Journal of Basics and Applied Sciences Research (JOBASR)
ISSN (print): 3026-9091, ISSN (online): 1597-9962
Volume 3(3) May 2025
DOI: https://dx.doi.org/10.4314/jobasr.v3i3.33

Multiobjective Optimization Solution for the Selection of Quasi Equally Informative Subsets in Classification Models

Abubakar I. Safyan[1*], Zaharaddeen Sani[2] and Mukhtar Abubakar[3]
[1, 2 & 3]Department of Computer Science and Information Technology, Faculty of Computing, Federal University Dutsin-Ma, Dutsin-Ma, Katsina State, Nigeria.
*Corresponding Author Email: asibrahim@fudutsinma.edu.ng

ABSTRACT
Feature selection is crucial in machine learning, particularly for high-dimensional data. This study presents two advanced multi-objective techniques—Improved Wrapper QEISS (IW-QEISS) and Improved Filter QEISS (IF-QEISS)—designed to identify multiple quasi-equally informative feature subsets. Unlike traditional methods, which focus solely on accuracy and subset size, our approach enhances robustness and interpretability. Using a four-objective NSGA-II framework with a population of 100 and 100 generations, we optimize for accuracy, redundancy, and feature importance (threshold = 0.05). Experiments show IW-QEISS identified seven subsets with a cardinality of four on the Heart dataset, achieving 0.836 accuracy—on par with W-MOSS. IF-QEISS offered similar accuracy with reduced computation. These results validate the efficiency and effectiveness of our proposed methods.

Keywords:
Quasi Equally, Informative Subsets,
Extreme Learning Machine (ELM),
Multi-objective Optimization,
Pareto Efficiency
Machine Learning.

INTRODUCTION

Feature selection is crucial in building efficient and precise machine learning models, especially with large, multidimensional datasets. The challenge lies in optimizing multiple objectives like model performance, feature count, and feature relevance versus redundancy. A potential strategy that has gained traction is multi-objective feature selection. To address these trade-offs, offering a better understanding of feature subset interactions has become necessary (Dowlatshahi and Hashemi 2024). Traditional methods like Elastic Net Regularization, Random Forest, and Gradient Boosting Machines often focus on maximizing classification accuracy while minimizing feature count, overlooking the potential for multiple feature subsets with similar information richness (Agrawal *et al.*, 2023). Entropy, a measure of the unpredictability or information content of a variable, is central to comparing the information richness of feature subsets. If two variables, X and Y, have similar entropy values, they carry approximately equal information. This concept is valuable in domains like machine learning and data compression when assessing the informativeness of different features (Chandrashekar and Sahin 2014). This paper introduces the Improved Wrapper for Quasi Equally Informative Subset Selection (IW-QEISS), a novel multi-objective wrapper feature selection method. IW-QEISS expands the search space to include quasi-equally informative subsets, evolving a diverse Pareto front that optimizes multiple objectives, including relevance and redundancy measures.

Generally speaking, there are three types of feature selection techniques: wrappers, filters, and embedding approaches. Wrappers interact directly with the model, identifying optimal subsets but at a higher computational cost. Filters use statistical measures, which are computationally efficient but may not align with the model's needs. By including feature selection into model training, embedded approaches provide customized feature selection at the price of generalizability. (Muñoz-Romero *et al.*, 2020). Conventional feature selection focused on individual objectives like accuracy or subset size (Omolara et al. 2021), but multi-objective approaches now estimate the Pareto front to evaluate trade-offs between objectives (Gunantara 2018).

Hybrid approaches combining population-based multi-objective optimization with wrapper-based evaluation are gaining traction. While evolutionary algorithms efficiently explore complex search spaces, wrappers provide precise model estimates (Sharma and Kumar 2022). However, existing methods have limitations: they often overlook intermediate solutions with larger subsets that maintain accuracy and fail to identify multiple quasi-equally informative subsets for a given cardinality. These gaps highlight the need for improved methods that consider these factors to enhance our understanding of feature relevance, dependencies, and

How to cite this article: Safyan, A. I., Sani, Z. and Abubakar, M. (2025). Multiobjective Optimization Solution for the Selection of Quasi Equally Informative Subsets in Classification Models. *Journal of Basics and Applied Sciences Research, 3*(3), 322-330. https://dx.doi.org/10.4314/jobasr.v3i3.33

redundancy (Cervantes *et al.*, 2022; Sharma and Kumar 2022).

Advancing feature selection methods is essential for the continued progress of machine learning. Methods like IW-QEISS that explore multiple quasi-equally informative feature subsets offer a promising direction for more transparent and effective model development (Wang *et al.*, 2022; Mukhopadhyay *et al.*, 2019).

Dependence between variables is often measured using information-theoretic criteria like Mutual Information and Partial Mutual Information due to their flexibility in modeling various situations without assuming a specific functional relationship (Taormina 2016; Mielniczuk 2022). The Shannon entropy, which measures a random variable's uncertainty, is the foundation of these standards (Saraiva 2023). Correlation-based measures, particularly the linear correlation coefficient, are commonly used for analyzing redundancy and relevance, although they may not effectively capture nonlinear relationships (Han *et al.*, 2024; Schober *et al.*, 2018). Information-theoretic criteria, such as mutual information and symmetric uncertainty, which are rooted in entropy concepts, have gained popularity in machine learning for their ability to handle both linear and nonlinear relationships (Auffarth *et al.*, 2010; Han *et al.*, 2024).

Entropy is used in information theory to quantify the degree of uncertainty related to a random variable (Saraiva 2023). A discrete random variable X's entropy H(X) has the following formal definition:

$$H(X) = -\int p(x) log p(x) dx \tag{1}$$

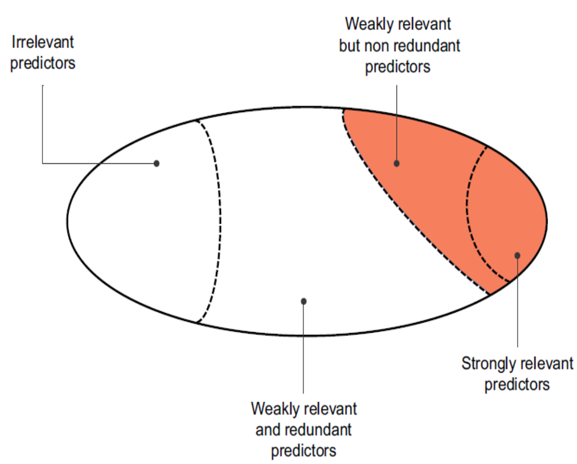

Figure 1. A classification of predictors

A classification of predictors according to redundancy and relevance is shown in Figure 1 (derived from Yu and Liu 2004). Strongly relevant and non-redundant weakly relevant predictors make up the ideal subset, which is shown in red.

The set of all potential values of X is represented by X, with $p(x) = \Pr(X = x), x\epsilon X$. Similarly, the joint entropy $H(X,Y)$, of two continuous random variables X and Y is defined as:

$$H(X,Y) = -\iint p(x,y) log p(x,y) dx dy \tag{2}$$

where $p(x,y)$ represents the joint probability density function of X and Y. When provided with a univariate sample of X and a bivariate sample of (X, Y), the discrete form of Equations (1) and (2) can be expressed as follows:

$$H(X) = -\sum_{x\epsilon X} p(x) \log p(x), \tag{3}$$

and

$$H(X,Y) = -\sum_{x\epsilon X}\sum_{y\epsilon Y} p(x,y) log p(x,y) \tag{4}$$

where $p(x,y) = Pr(X = x, Y = y)$, $x\epsilon X$, and $y\epsilon Y$, while the mutual information between X and Y is expressed as:

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \tag{5}$$

Mutual information has been standardized to create symmetric uncertainty, which has values between 0 and 1. According to Gomes and Figueiredo (2024), a value of 0 denotes independence between the variables, but a value of 1 indicates that knowledge of one variable is sufficient to accurately determine the other. Formally speaking, symmetric uncertainty is defined as:

$$SU(X,Y) = \frac{2I(X,Y)}{H(X) + H(Y)} \tag{6}$$

When dealing with continuous data, it becomes crucial to employ appropriate techniques for estimating the joint and marginal probability density functions $p(x,y), p(x)$, and $p(y)$.

The characteristics of multiobjective optimization issues include the presence of two or more objectives that frequently clash (Coello *et al.*, 2019). A multiobjective optimization issue can generally be stated as follows:

$$\text{"}minimize/Maximize\text{"}\ f(x) = [f_1(x), f_2(x), \dots, f_k(x)] \tag{7}$$

The j-th objective function is denoted by f_j, the feasible decision space by X, the decision vector by x, and the total number of objectives by p. A mathematical operation that is not clearly defined is vector minimization, as the quote marks in the above equation show. Only when no alternative solution exists does a given solution X qualify as Pareto-efficient $x\epsilon X$ such that $f_j(\hat{x}) < f_j(x)$ for all j and $f_j(\hat{x}) < f_j(x)$ for at least one j.

Khan *et al.*, 2019 developed a computational model for distinguishing cancer lectins with high accuracy, sensitivity, and specificity. Wang *et al.*, 2021 combined ensemble learning with multi-objective optimization to develop a fault diagnosis method for planetary gearboxes, achieving a classification accuracy of 99.62%. Prina et al. 2018 introduced EPLANopt, a model combining EnergyPLAN with a genetic algorithm, optimizing energy systems to reduce costs and emissions, particularly in South Tyrol by 2050. Audet *et al.,* 2021 provided a comprehensive review of performance metrics for Pareto front approximations, classifying 63 indicators into four categories. MOSHO, a multi-objective version of the Spotted Hyena Optimizer, was proposed by Dhiman and Kumar (2018). It performed better in terms of convergence and diversity than MOPSO and NSGA-II. For microarray gene data, Chaudhuri and Sahu 2022 created a hybrid feature selection technique, which reduced error rates and improved selection speed. Schweidtmann *et al.*, 2018 introduced TS-EMO, a Bayesian optimization technique for chemical reactions, demonstrating the effective trade-offs between objectives with minimal experimentation.

Wang *et al.*, 2023 developed NMDE, a niching multi-objective differential evolution approach, which found diverse and high-quality feature subsets for classification, outperforming other methods. In 2020, Blank and Deb presented Pymoo, a versatile multi-objective optimization framework for testing algorithms and making decisions. Emmerich and Deutz 2018 reviewed fundamental concepts and algorithms in evolutionary multi-objective optimization, including NSGA-II and MOEA/D.

Many studies have shown that using multi-objective optimization in structural engineering and text classification, respectively, improves performance over traditional methods (Afshari *et al.,* 2019; Labani et al. 2020). Other notable contributions include methods for constrained multi-objective optimization (Li *et al.*, 2019), dynamic optimization using manifold transfer learning (Jiang *et al.,* 2021), and multi-population approaches (Ma *et al.,* 2020), all of which have shown significant improvements in computational efficiency and optimization performance.

Despite these advancements, many approaches remain limited to two or three objectives. This research aims to extend the complexity of multi-objective optimization by developing a four-objective wrapper-based and filter-based feature selection technique, leveraging enhanced Extreme Learning Machine (ELM) models to improve search space exploration and computational efficiency (Taormina *et al.,* 2016; Sun *et al.,* 2020; Kale and Sonavane 2019; Khan *et al.,* 2019; Wang *et al.,* 2021).

Research Gap

Despite progress in multi-objective feature selection, most existing methods are limited to optimizing only two or three objectives and often fail to identify multiple quasi-equally informative subsets with similar predictive power. Traditional approaches focus narrowly on maximizing accuracy and minimizing feature count, overlooking feature redundancy and the potential for diverse informative subsets. Moreover, many models do not adequately explore the trade-offs between relevance, redundancy, and interpretability. This leaves a gap in discovering alternative feature sets that are equally valid yet structurally different. Addressing this requires more comprehensive and computationally efficient optimization techniques.

MATERIALS AND METHODS

Here, we define the notion of quasi-equivalency between subsets and explain the importance, redundancy, and precision metrics that were applied in this paper. We explain a computationally less demanding filter, provide the improved wrapper for recognizing quasi-equally informative subsets, and explore implementation details of the learning algorithm and global optimization technique.

Problem Definition

We assume that $\hat{f}(.)$in this study is a metric that, on a scale from 0 to 1, assesses prediction accuracy. A model with no predictive ability is represented by a value of 0, whereas a model with perfect predictive accuracy is represented by a value of 1. We present the idea of δ-quasi equivalency between predictor variable subsets. When two subsets of predictors, S_j and S_i, are used to create models for a particular model class, and if the models produced by the two subsets have extremely similar levels of predictive accuracy, consequently $S_i \delta$-quasi equally informative to S_j. More specifically, if subset if: $\hat{f}(S_i) \geq (1-\delta)\hat{f}(S_j) for\ 0 \leq \delta \leq 1$ then subset S_i is d-quasi equally informative to subset S_j. Where $\delta$is a predetermined threshold for the permitted variation in predictive accuracy, and $f(S_i)$ and $f(S_j)$ represent the predictive accuracy obtained using subsets S_i and S_j, respectively. The observational dataset's unique variable interactions determine whether quasi-equivalent subsets exist. Depending on their significance in connection to the output variable, predictors can be categorized as strongly relevant, weakly relevant, or irrelevant, according to research by Pudjihartono *et al.,* 2022.

Objective Function

Let X be the set of candidate features, S be a subset of X, and y be the target classification variable. The relevance metric should be optimized throughout the search process, represented by $f_1(S)$, which is defined as follows:

$$f_1(S) = \sum_{x_i \in S} SU(x_i, y) \tag{8}$$

where $SU(x_i, y)$ is the symmetric uncertainty between the ith predictor $x_i$and the output y. The metric of redundancy$f_2(S)$ to be minimized is:

$$f_2(S) = \sum_{x_i, x_j \in S,\ i<j} SU(x_i, y) \quad (9)$$

The function $SU(x_i, x_j)$ represents the symmetric uncertainty between the predictors x_i and x_j. Put another way, for a given subset S, $f_1(S)$ evaluates the extent to which the predictors in S explain the outcome, whereas $f_2(S)$ quantifies the degree of likeness. Consequently, the search is directed toward choosing predictors that are very different from one another by minimizing $f_2(S)$ (Hanke *et al.*, 2023).

$$f_3(S) = |S| \quad (10)$$

The learning algorithm's performance in classification is evaluated using the accuracy measure that follows for the fourth objective:

$$f_4(S) = \frac{tp + tn}{tp + tn + fp + fn} \quad (11)$$

Using a pre-defined model architecture, the symmetric uncertainty $SU(y, \hat{y}(S))$ measures the information shared between the observed output y and the prediction $\hat{y}(S)$ acquired through the subset S. An important benefit of using an entropy-based metric for predictive accuracy is that it captures the distributional features present in the flow duration curve of both the observed and forecasted streamflow (Zhang 2024; Zhou *et al.*, 2024). The use of information-theoretic measures for hydrologic model evaluation is becoming more popular, as shown by works like Teegavarapu et al. 2022, Beven 2024, Brauman *et al.*, 2021, and Gupta *et al.*, 2024.

Datasets and Parameter Setting

Heart and Concrete, two distinct datasets from the UCI machine learning library, were used to test the algorithms. A variety of binary and multiclass classification tasks were purposefully represented in these datasets to provide a thorough evaluation of the algorithms' performance. The selected datasets will show varying degrees of complexity in terms of the number of classes; these will cover binary classification jobs with two classes to more challenging issues with up to ten classes.

RESULTS AND DISCUSSION

The IW-QEISS (IF-QEISS) algorithm identifies multiple feature subsets nearly as informative as the best subset using two UCI datasets. It enables superior trade-offs between classification metrics and assesses feature importance, while also comparing the computational power required by each algorithm. Subset sizes are capped at 20 or 50 features, depending on the dataset.

Quasi Equally Informative Subset Comparison

The number of features in the datasets varies, affecting classifier accuracy, which generally improves with more features until overfitting occurs. The W-QEISS algorithm often identifies multiple nearly equivalent subsets for each feature count, as seen in the Heart dataset where it found seven informative subsets with a cardinality of four (features 3, 9, 12, and 13), as opposed to one subset that the W-MOSS algorithm discovered with 0.836 accuracy:

$$S1 = \{2, 3, 12, 13\}, S2 = \{3, 8, 12, 13\},$$
$$S3 = \{3, 9, 12, 13\},$$
$$S4 = \{5, 11, 12, 13\},$$
$$S5 = \{7, 8, 12, 13\},$$
$$S6 = \{8, 9, 12, 13\},$$
$$S7 = \{9, 10, 12, 13\}$$

The IW-QEISS algorithm identified subsets with accuracy ranging from 0.807 (S4) to 0.836 (S3), matching the W-MOSS solution. It also found larger subsets (cardinalities 7, 8, and 9) and achieved better classification accuracy for certain datasets due to its four-objective optimization approach. This expands the search space, enabling the discovery of subsets with higher discriminating power. The mRMR algorithm's performance varied, sometimes matching W-MOSS accuracy, but often lower for other subsets.

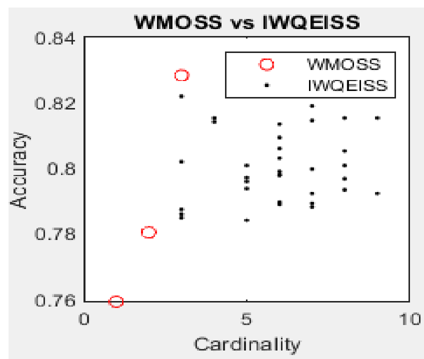

Fig 2: Binary Classification 1

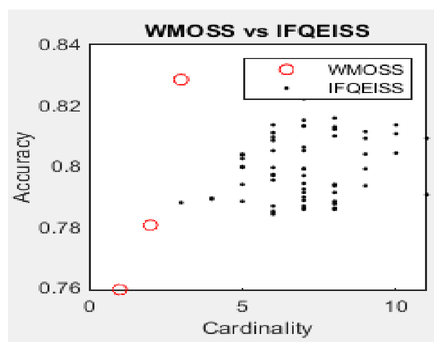

Figure 3: Binary Classification 2

The W-QEISS (F-QEISS) and W-MOSS plots in Figures 1 and 2 illustrate subsets that are nearly equally informative at the 0.05 level.

Table 1: Comparison of the Four Algorithms

Data set	Algorithm	Sub sets	Avg Sub sets	Best Accuracy	δ – Quasi Equally informative subsets			
					0-1 %	1-5 %	5-10 %	10-20 %
Heart	IW-QEISS IF-QEISS W-MOSS mRMR	36 20 4 1	5.14 4.00 1.00 1.00	0.8430: 7 0.8400: 6 0.8430: 6 0.8289: 6	20 (20) 9 (7) 31 (0)	16 (16) 11 (11) 10 (1)	21 (21) 6 (8) 10 (0)	13 (13) 3 (3) 10 (0)
Concrete	IW-QEISS IF-QEISS W-MOSS mRMR	61 49 5 1	11.87 7 1.00 1.00	0.8430: 5 0.8240:: 7 0.7590: 6 0.7538: 6	55 (46) 27 (27) 31 (1)	35 (35) 12 (12) 20 (0)	0 (0) 0 (0) 10 (0)	0 (0) 0 (0) 00 (0)

The IF-QEISS algorithm, like IW-QEISS, can find multiple feature subsets per cardinality level, but the accuracy of linked classifiers is generally lower since the search focuses on relevance, redundancy, and feature count rather than classification accuracy. However, this trade-off results in improved processing performance. As the number of candidate characteristics rises, the IW-QEISS and IF-QEISS approaches frequently identify several quasi-equally informative subsets. For the Heart dataset, for instance, IW-QEISS discovered an average of 5.14 subsets per cardinality level, frequently matching or surpassing the accuracy of the W-MOSS method. The table summarizes the number of subsets, average subsets per cardinality level, and the highest accuracy achieved by each method, highlighting that IW-QEISS typically identifies more accurate subsets with higher cardinality. The final columns show that IW-QEISS and IF-QEISS can discover many δ-quasi equally informative subsets, with the IW-QEISS method's maximum accuracy closely matching that of the W-MOSS algorithm, making changes in reference minimal.

Categorization of Features

Identifying multiple equally informative feature subsets at each cardinality level offers insights into the importance of selected attributes. A recent technique by Wang *et al.*, 2023 categorizes features into strongly relevant, weakly relevant (non-redundant), and irrelevant groups. The Heart dataset illustrates this approach.

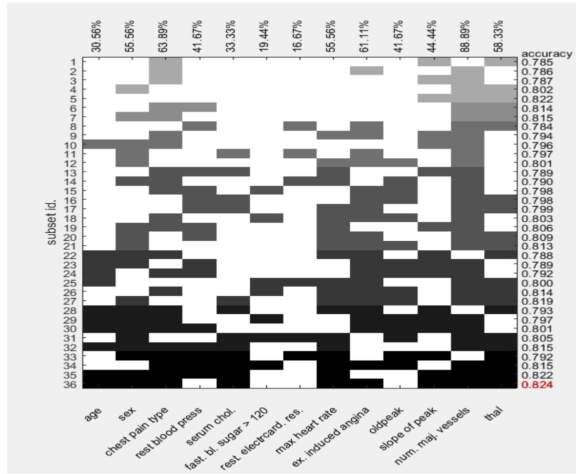

Figure 4: The Heart Dataset's Quasi-Equally Informative Subsets' Feature Selection Frequency

The IW-QEISS algorithm determined that 36 feature subsets were 0.05-quasi equally informative to the best subset, as shown in Figure 4. Each row represents a subset, and each column corresponds to one of the 13 possible features. The presence of a feature in a subset is highlighted, with darker shades indicating higher cardinality. The figure shows that certain features, like 12 and 13, consistently appear in subsets with cardinality 3 or higher, suggesting their strong relevance. Features like 8 and 9 enter the subsets at higher cardinalities, indicating varying levels of importance. The persistent presence of certain features across multiple subsets suggests they are more likely to be highly relevant, while infrequent features are likely less important. This approach allows for a nuanced categorization of features, distinguishing between strongly relevant, weakly relevant, and irrelevant features, aiding in the interpretation of the decision-making process.

CONCLUSION

The development of the Improved Wrapper for Quasi Equally Informative Subset Selection (IW-QEISS) and its filter-based counterpart, F-QEISS, offers a robust solution for addressing multiple objectives in feature selection, including relevance, redundancy, and classification accuracy. By employing the concept of quasi-equivalency, these models identify subsets of features that maintain high predictive performance while improving interpretability and reducing redundancy. Extensive experimentation with high-dimensional datasets showcases the method's ability to outperform traditional single- and bi-objective approaches. This advancement underscores the importance of multi-objective optimization in creating more stable, interpretable, and accurate machine learning models. Future research should continue to explore the method's applicability to diverse datasets and investigate further optimization strategies to enhance performance.

Recommendation

Based on the findings of this study, it is recommended that future feature selection efforts in machine learning, particularly for high-dimensional datasets, adopt multi-objective optimization frameworks like IW-QEISS and IF-QEISS to identify multiple quasi-equally informative subsets. This approach enhances model robustness, interpretability, and flexibility in real-world deployment. Additionally, integrating these methods into various machine learning pipelines can improve decision-making by offering alternative feature configurations without compromising accuracy. Researchers and practitioners are encouraged to explore further refinement of the algorithms to support larger datasets and more complex classification tasks, while also considering hybrid integration with deep learning models for broader applicability.

REFERENCE

Afshari, H., Hare, W., & Tesfamariam, S. (2019). Constrained multi-objective optimization algorithms: Review and comparison with application in reinforced concrete structures. Applied Soft Computing, 83, 105631. https://doi.org/10.1016/j.asoc.2019.105631

Al-Ani, A. (2005). Feature subset selection using ant colony optimization algorithm. International journal of computational intelligence, 2(1), 53-58

Asha, L. N., Dey, A., Yodo, N., & Aragon, L. G. (2022). Optimization approaches for multiple conflicting objectives in sustainable green supply chain management. *Sustainability*, *14*(19), 12790. https://doi.org/10.3390/su141912790

Audet, C., Bigeon, J., Cartier, D., Digabel, S. L., & Salomon, L. (2021). Performance indicators in multiobjective optimization. European Journal of Operational Research, 292(2), 397–422. https://doi.org/10.1016/j.ejor.2020.11.016

Azadifar, S., & Ahmadi, A. (2020). A Graph Theoretic Based Feature Selection Method Using Multi Objective PSO. *2020 28th Iranian Conference on Electrical Engineering (ICEE).* Tabriz, Iran, 2020, pp. 1-5, https://doi.org/10.1109/icee50131.2020.9260948

Babor, M., Pedersen, L., Kidmose, U., Paquet-Durand, O., & Hitzmann, B. (2022). Application of Non-Dominated Sorting Genetic Algorithm (NSGA-II) to increase the efficiency of bakery production: a case study. *Processes*, *10*(8), 1623. https://doi.org/10.3390/pr10081623

Beven, K. (2024). A brief history of information and disinformation in hydrological data and the impact on the evaluation of hydrological models. *Hydrological Sciences Journal*, 1–9. https://doi.org/10.1080/02626667.2024.2332616

Blank, J., & Deb, K. (2020). PyMoO: Multi-Objective Optimization in Python. IEEE Access, 8, 89497–89509. https://doi.org/10.1109/access.2020.2990567

Brauman, K. A., Bremer, L. L., Hamel, P., Ochoa-Tocachi, B. F., Roman-Dañobeytia, F., Bonnesoeur, V., Arapa, E., & Gammie, G. (2021). Producing valuable information from hydrologic models of nature-based solutions for water. *Integrated Environmental Assessment and Management*, *18*(1), 135–147. https://doi.org/10.1002/ieam.4511

Chaudhuri, A., & Sahu, T. P. (2022b). Multi-objective feature selection based on quasi-oppositional based Jaya algorithm for microarray data. Knowledge Based Systems, 236, 107804. https://doi.org/10.1016/j.knosys.2021.107804

Coello, C. a. C., Brambila, S. G., Gamboa, J. F., Tapia, M. G. C., & Gómez, R. H. (2019). Evolutionary multiobjective optimization: open research areas and some challenges lying ahead. *Complex & Intelligent Systems*, 6(2), 221–236. https://doi.org/10.1007/s40747-019-0113-4

Cui, Z., Zhang, J., Wang, Y., Cao, Y., Cai, X., Zhang, W., & Chen, J. (2019). A pigeon-inspired optimization algorithm for many-objective optimization problems. Science China Information Sciences, 62(7). https://doi.org/10.1007/s11432-018-9729-5

Dhiman, G., & Kumar, V. (2018). Multi-objective spotted hyena optimizer: A Multi-objective optimization algorithm for engineering problems. Knowledge Based Systems, 150, 175–197. https://doi.org/10.1016/j.knosys.2018.03.011

Emmerich, M., & Deutz, A. H. (2018). A tutorial on multiobjective optimization: fundamentals and evolutionary methods. Natural Computing, 17(3), 585–609. https://doi.org/10.1007/s11047-018-9685-y

Epstein, E., Nallapareddy, N., & Ray, S. (2023). On the Relationship between Feature Selection Metrics and

Accuracy. *Entropy*, *25*(12), 1646. https://doi.org/10.3390/e25121646

Gaspar-Cunha, A., Costa, P., Monaco, F., & Delbem, A. (2023). Many-Objectives Optimization: a machine learning approach for reducing the number of objectives. *Mathematical and Computational Applications*, *28*(1), 17. https://doi.org/10.3390/mca28010017

Gomes, A. F. C., & Figueiredo, M. a. T. (2024). A measure of synergy based on union information. *Entropy*, *26*(3), 271. https://doi.org/10.3390/e26030271

Got, A., Moussaoui, A., & Zouache, D. (2020b). A guided population archive whale optimization algorithm for solving multiobjective optimization problems. Expert Systems with Applications, 141, 112972. https://doi.org/10.1016/j.eswa.2019.112972

Gu, F., Liu, H., Cheung, Y., & Zheng, M. (2022). A Rough-to-Fine evolutionary multiobjective optimization algorithm. IEEE Transactions on Cybernetics, 52(12), 13472–13485. https://doi.org/10.1109/tcyb.2021.3081357

Gu, X., Guo, J., Xiao, L., Tao, M., & Li, C. (2019). A feature selection algorithm based on equal interval division and Minimal-Redundancy–Maximal-Relevance. *Neural Processing Letters*, 51(2), 1237–1263. https://doi.org/10.1007/s11063-019-10144-3

Gunantara, N. (2018). A review of multi-objective optimization: Methods and its applications. Cogent Engineering, 5(1), 1502242. https://doi.org/10.1080/23311916.2018.1502242

Gupta, A., Hantush, M. M., Govindaraju, R. S., & Beven, K. (2024). Evaluation of hydrological models at gauged and ungauge-d basins using machine learning-based limits-of-acceptability and hydrological signatures. *Journal of Hydrology*, 131774. https://doi.org/10.1016/j.jhydrol.2024.131774

Guyon, I, and Elisseeff, A. (2003). An introduction to variable and feature selection. Journal of Machine Learning Research. https://doi.org/10.5555/944919.944968

Hanke, M., Dijkstra, L., Foraita, R., & Didelez, V. (2023). Variable selection in linear regression models: Choosing the best subset is not always the best choice. *Biometrical Journal, 66*(1). https://doi.org/10.1002/bimj.202200209 https://doi.org/10.1109/tcbb.2020.2974953

Jia, W., Sun, M., Lian, J., & Hou, S. (2022b). Feature dimensionality reduction: a review. *Complex & Intelligent Systems*, 8(3), 2663–2693. https://doi.org/10.1007/s40747-021-00637-x

Jiang, M., Wang, Z., Qiu, L., Guo, S., Gao, X., & Tan, K. C. (2021). A fast dynamic evolutionary multiobjective algorithm via manifold transfer learning. IEEE Transactions on Cybernetics, 51(7), 3417–3428. https://doi.org/10.1109/tcyb.2020.2989465

Jiménez, F., Martínez, C., Marzano, E., Palma, J., Sánchez, G., & Sciavicco, G. (2019). Multiobjective evolutionary feature selection for fuzzy classification. IEEE Transactions on Fuzzy Systems, 27(5), 1085–1099. https://doi.org/10.1109/tfuzz.2019.2892363

Kale, A., & Sonavane, S. (2019b). IoT based Smart Farming: Feature subset selection for optimized high-dimensional data using improved GA based approach for ELM. Computers and Electronics in Agriculture, 161, 225–232. https://doi.org/10.1016/j.compag.2018.04.027

Kesireddy, A., & Medrano, F. A. (2024). Elite Multi-Criteria Decision Making—Pareto Front Optimization in Multi-Objective Optimization. *Algorithms*, *17*(5), 206. https://doi.org/10.3390/a17050206

Khan, Z., Ali, F., Ahmad, I., Hayat, M., & Pi, D. (2019c). iPredCNC: Computational prediction model for cancerlectins and non-cancerlectins using novel cascade features subset selection. Chemometrics and Intelligent Laboratory Systems, 195, 103876. https://doi.org/10.1016/j.chemolab.2019.103876

Kızılöz, H. E., Deniz, A., Dökeroğlu, T., & Çoşar, A. (2018). Novel multiobjective TLBO algorithms for the feature subset selection problem. Neurocomputing, 306, 94–107. https://doi.org/10.1016/j.neucom.2018.04.020

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, *97*(1–2), 273–324. https://doi.org/10.1016/s0004-3702(97)00043-x

Labani, M., Moradi, P., & Jalili, M. (2020). A multi-objective genetic algorithm for text feature selection using the relative discriminative criterion. *Expert Systems With Applications*, 149, 113276. https://doi.org/10.1016/j.eswa.2020.113276

Li, D., Zhang, X., Xu, Z., Liu, X., Yang, J., & Zhang, D. (2020). Adversarial feature selection for trustworthy classification. IEEE transactions on knowledge and data engineering.

Li, K., Chen, R., Fu, G., & Yao, X. (2019). Two-Archive Evolutionary Algorithm for constrained multiobjective optimization. IEEE Transactions on Evolutionary Computation, 23(2), 303–315. https://doi.org/10.1109/tevc.2018.2855411

Li, X., Lu, C., Gao, L., Xiao, S., & Wen, L. (2018). An effective multiobjective algorithm for Energy-Efficient scheduling in a Real-Life welding shop. IEEE Transactions on Industrial Informatics, 14(12), 5400–5409. https://doi.org/10.1109/tii.2018.2843441

Liu, Y., Yen, G. G., & Gong, D. (2019). A multimodal multiobjective evolutionary algorithm using Two-Archive and recombination strategies. IEEE Transactions on Evolutionary Computation, 23(4), 660–674. https://doi.org/10.1109/tevc.2018.2879406

Ma, H., Fei, M., Jiang, Z., Li, L., Zhou, H., & Crookes, D. (2020b). A Multipopulation-Based multiobjective evolutionary algorithm. IEEE Transactions on Cybernetics, 50(2), 689–702. https://doi.org/10.1109/tcyb.2018.2871473

Mathotaarachchi, K. V., Hasan, R., & Mahmood, S. (2024). Advanced machine learning techniques for predictive modeling of property prices. *Information*, *15*(6), 295. https://doi.org/10.3390/info15060295

Mirjalili, S. Z., Mirjalili, S., Saremi, S., Faris, H., & Aljarah, I. (2017). Grasshopper optimization algorithm for multi-objective optimization problems. Applied Intelligence, 48(4), 805–820. https://doi.org/10.1007/s10489-017-1019-8

Nematollahi, A. F., Rahiminejad, A., & Vahidi, B. (2019). A novel multi-objective optimization algorithm based on Lightning Attachment Procedure Optimization algorithm. Applied Soft Computing, 75, 404–427. https://doi.org/10.1016/j.asoc.2018.11.032

Prina, M. G., Cozzini, M., Garegnani, G., Manzolini, G., Moser, D., Oberegger, U. F., Pernetti, R., Vaccaro, R., & Sparber, W. (2018). Multi-objective optimization algorithm coupled to EnergyPLAN software: The EPLANopt model. Energy, 149, 213–221. https://doi.org/10.1016/j.energy.2018.02.050

Qi, F., Wu, W., Yu, Z. L., Gu, Z., Wen, Z., Yu, T., & Li, Y. (2021). Spatiotemporal-Filtering-Based channel selection for Single-Trial EEG classification. IEEE Transactions on Cybernetics, 51(2), 558–567. https://doi.org/10.1109/tcyb.2019.2963709

Qi, N., Li, X., Wu, Z., Wan, Y., Wang, N., Duan, G., Wang, L., Xiang, J., Zhao, Y., & Zhan, H. (2024). Machine Learning-Based Research for predicting shale Gas well production. *Symmetry*, *16*(5), 600. https://doi.org/10.3390/sym16050600

Saraiva, P. (2023). On Shannon entropy and its applications. *Kuwait Journal of Science*, *50*(3), 194–199. https://doi.org/10.1016/j.kjs.2023.05.004

Schweidtmann, A. M., Clayton, A. D., Holmes, N. P., Bradford, E., & Bourne, R. A. (2018). Machine learning meets continuous flow chemistry: Automated optimization towards the Pareto front of multiple objectives. Chemical Engineering Journal, 352, 277–282. https://doi.org/10.1016/j.cej.2018.07.031

Sun, C., Wang, Y., & Sun, G. (2020). A multi-criteria fusion feature selection algorithm for fault diagnosis of helicopter planetary gear train. Chinese Journal of Aeronautics, 33(5), 1549–1561. https://doi.org/10.1016/j.cja.2019.07.014

Teegavarapu, R. S., Sharma, P. J., & Patel, P. L. (2022). Frequency-based performance measure for hydrologic model evaluation. *Journal of Hydrology*, *608*, 127583. https://doi.org/10.1016/j.jhydrol.2022.127583

Wang, P., Xue, B., Liang, J., & Zhang, M. (2023b). Feature clustering-Assisted feature selection with differential evolution. *Pattern Recognition*, *140*, 109523. https://doi.org/10.1016/j.patcog.2023.109523

Wang, Y., Liu, B., Ma, Z., Wong, K., & Li, X. (2019). Nature-Inspired Multiobjective Cancer subtype diagnosis. *IEEE Journal of Translational Engineering in Health and Medicine*, 7, 1–12. https://doi.org/10.1109/jtehm.2019.2891746

Xue, B., Zhang, M., Browne, W. N., & Yao, X. (2016). A survey on Evolutionary Computation Approaches to feature selection. IEEE Transactions on Evolutionary Computation, 20(4), 606–626. https://doi.org/10.1109/tevc.2015.2504420

Yu, J., Pan, J., & Lv, Y. (2020). Solving the ED problem with ACO algorithm modified by mRMR and local search method. *2020 Chinese Automation Congress (CAC)*, Shanghai, China, 2020, pp. 2131-2136. https://doi.org/10.1109/cac51589.2020.9326683

Yue, C., Liang, J., Qu, B., & Song, H. (2019). Multimodal Multiobjective Optimization in Feature Selection. IEEE Congress on Evolutionary Computation (CEC). https://doi.org/10.1109/cec.2019.8790329

Yue, C., Qu, B., Yu, K., & Li, X. (2019). A novel scalable test problem suite for multimodal multiobjective optimization. Swarm and Evolutionary Computation, 48, 62–71. https://doi.org/10.1016/j.swevo.2019.03.011

Zhang, Y., Zhou, Z., Deng, Y., Pan, D., Van Griensven Thé, J., Yang, S. X., & bermGharabaghi, B. (2024). Daily streamflow forecasting using networks of Real-Time monitoring stations and hybrid machine learning methods. *Water*, *16*(9), 1284. https://doi.org/10.3390/w16091284

Zhou, Y., Kang, J., Kwong, S., Wang, X., & Zhang, Q. (2021). An evolutionary multi-objective optimization framework of discretization-based feature selection for classification. Swarm and Evolutionary Computation, 60,100770. https://doi.org/10.1016/j.swevo.2020.100770

Zhu, Q., Zhang, Q., & Lin, Q. (2020). A constrained multiobjective evolutionary algorithm with Detect-and-Escape strategy. IEEE Transactions on Evolutionary Computation, 24(5), 938–947. https://doi.org/10.1109/tevc.2020.2981949