1), 64–75.

[155] Zhang, Y., Gong, D., Rong, M., & Guo, Y. (2016). Interval cost feature selection using multi-objective PSO and linear interval programming. Advances in Swarm Intelligence: 7th International Conference, ICSI 2016, Bali, Indonesia, June 25-30, 2016, Proceedings, Part I 7, 579–586.

[156] Zhang, Y., Song, X., & Gong, D. (2017). A return-cost-based binary firefly algorithm for feature selection. Information Sciences, 418, 561–574.

[157] Zhang, Y., Zhang, J., Guo, Y., & Sun, X. (2016). Fuzzy cost-based feature selection using interval multi-objective particle swarm optimization algorithm. Journal of Intelligent & Fuzzy Systems, 31(6), 2807–2812.

[158] Zhao, H., Min, F., & Zhu, W. (2013a). A backtracking approach to minimal cost feature selection of numerical data. JOURNAL OF INFORMATION &COMPUTATIONAL SCIENCE, 10(13), 4105–4115.

[159] Zhao, H., Min, F., & Zhu, W. (2013b). Cost-Sensitive Feature Selection of Numeric Data with Measurement Errors. Journal of Applied Mathematics, 2013(1), 754698.

[160] Zhao, H., Wang, P., & Hu, Q. (2016). Cost-sensitive feature selection based on adaptive neighborhood granularity with multi-level confidence. Information Sciences, 366, 134–149.

[161] Zhao, H., & Yu, S. (2019). Cost-sensitive feature selection via the L2, 1-norm. International Journal of Approximate Reasoning, 104, 25–37.

[162] Zhao, H., & Zhu, W. (2014). Optimal cost-sensitive granularization based on rough sets for variable costs. Knowledge-Based Systems, 65, 72–82.

[163] Zhou, Q., Zhou, H., & Li, T. (2016). Cost-sensitive feature selection using random forest: Selecting low-cost subsets of informative features. Knowledge-Based Systems, 95, 1–11.

Saba Beiranvand
Orcid number: 0000-0002-3468-2981
Department of Computer Engineering
National University of Skills (NUS)
Tehran , Iran
Email address: bir.saba@yahoo.com

Mohammad Bagher Dowlatshahi
Orcid number: 0000-0002-4862-0626
Department of Computer Engineering
Lorestan University
Khorramabad, Iran
Email address: dowlatshahi.mb@lu.ac.ir

Amin Hashemi
Orcid number: 0000-0002-3768-0615
Department of Computer Engineering
Lorestan University
Khorramabad, Iran
Email address: hashemi.a@lu.ac.ir