Ant colony optimization plus temporal difference and opposition-based learning for semi-supervised multi-label feature selection

Sahar Shams Beiranvand [1], Vahid Mehrdad [2,*], Mohammad Bagher Dowlatshahi [3]

[1] *Department of Electrical Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran.*
Shams.s@fe.lu.ac.ir
[2] *Department of Electrical Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran.*
mehrdad.v@lu.ac.ir
[3] *Department of Computer Engineering, Faculty of Engineering, Lorestan University, Khorramabad, Iran.*
dowlatshahi.mb@lu.ac.ir

* *Corresponding author*

Abstract

This paper introduces a novel feature selection method for multi-label data, using Ant Colony Optimization (ACO) enhanced with Temporal Difference (TD) reinforcement learning, opposition-based learning (OBL), and semi-supervised learning. The key innovation lies in the application of semi-supervised learning, where we explore the impact of different proportions of labeled data—specifically 20% and 40%—on the algorithm's performance. This marks the first time such a model has been applied to multi-label datasets. In our approach, ACO is used to navigate the feature space, while TD learning dynamically updates heuristic functions, specifically the state value V, to better predict future rewards for selected feature subsets. OBL further enhances the exploration process by considering both original and opposite solutions, ensuring a more comprehensive search. The semi-supervised learning aspect of our method is particularly critical, as it allows the algorithm to leverage the structure of the unlabeled data, improving the quality of feature selection even when labeled data is scarce. We evaluated our methods on nine diverse multi-label datasets using the Multi-Label K-Nearest Neighbors (MLKNN) classifier, measuring performance with accuracy and Hamming loss metrics. The results demonstrate that our hybrid approach, which uniquely combines ACO, TD learning, OBL, and semi-supervised learning, significantly outperforms traditional feature selection techniques. The experiments with 20% and 40% labeled data highlight the effectiveness of the semi-supervised approach in enhancing classification performance and identifying more relevant feature subsets, establishing a new benchmark for multi-label feature selection.

Introduction

The rapid growth of data volume and complexity poses challenges in machine learning, particularly in model development and optimization. Feature selection addresses these challenges by identifying and retaining the most relevant features while discarding redundant or irrelevant ones [1]. This process improves model performance by reducing overfitting, enhancing accuracy, and simplifying learning [2]. Feature selection also mitigates the "curse of dimensionality," reduces computational costs, and makes models faster, more efficient, and interpretable [3]. By focusing on impactful features, data scientists can develop robust and efficient models that yield better predictions and insights [4]. Multi-label data involves instances associated with multiple labels, differing from single-label datasets where each instance has one label [5]. This paradigm better represents real-world complexities where categories often overlap [6]. Multi-label analysis demands specialized techniques to handle label dependencies, imbalances, and complexity, enabling nuanced and accurate models for diverse applications [7].

Ant Colony Optimization (ACO), inspired by ant foraging behavior, is a metaheuristic algorithm for solving combinatorial problems [8]. Artificial ants explore solutions guided by pheromone trails, enabling efficient exploration and exploitation of promising regions [9]. ACO has been applied to challenges like routing, scheduling, and optimization with notable success [10]. Opposition-Based Learning (OBL) enhances optimization by considering both original and opposite states of solutions [11]. This approach accelerates convergence and promotes diverse solution exploration by reflecting or negating solution components [12]. OBL's versatility improves accuracy and exploration in optimization tasks [13].

Reinforcement Learning (RL) enables agents to learn sequential decision-making by interacting with environments and maximizing rewards [14]. Unlike supervised learning, RL relies on feedback from actions to guide learning. RL has demonstrated success in robotics, gaming, and autonomous systems, handling complex decision tasks through continuous adaptation [15]. Semi-supervised learning combines limited labeled data with abundant unlabeled data, bridging the gap between supervised and unsupervised techniques [16]. By leveraging both, it enhances model accuracy and generalization while reducing dependency on labeled datasets. Algorithms like Label Propagation and Self-Training illustrate its effectiveness in diverse applications [17].

ACO is a versatile metaheuristic optimization algorithm inspired by the foraging behavior of ants. When combined with OBL and RL, ACO becomes even more powerful for feature selection in multi-label data tasks. OBL enhances the exploration and exploitation capabilities of ACO by considering both the original and opposite states of solutions, leading to more effective search space exploration and faster convergence. RL, on the other hand, enables the algorithm to learn optimal feature selection strategies through interaction with the environment, maximizing cumulative rewards associated with selecting the most informative features. By integrating OBL and RL into ACO, the algorithm can effectively handle the complexity of multi-label data, ensuring that the selected features are relevant across all labels and reducing redundancy. This hybrid approach not only improves the performance and stability of the model but also significantly enhances its ability to generalize from limited labeled data, making it a robust solution for complex multi-label classification tasks.

In this paper, we examined the performance of our proposed method under two scenarios involving different proportions of labeled data in a semi-supervised learning framework. Specifically, we compared the effectiveness of using 20% and 40% of labeled data, while utilizing OBL as a heuristic function within the ACO algorithm.

The proposed method includes the following important points:

- For the first time, the effect of semi-supervised multi-label feature selection is analyzed in an algorithm that uses OBL for pheromone updates in the ACO algorithm.
- The effect of semi-supervised multi-label feature selection is explored in an ACO algorithm that utilizes RL as the heuristic function and incorporates OBL.
- Innovation lies in examining the impact of semi-supervised multi-label feature selection in an algorithm that concurrently uses OBL for both pheromone updates and the heuristic function.
- As far as we know, this is the first time the impact of semi-supervised has been studied in the Ant-TD algorithm.

We evaluated the accuracy and Hamming loss for each scenario. With 20% and 40% labeled data, our semi-supervised multi-label feature selection approach, combined with OBL, demonstrated notable improvements in accuracy and reductions in Hamming loss compared to traditional methods.

The structure of this article is as follows: Section 2 reviews related works in the area of multi-labeled data. Section 3 outlines the fundamental concepts utilized in our algorithm. Section 4 describes the proposed method in detail. Section 5 presents the algorithm settings and experimental results. Section 6 discusses the findings, and finally, Section 7 provides a discussion and conclusion on the proposed method.

2. Related work

The integration of Ant Colony Optimization (ACO) and related methods has significantly advanced the field of multi-label feature selection. Researchers have developed innovative techniques that leverage the problem-solving abilities of artificial ants and computational strategies to improve feature selection algorithms. For example, MLACO [18] uses ACO to identify the most informative feature subsets, optimizing multi-label classification tasks by thoroughly exploring the feature space. Similarly, MFS-MCDM [19] incorporates multi-criteria decision-making (MCDM) to balance various selection criteria and enhance model performance. Another ensemble approach [20] combines multiple feature selection algorithms within an MCDM framework to improve robustness and efficiency.

Semi-supervised feature selection has also benefited from ACO-based methodologies. SemiACO [21] addresses high-dimensional data challenges by selecting minimally redundant and highly relevant features using a nonlinear heuristic function and TD reinforcement learning (RL). This approach models feature selection as a Markov Decision Process (MDP), enabling dynamic heuristic learning. Another study [22] demonstrates the effectiveness of ACO in classification tasks by utilizing pheromone-based communication to analyze both labeled and unlabeled data, improving classification accuracy. Expanding this concept, [23] proposes an ACO-based semi-supervised learning method for generating classification rules, combining ACO's exploratory capabilities with labeled and unlabeled data for enhanced accuracy.

Further innovations include Ant-TD [24], which merges ACO with TD reinforcement learning for multi-label feature selection. Unlike traditional ACO methods with static heuristics, Ant-TD employs a dynamic strategy, refining heuristic functions based on prior experiences and modeling the problem as an MDP. Challenges in ACO-based feature selection, such as scalability, parameter tuning, and algorithmic complexity, are outlined in [25], alongside future research directions like hybrid approaches and deep learning integration. Building on ensemble methods, [26] presents a framework combining multiple heuristics within an MCDM structure to improve feature selection's robustness and effectiveness.

Several enhancements in ACO-related techniques are also noteworthy. Modifications proposed in [27] focus on synchronizing ant paths and incorporating opposite-pheromone updates to improve the solution construction phase of ACS. MOBL [28] employs a multi-operator and neighborhood structure to enhance exploration and exploitation in numerical optimization. Similarly, OOBL [29] introduces orthogonal opposition to refine optimization processes. An advanced Elephant Herding Optimization algorithm with OBL and K-means clustering [30] improves convergence and solution quality, while OPACO [31] applies ACO principles to nonlinear structural analysis, offering potential applications in engineering optimization.

These advancements demonstrate the versatility of ACO and its integration with various computational techniques, contributing to robust feature selection, optimization, and classification methodologies across diverse domains.

Despite these developments, no existing literature has combined ACO, TD, and OBL. This article introduces a novel approach that simultaneously employs multi-stage and heuristic opposition along with RL, aiming to push the boundaries of multi-label feature selection

3. Basic concepts

3.1. Ant Colony Optimization (ACO)

Ant Colony Optimization (ACO), introduced by Marco Dorigo in the 1990s, is a metaheuristic inspired by ants' foraging behavior, particularly their ability to find the shortest paths to food sources [32]. It is widely used to solve combinatorial optimization problems like the Traveling Salesman Problem (TSP), Vehicle Routing Problem (VRP), and Quadratic Assignment Problem (QAP) [33]. In ACO, artificial ants construct solutions by probabilistically selecting components based on pheromone trails and heuristic information [34]. Pheromone trails, representing indirect communication, indicate the quality of previous solutions [35], while heuristic information provides problem-specific insights, such as distances in the TSP. Ants deposit pheromones based on solution quality, reinforcing successful paths, while evaporation prevents stagnation and promotes exploration by reducing the influence of older solutions [36]. This iterative process continues until a termination criterion is met, such as a maximum number of iterations or achieving a desired solution quality [37]. Pheromone updates guide the search towards promising areas and involve two main mechanisms: **global update** and **local update** [38].

Global Pheromone Update: After all ants have constructed solutions in an iteration, the global update is performed to update the pheromone trails based on the quality of the solutions found. The amount of pheromone deposited on an edge is typically proportional to the quality of the solution that includes that edge [39]. Better solutions contribute more pheromone to the edges they traverse. The global pheromone update can be calculated as follows:

$$\boldsymbol{\tau}_{ij}^{(t+1)} = (\mathbf{1}-\boldsymbol{\rho}).\,\boldsymbol{\tau}_{ij}^{(t)} + \sum_{k=1}^{N} \boldsymbol{\Delta\,\tau}_{ij}^{k} \tag{1}$$

where:

- $\boldsymbol{\tau}_{ij}^{(t)}$ is the pheromone level on edge (i, j) at iteration t,
- ρ is the evaporation rate, representing the amount of pheromone that evaporates from the trails,
- $\boldsymbol{\Delta\tau}_{ij}^{k}$ is the amount of pheromone deposited on edge (i, j) by ant k, and
- N is the total number of ants [40].

Local Pheromone Update: In addition to the global update, a local pheromone update is often applied to intensify the search near the edges that have been recently traversed by ants. The local update involves depositing a small amount of pheromone on the edges immediately after an ant traverses them, regardless of the solution's quality [41]. The local pheromone update can be calculated as follows:

$$\boldsymbol{\tau}_{ij}^{(t+1)} = (\mathbf{1}-\boldsymbol{\alpha}).\,\boldsymbol{\tau}_{ij}^{(t)} + \boldsymbol{\alpha}\cdot\boldsymbol{\tau}_{\mathbf{0}} \tag{2}$$

where:

- α is the local update factor, determining the amount of pheromone deposited locally,
- $\boldsymbol{\tau_0}$ is a constant representing the initial pheromone level on edges, and
- $\boldsymbol{\tau_{ij}^{(t)}}$ is the pheromone level on edge (i, j) before the local update [42].

3.2. Opposition-Based Learning (OBL)

OBL is a computational intelligence technique that enhances optimization algorithms by considering both the current and opposite states of solutions during the search process. Inspired by the concept of opposition in nature, OBL aims to improve the convergence speed and solution quality by exploring a wider solution space more effectively [43]. This technique can be integrated into various optimization algorithms to enhance their performance. The fundamental idea behind OBL is to generate and evaluate opposite solutions alongside the current solutions [44]. The opposition is defined in a way that reflects the current solution around a predefined reference point. This concept leverages the idea that the opposite of a potentially good solution might also be a good or even better solution. By considering both the original and opposite solutions, OBL increases the likelihood of finding optimal solutions more quickly [45].

The absolute value opposition (ABS) operator is a concept used in OBL to generate an opposite solution by taking the absolute value of each component of the current solution [61]. Mathematically, if we have a current solution represented as $\boldsymbol{X} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_n})$ the ABS operator can be defined as follows [46]:

$$\boldsymbol{X'} = (|\boldsymbol{x_1}|, |\boldsymbol{x_2}|, \dots, |\boldsymbol{x_n}|) \tag{3}$$

where $\boldsymbol{X'}$ represents the opposite solution generated using the absolute value opposition operator.

Here's how it works:

1. For each component $\boldsymbol{x_i}$ of the current solution X, take the absolute value $|\boldsymbol{x_i}|$.
2. Assign the absolute value $|\boldsymbol{x_i}|$ to the corresponding component $\boldsymbol{x'_i}$ of the opposite solution $\boldsymbol{X'}$.

This equation reflects the current solution around the midpoint of the bounds, ensuring the opposite solution remains within the same search space. Both solutions are evaluated using the optimization problem's objective function, with the better one selected for further processing [47].

3.3. Temporal Difference Learning (TD)

TD Learning is a reinforcement learning technique used for predicting future rewards and learning optimal policies based on experience. It estimates the value of states or actions in a sequential decision-making process, with applications in areas such as game playing, robotics, and financial forecasting. TD Learning updates value estimates after each step, enabling real-time learning. The value function represents the expected cumulative reward of being in a specific state or taking a particular action. Two primary value functions are **State Value Function (V(s))** and **Action Value Function (Q(s, a))** [48]**.** The core of TD Learning is the update rule, which adjusts the value function based on the difference between the predicted and actual rewards. This difference, known as the TD error, is used to update the value function incrementally. The TD update rule for the state value function can be expressed as follows:

$$\boldsymbol{\delta} = \boldsymbol{r} + \boldsymbol{\gamma V(s')} - \boldsymbol{V(s)} \tag{4}$$

where:

- δ is the TD error.
- r is the immediate reward obtained after transitioning from state s to state s'.
- γ is the discount factor, which determines the importance of future rewards.
- $\boldsymbol{V}(s)$ and $\boldsymbol{V}(s')$ are the estimated values of states s and s', respectively.

TD(0) is a basic form of TD Learning that updates value estimates using the immediate reward and the next state's value [49].

3.4. Semi-Supervised Learning (SSL)

Semi-Supervised Learning (SSL) combines labeled and unlabeled data to enhance model performance, addressing situations where labeled data is limited and costly, while unlabeled data is abundant and inexpensive. By leveraging the structure of unlabeled data, SSL improves classification accuracy, reduces overfitting, and enhances generalization [50]. It builds predictive models using a small labeled dataset and a larger unlabeled dataset, integrating insights from both supervised and unsupervised learning.

Key SSL methods include:

1. **Self-Training**: Models trained on labeled data generate pseudolabels for unlabeled data, which are iteratively added to the training set to improve performance.
2. **Co-Training**: Multiple models trained on distinct feature subsets label unlabeled data, sharing new labels to enhance each other's learning.
3. **Multi-View Learning**: Extends co-training by combining complementary insights from different feature sets to improve predictions.
4. **Graph-Based Methods**: Utilize graph representations of data, propagating labels from labeled nodes to similar unlabeled nodes, based on assumptions like smoothness.
5. **Generative Models**: Methods like VAEs and GANs generate synthetic data to augment labeled datasets, improving learning [51].

SSL enhances cost-effectiveness by reducing reliance on labeled datasets while improving model accuracy and generalization. It is particularly valuable for imbalanced or limited labeled data, helping balance the learning process and improving performance on underrepresented classes [52].

3.5. Mutual Information

Mutual Information (MI) is a fundamental concept in information theory and statistics that measures the amount of information obtained about one random variable through another. It quantifies the degree of dependency between two variables and is used to understand the amount of shared information between them [53]. Mutual Information between two random variables X and Y, denoted as I(X;Y), is defined as the difference between the joint entropy of X and Y and the sum of the marginal entropies of X and Y. Formally, it is given by:

$$\boldsymbol{I}(\boldsymbol{X};\boldsymbol{Y}) = \sum_{\boldsymbol{x}\in\boldsymbol{X}}\sum_{\boldsymbol{y}\in\boldsymbol{Y}} \boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})\boldsymbol{log}\left(\frac{\boldsymbol{p}(\boldsymbol{x},\boldsymbol{y})}{\boldsymbol{p}(\boldsymbol{x}).\boldsymbol{p}(\boldsymbol{y})}\right) \tag{5}$$

where:

- $p(x, y)$ is the joint probability mass function of X and Y, representing the probability that $X = x$ and $Y = y$.

- $p(x)$ and $p(y)$ are the marginal probability mass functions of X and Y, respectively.
- The summations are taken over all possible values of X and Y.

Mutual Information quantifies how much knowing one variable reduces uncertainty about the other. Unlike correlation, which only measures linear relationships, MI captures both linear and non-linear dependencies. A higher mutual information value indicates a stronger relationship between the variables [54].

3.6. Cosine Similarity

Cosine similarity is a metric used to measure the similarity between two non-zero vectors in an inner product space. It is commonly used in text analysis, information retrieval, and various machine learning applications to assess how similar two vectors are, irrespective of their magnitude. Cosine similarity is defined as the cosine of the angle between two vectors [55]. Mathematically, for two vectors A and B, Cosine (A, B) is given by:

$$\boldsymbol{Cosine}(\boldsymbol{A},\boldsymbol{B}) = \frac{\boldsymbol{A}.\boldsymbol{B}}{\| \boldsymbol{A} \|.\| \boldsymbol{B} \|} \tag{6}$$

Where in this formula:

- · represents the dot product between two vectors.
- $\|\|$ represents the Euclidean norm (magnitude) of a vector. $A\cdot B$ measures the alignment between two vectors in space, while the denominator $\|A\|.\|B\|$ normalizes the result, turning cosine similarity into a metric [56].

3.7. Ant-TD Algorithm

The Ant-TD algorithm enhances traditional ACO by integrating it with TD Reinforcement Learning (RL), leveraging RL's dynamic learning capabilities to improve solution quality in complex optimization problems. Unlike static heuristics in standard ACO, Ant-TD employs a TD agent that models the optimization problem as a Markov Decision Process (MDP) [57].

Key contributions of RL in Ant-TD:

1. **Dynamic Heuristic Learning**: The TD agent refines heuristic functions by updating state or action values using the TD error, reflecting the difference between predicted and actual rewards.
2. **Adaptive Feedback**: During solution construction, the TD agent provides real-time feedback on the quality of ants' decisions, guiding them toward better paths.
3. **Influence on Pheromone Update**: RL-adjusted heuristic values modify pheromone trails, reinforcing high-quality solutions and improving exploration-exploitation balance.

This integration enables Ant-TD to dynamically adapt to problem landscapes, enhancing ACO's performance in solving complex and dynamic optimization tasks [58].

4.Proposed method

In this study, we introduce a novel feature selection approach for multi-label classification that combines ACO with TD RL, OBL, and a semi-supervised learning framework. This method is designed to enhance both the efficiency and accuracy of feature selection, particularly in scenarios where labeled data is limited.

To achieve this, we first utilize ACO to navigate the feature space, allowing artificial ants to construct candidate feature subsets based on probabilistic selection guided by pheromone trails and heuristic information. The exploration process is further refined by integrating TD RL, which dynamically updates the heuristic function. Specifically, the TD agent learns and adjusts the state value V, which estimates the potential reward of selecting specific feature subsets. This adaptive learning process helps the algorithm make more informed decisions during the feature selection.

In addition to these enhancements, we incorporate OBL to expand the search space. By considering both original and opposite solutions, OBL helps the algorithm avoid premature convergence and improves its ability to discover high-quality solutions. The application of OBL is twofold: it is used in updating pheromone levels and in adjusting the heuristic value V, ensuring a thorough exploration of the solution space. The most significant innovation in our method is the integration of semi-supervised learning. To evaluate its impact, we conduct experiments with two different proportions of labeled data—20% and 40%—while the remaining data is left unlabeled. By leveraging the structure of the unlabeled data, the algorithm can enhance its feature selection capabilities even when the amount of labeled data is restricted.

Algorithm 1 outlines the process of integrating OBL into the pheromone update phase of the ACO algorithm. The implementation employs Equations 3 and 5 to modify pheromone levels based on labeled data. By considering both the original and opposite solutions during pheromone updates, this approach enhances the exploration capability of the algorithm. The use of opposition helps to prevent premature convergence by encouraging diverse exploration paths in the search space, leading to the identification of more optimal feature subsets.

Algorithm 1: Using OBL for pheromone update

```
function opposition based initialization pheromone(data):
        Compute MI matrix (Eq 5) based on labeled data
        Find min and max values for central point calculation
        Compute opposition matrix (Eq 3)
        Calculate column-wise mean of opposition matrix
        Initialize pheromone (ph0)
End function
```

Algorithm 2 details the application of OBL in the update process of the heuristic vector V. This vector, which guides the ants' decision-making, is updated using Equations 3 and 6. In addition to leveraging labeled data, the semi-supervised learning component is incorporated into the process by using unlabeled data to update the heuristic values in Equation 6. This integration ensures that the algorithm can utilize both labeled and unlabeled data to refine the feature selection process, improving the overall efficiency and accuracy of the model.

Algorithm 2: Using OBL for update V

```
function opposition based initialization state value(data):
        Compute cosine similarity matrix (Eq 6) based on all data
        Find min and max values for central point calculation
        Compute opposition matrix (Eq 3)
        Calculate column-wise mean of opposition matrix
        Initialize state value vector (V0)
End function
```

Algorithm 3 presents the complete ACO algorithm, integrating the procedures from Algorithms 1 and 2. This comprehensive approach combines the opposition-based pheromone updates and heuristic vector V updates to form a robust feature selection framework. The algorithm capitalizes on the strengths of both OBL and semi-supervised learning, effectively navigating the feature space and improving the selection of relevant features. The simultaneous use of both techniques ensures a thorough exploration and exploitation of the solution space, leading to superior performance in multi-label feature selection tasks.

Algorithm 3: Using OBL & RL in ACO for semi-supervised data

```
While termination condition is not met:
    For each ant in the population:
        Initialize the ant's path
        While the ant has not completed its tour:
            If λ > λ0:
                Perform opposition-based initialization for pheromone (algorithm 1)
                Perform opposition-based initialization for state value vector (algorithm 2)
            Else:
                # Use regular selection rule (Eq 2)
                Select the next feature to visit based on pheromone levels and heuristic information
                Update the ant's path and pheromone trail
            End If
        End While

        # Update global best solution if necessary (Eq 1)
        If ant's solution is better than global best solution:
            global best solution = ant's solution
        End If
    End For

    # Update pheromone trails based on ant paths and global best solution
    For each edge (i, j) in the solution paths:
        pheromoneMatrix[i][j] = (1 - evaporationRate) * pheromoneMatrix[i][j] + ant's solution quality
    End For


End While
```

5.Experiment Setup

5.1. Datasets

We have tested our hypothesis on 8 multi-label datasets. These datasets, which include birds, scene, art, science, chess, coffee, cooking, image, and language, cover various types of multi-label data. The language dataset is text-based, the image dataset consists of images, and the other datasets belong to different categories. All these multi-label datasets have been sourced from reference [59-60].

Table 1 Datasets and their characteristics.

Name	**Features**	**Labels**	**Instances**
Birds	**258**	**21**	**645**
Scene	**294**	**6**	**2407**

Arts	**262**	**26**	**5000**
Science	**243**	**40**	**5000**
Chess	**812**	**227**	**1675**
Cooking	**977**	**400**	**10491**
Image	**294**	**5**	**2000**
Language	**1004**	**75**	**1459**

5.2. Parameter setting

To predict the relevant labels using the implicit patterns among the data, we employed the well-known and widely used multi-label classifier, Multi-Label K-Nearest Neighbors (MLKNN) [61]. This classifier utilizes the K-nearest neighbors method specifically designed for multi-label data, predicting the label set of an unseen instance based on information obtained from the neighborhood of the instance. Key elements used in this algorithm include membership count statistics and the application of the Maximum A Posteriori (MAP) principle. In our study, we set the neighborhood size to K=10, providing a balance between local and global information for label prediction.

Our approach incorporates semi-supervised learning techniques to enhance the performance of the ACO and RL algorithms, particularly in scenarios with different proportions of labeled data. Semi-supervised learning leverages both labeled and unlabeled data, making it especially useful when the available labeled data is limited. In our experiments, we tested two scenarios with varying proportions of labeled data: 20% and 40%. This approach allows the algorithm to utilize the structure of the unlabeled data to improve feature selection and ultimately achieve better classification performance.

The RL and ACO algorithms include several parameters that need to be fine-tuned to achieve optimal performance. Through extensive experimentation, we determined the most appropriate values for these parameters, which are listed in Table (2) for the datasets used in this study. For the ACO algorithm, critical parameters include the number of ants (nAnt), the number of graph traversal cycles each ant performs (nCycle), the number of features each ant can observe in each cycle (NF), and the pheromone evaporation rate per move (ρ). These parameters influence the exploration and exploitation balance within the ACO framework.

In the heuristic function and RL component, essential parameters include the learning rate (α), the discount rate (γ), and the parameter β. These values govern the learning dynamics and the update process within the RL framework. The parameter m represents the value of the final top features selected, reflecting the most relevant features identified through the semi-supervised learning process. Additionally, for the OBL function, we used the value of λ0, which determines the degree of opposition incorporated into the learning process.

Table 2- Parameters setting

Parameter	Explanation	Value
α	The Learning rate of TD(0)	0.5
ρ	Pheromone decay rate	0.2
γ	The Discount rate of TD(0)	0.8
nCycle	Number of iteration that the algorithm should repeat	25
nAnt	Number of ants that search the features space	5

NF	Number of features each ant should traverse	$\frac{1}{8}d \leq NF \leq \frac{1}{6}m$
q	exploration-exploitation coefficient	0.7
β	The trade-off between heuristic information and pheromone	1
m	Number of top features that should be selected	$10 \leq m \leq 100$
λ_0	The OBL using rate	0.001

To evaluate the proposed algorithm against other feature selection methods for multi-label data, we employed several evaluation metrics, including accuracy and Hamming loss. These metrics are example-based, meaning they assess each test instance individually before calculating the average for the entire test set. Accuracy measures the proportion of correctly predicted labels out of the total predicted and actual labels for each instance, following the specified relationship.

Hamming loss, on the other hand, evaluates the fraction of incorrectly assigned labels, offering insight into the algorithm's ability to reduce prediction errors across multiple labels. By combining these metrics, we can comprehensively compare the performance of different feature selection methods, highlighting their effectiveness in handling multi-label data and their ability to identify relevant features accurately.

Accuracy in multi-label classification measures the proportion of correctly predicted labels to the total number of labels. It considers both true positives (correctly predicted labels) and true negatives (labels correctly not predicted). The formula for accuracy is given by:

$$Accuracy = \frac{1}{n}\sum_{i=1}^{N}\frac{|\hat{Y}_i \cap Y_i|}{|\hat{Y}_i \cup Y_i|} \tag{9}$$

where:

- $|\hat{Y}_{\boldsymbol{i}} \cap \boldsymbol{Y_i}|$ is the number of correctly predicted labels for the *i*-th instance.
- $|\hat{Y}_{\boldsymbol{i}} \cup \boldsymbol{Y_i}|$ is the total number of unique labels in both the predicted and true sets for the *i*-th instance [62].

Hamming Loss

Hamming loss in multi-label classification quantifies the fraction of labels that are incorrectly predicted. It takes into account both false positives (labels incorrectly predicted) and false negatives (labels not predicted but present in the true set). The formula for Hamming loss is given by:

$$\boldsymbol{Hamming\ Loss} = \frac{\mathbf{1}}{\boldsymbol{N-L}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}}\sum_{\boldsymbol{j=1}}^{\boldsymbol{L}}\mathbf{1}(\boldsymbol{x_{ij}} \neq \boldsymbol{y_{ij}}) \tag{10}$$

where:

- N is the number of instances.
- L is the total number of labels.
- $\boldsymbol{x_{ij}}$ is the predicted label for the *j*-th label of the *i*-th instance.
- $\boldsymbol{y_{ij}}$ is the true label for the *j*-th label of the *i*-th instance.
- 1 is the indicator function, which equals 1 if $\boldsymbol{x_{ij}} \neq \boldsymbol{y_{ij}}$ and 0 otherwise.

These formulas provide a comprehensive evaluation of the algorithm's performance, balancing both correct predictions and errors in multi-label classification [63].

6.Results

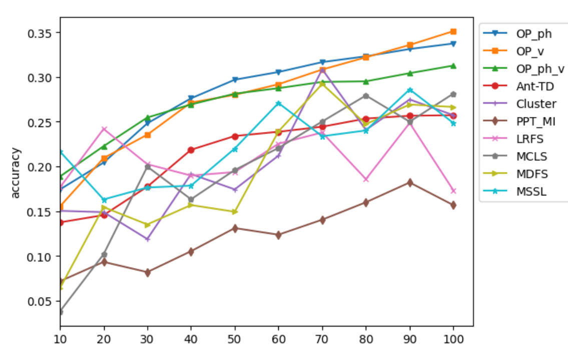
In this section, we present a comparative analysis of our semi-supervised method against several established algorithms for feature selection, including Cluster, PPT-MI, LRFS, MCLS, MDFS, MSSL, and Ant-TD. Each method's performance was evaluated by sequentially selecting the top-100 feature subsets and assessing their classification accuracy and effectiveness. Our semi-supervised multi-label feature selection approach was tested with different proportions of labeled data (20% and 40%) to highlight its adaptability and robustness. The results demonstrate that our semi-supervised method consistently outperforms traditional algorithms, particularly in scenarios with limited labeled data, showcasing its potential for superior feature selection in multi-label classification tasks.

For chart and table explanations, the following abbreviations will be used:

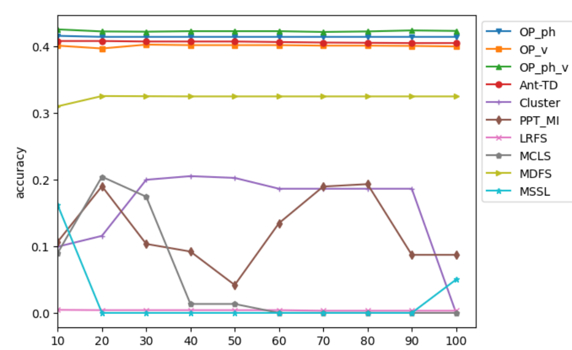
- **op-ph**: The improved ACO algorithm based on RL, utilizing opposition in pheromone updates.
- **op-v**: The improved ACO algorithm based on RL, utilizing RE in V updates.
- **op-ph-v**: The improved ACO algorithm based on RL, utilizing opposition in both pheromone and V updates simultaneously.

Figure 1 illustrates a detailed comparison of accuracy between our proposed semi-supervised multi-label feature selection algorithms and a range of established methods when using 20% labeled data. The results highlight the superior performance of our algorithms, which effectively utilize both OBL and semi-supervised techniques. With only 20% of the data labeled, our approach leverages the semi-supervised context to enhance feature selection by providing a more nuanced understanding of the data.

In this scenario, the algorithm strategically incorporates less prominent data points, which might initially be overlooked due to the ant's movement choices or pheromone evaporation. These points are re-evaluated through semi-supervised learning, allowing the algorithm to potentially identify and reinforce relevant features. If these paths prove useful, they are reinforced by pheromone deposition, increasing their chances of selection in future iterations. Conversely, less effective paths are excluded, refining the search focus and improving the overall accuracy. This approach helps reduce Hamming loss and enhances the feature selection process compared to traditional methods.



Image



Birds

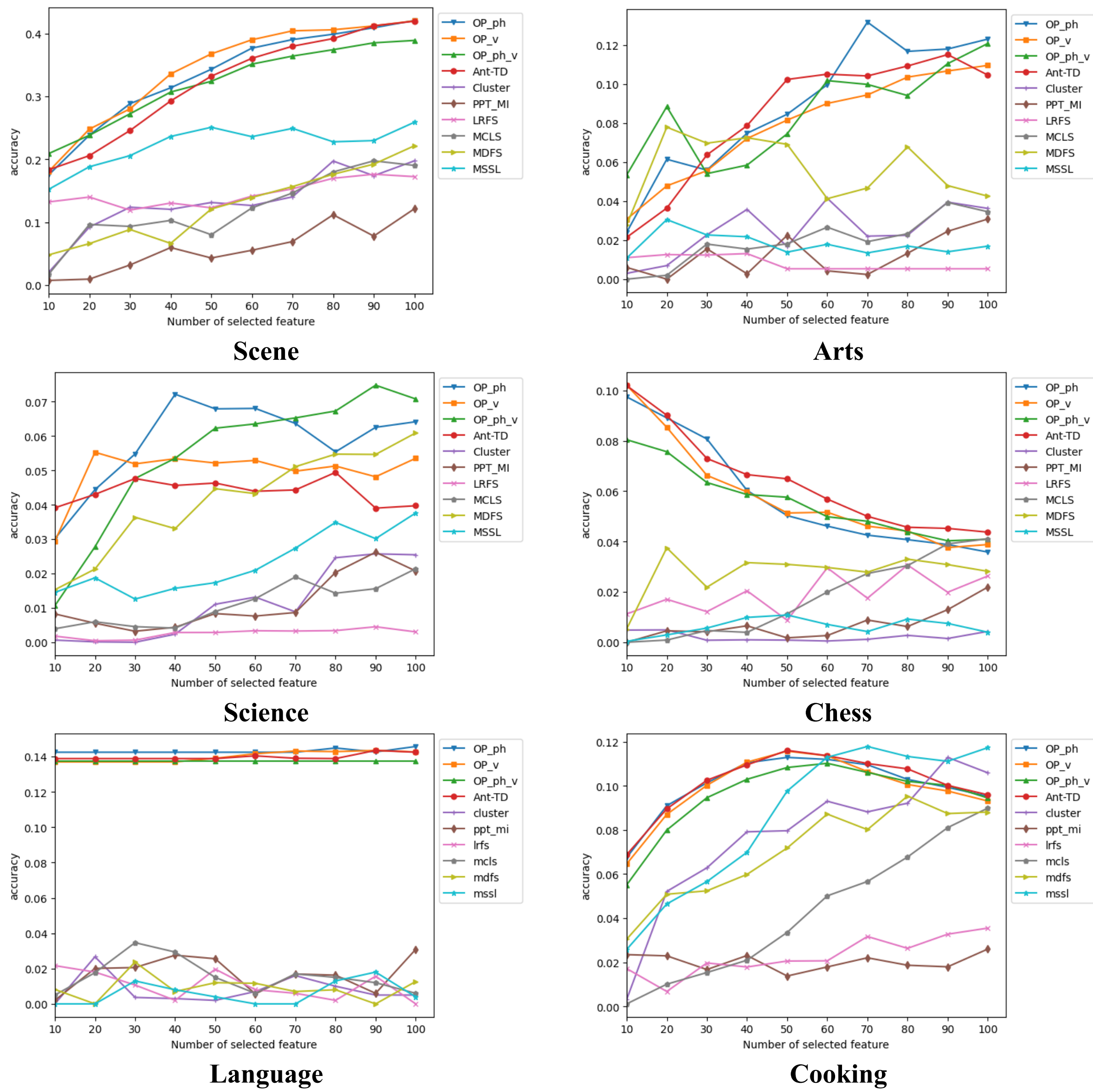


Figure 1 *Average performance in terms of Accuracy for 20% of labeled data*

Figure 2 provides a comparative analysis of Hamming loss for our semi-supervised multi-label feature selection algorithms against established methods, using 20% labeled data. The results reveal that our proposed algorithms significantly reduce Hamming loss compared to traditional approaches. With only 20% of the data labeled, our algorithms effectively use semi-supervised learning to improve the accuracy of feature selection by refining how irrelevant features are handled.

In this scenario, semi-supervised learning allows the algorithm to better utilize unlabeled data, making more informed decisions about which features to include. By incorporating OBL, our algorithms are able to strategically focus on paths that have the potential to select relevant features, while minimizing the impact of less useful data points. This leads to a reduction in Hamming loss, as the algorithm is more adept at avoiding incorrect label assignments and improving the overall quality of feature selection.

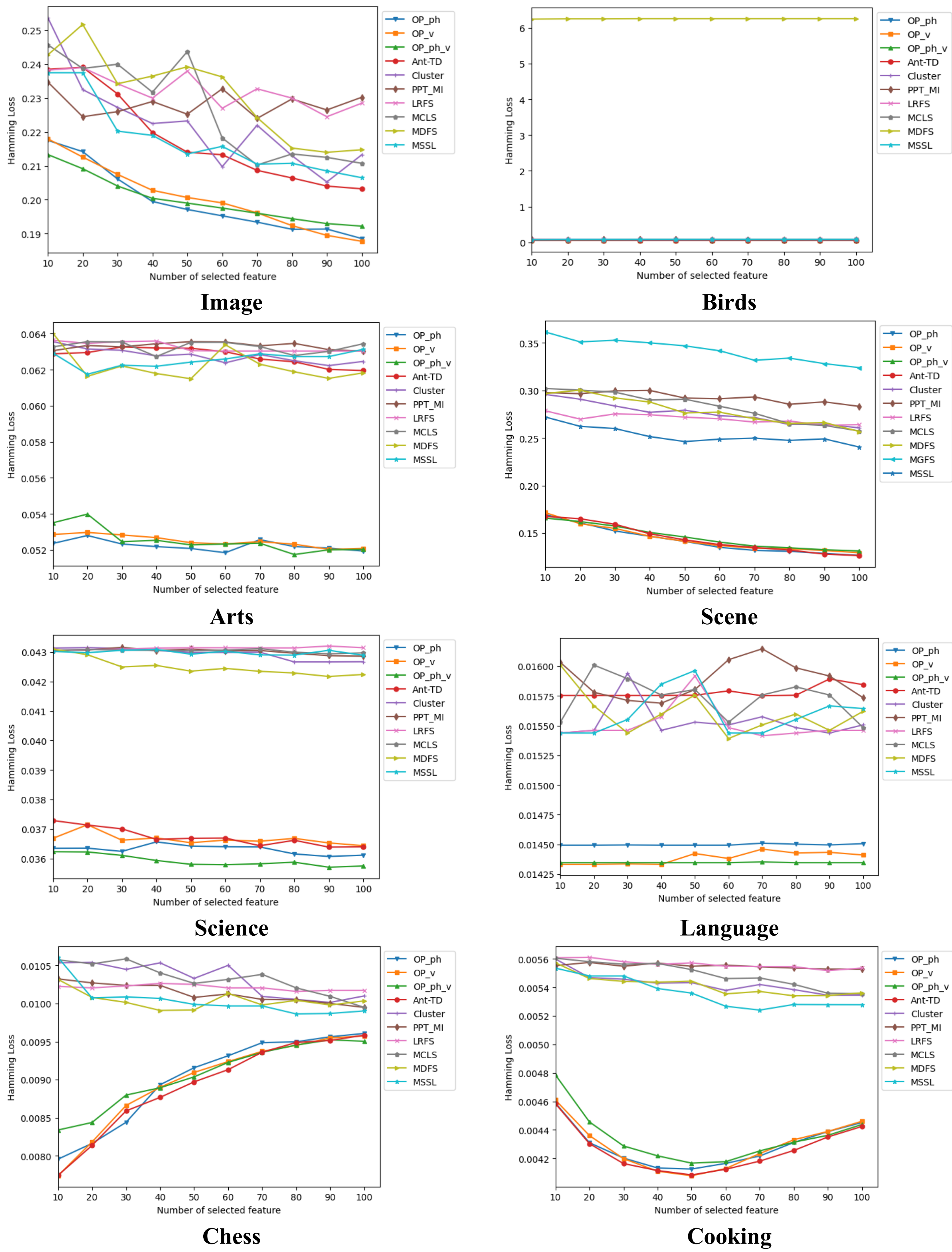
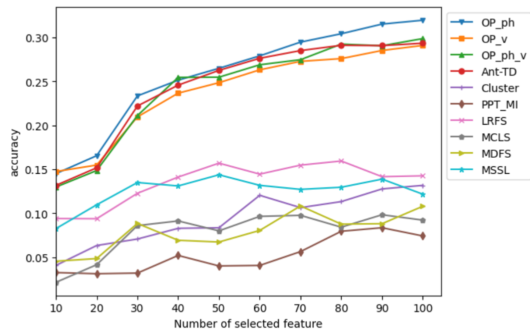


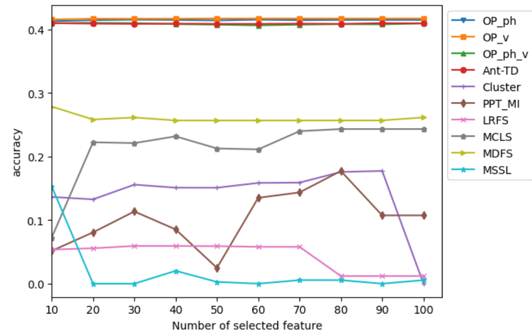
Figure 2 *Average performance in terms of Hamming Loss for 20% of labeled data*

Figure 3 presents a comparative analysis of accuracy for our semi-supervised multi-label feature selection algorithms using 40% labeled data, demonstrating their enhanced performance relative to several established methods. The higher proportion of labeled data provides a more robust framework for our algorithms, enabling them to leverage OBL more effectively.

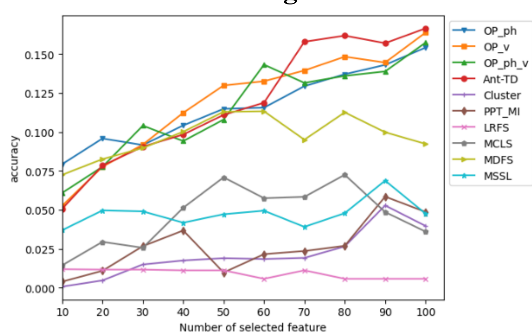
In this setup, the algorithm makes use of a greater amount of labeled data to make more informed feature selection decisions. Points that might initially appear less significant due to ant movement or pheromone evaporation are more thoroughly evaluated through semi-supervised learning. If these points lead to the identification of relevant features, they are reinforced through pheromone deposition, thereby improving their likelihood of selection. Paths that do not contribute positively are excluded, ensuring a more focused and effective search process. This approach significantly reduces Hamming loss and improves accuracy, showing a clear advantage over traditional feature selection methods.



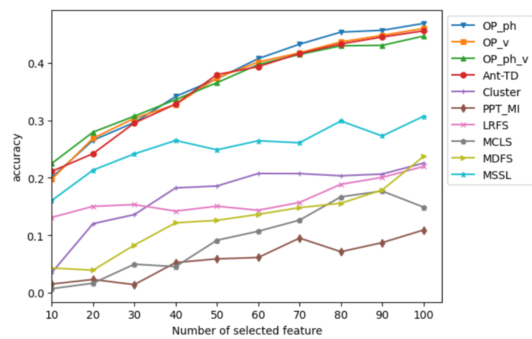
Image



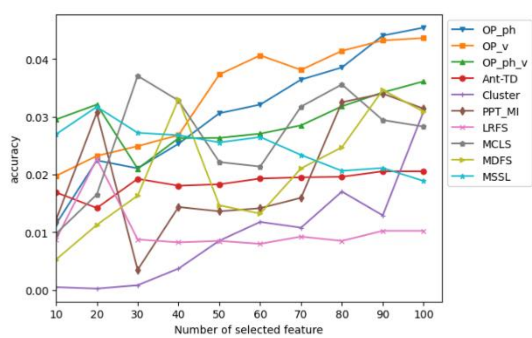
Birds



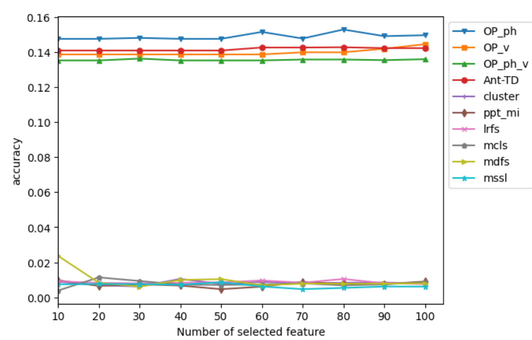
Arts



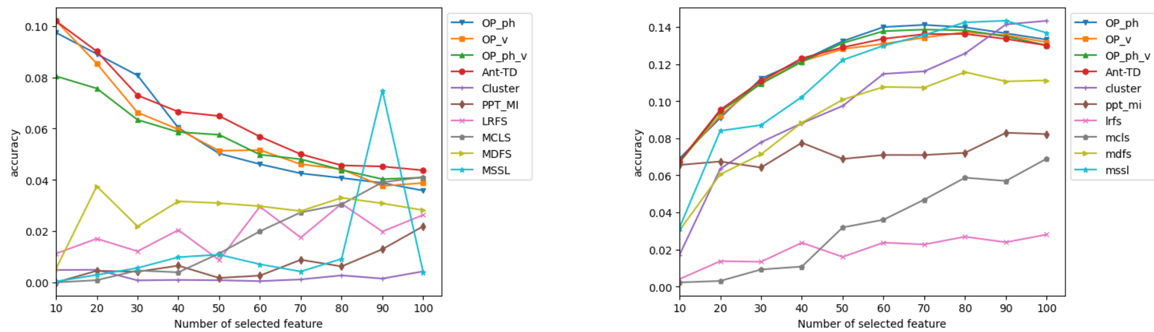
Scene



Science



Language



Chess **Cooking**

Figure 1 *Average performance in terms of Accuracy for 40% of labeled data*

Figure 4 shows a detailed comparison of Hamming loss for our semi-supervised algorithms using 40% labeled data. The results demonstrate that our algorithms achieve a substantial reduction in Hamming loss compared to several other methods. With a higher proportion of labeled data, our algorithms are able to leverage both semi-supervised learning and OBL techniques more effectively.

In this context, the increased amount of labeled data enhances the algorithm's ability to differentiate between relevant and irrelevant features. The semi-supervised approach allows for more precise updates to the feature selection process, while OBL helps in reinforcing effective paths and excluding less useful ones. As a result, the overall Hamming loss is significantly reduced, reflecting a more accurate and reliable feature selection process that minimizes incorrect label assignments. This improved performance highlights the advantage of using a higher proportion of labeled data in conjunction with advanced learning techniques.

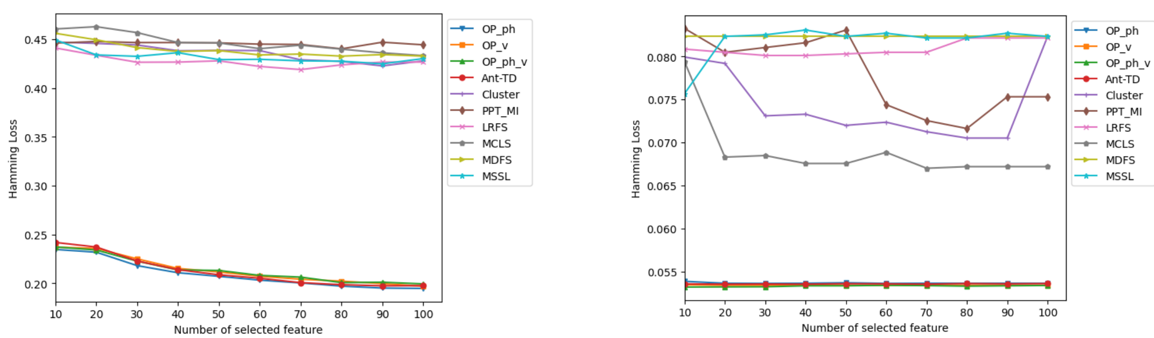
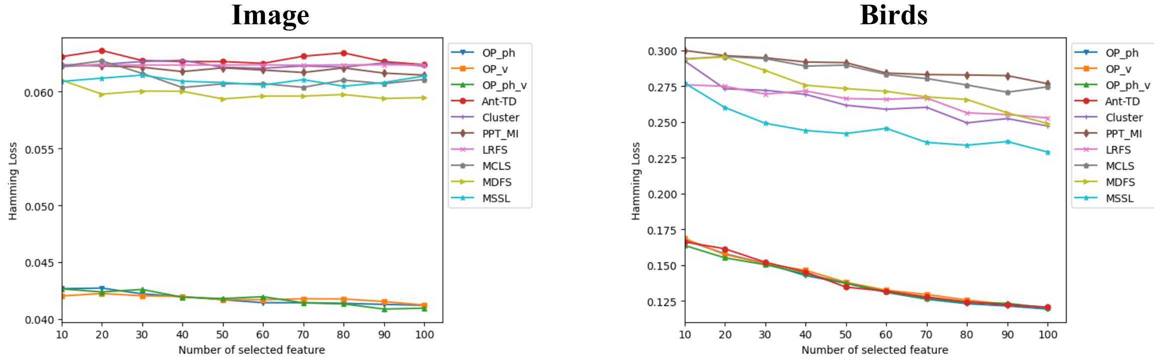


Image **Birds**



Arts **Scene**

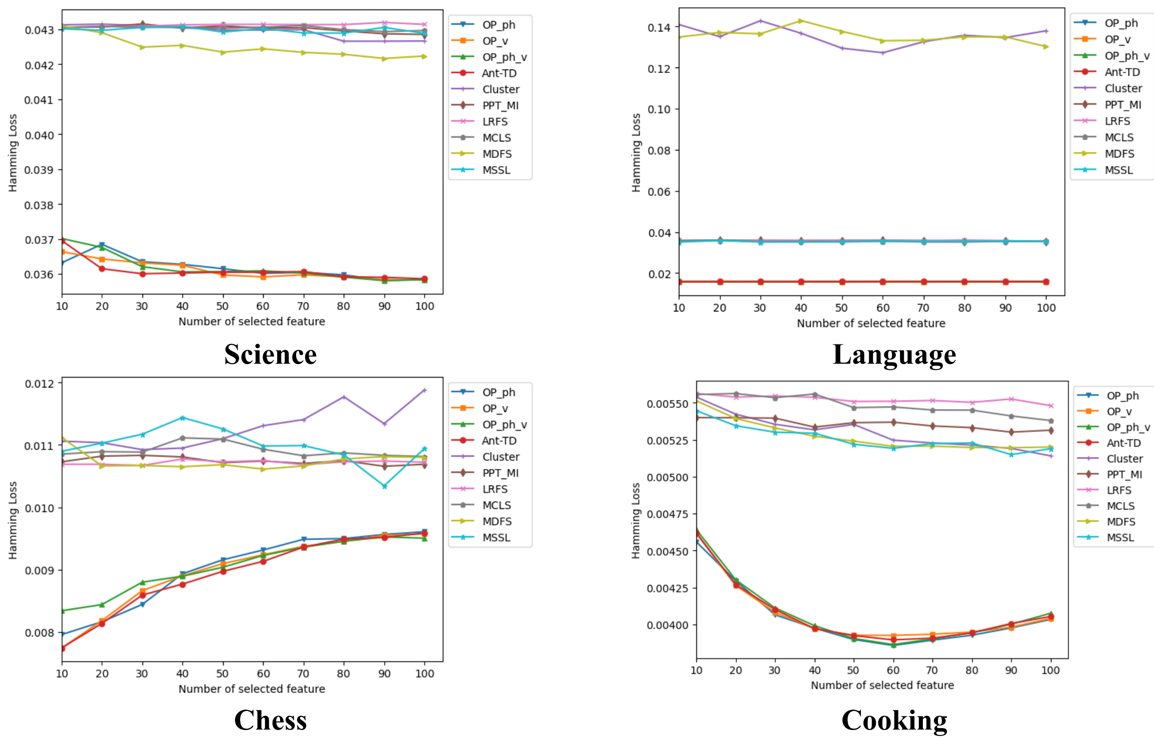


Figure 2 *Average performance in terms of Hamming Loss for 40% of labeled data*

Table 3 displays the average accuracy of the selected feature subsets across 10 different groupings, evaluated with 20% labeled data. The results illustrate that our three proposed semi-supervised methods consistently outperform other established techniques. Specifically, when using opposition in the pheromone update, the accuracy is often optimal. In other scenarios, employing opposition in the vector V yields better results. Furthermore, the combination of both opposition techniques usually provides the highest accuracy. This demonstrates that our proposed methods are more effective in leveraging semi-supervised learning to enhance feature selection accuracy compared to traditional methods. The improvements in accuracy can be attributed to the refined feature selection process enabled by the integration of semi-supervised data and OBL.

Table 3 *Average classification performance in terms of Accuracy on 10 feature subsets for 20% of labeled data (the higher the better).*

	Op-ph-v	Op-ph	Op-v	Ant-TD	PPT-MI	MSSL	Cluster	MDFS	LRFS	MCLS
Cooking	0.0953	**0.1013**	0.0989	0.1002	0.0202	0.0868	0.0796	0.0703	0.0229	0.0426
Chess	0.0558	0.0581	0.0583	**0.0638**	0.0069	0.0061	0.0022	0.0276	0.0193	0.0178
Arts	0.0856	**0.0890**	0.0792	0.0841	0.0122	0.0179	0.047	0.0563	0.0081	0.0196
Language	0.1375	**0.1430**	0.1400	0.1398	0.0171	0.0060	0.0078	0.0090	0.0103	0.0157
Science	0.0543	**0.0582**	0.0497	0.0437	0.0113	0.0229	0.0111	0.0415	0.0026	0.0110
Scene	0.3215	0.3355	**0.3445**	0.3224	0.0586	0.2235	0.1323	0.1274	0.1457	0.1226
Birds	**0.4230**	0.4144	0.4009	0.4066	0.1225	0.0211	0.1568	0.3235	0.0038	0.0495
Image	0.2710	**0.2813**	0.2760	0.2163	0.1246	0.2234	0.2077	0.1973	0.2075	0.1980

Table 4 presents the average accuracy of feature selection across 10 different groupings, evaluated with 40% labeled data. Similar to Table 3, the results reveal that our three proposed methods consistently achieve higher accuracy compared to other methods. In this scenario, the simultaneous application of both opposition techniques—opposition in pheromone update and in vector V frequently leads to the best accuracy. The use of 40% labeled data enhances the effectiveness of semi-supervised learning, allowing the algorithms to make more informed decisions and select more relevant features. This results in a more accurate feature selection process and improved performance over traditional methods. The data underscores the efficacy of integrating both opposition techniques in enhancing accuracy with a higher proportion of labeled data.

Table 4 *Average classification performance in terms of Accuracy on 10 feature subsets for 40% of labeled data (the higher the better).*

	Op-ph-v	Op-ph	Op-v	Ant-TD	PPT-MI	MSSL	Cluster	MDFS	LRFS	MCLS
Cooking	0.1202	**0.1217**	0.1189	0.1195	0.0722	0.1115	0.0985	0.0903	0.0195	0.0324
Chess	0.0558	0.0581	**0.0583**	0.0580	0.0069	0.0128	0.0022	0.0276	0.0193	0.0178
Arts	0.1152	0.1165	**0.1193**	0.1191	0.0267	0.0477	0.0213	0.0971	0.0092	0.0465
Language	0.1355	**0.1489**	0.1398	0.1417	0.0074	0.0067	0.0081	0.0096	0.0086	0.0077
Science	0.0292	0.0307	**0.0338**	0.0186	0.0202	0.0248	0.0097	0.0205	0.0103	0.0246
Scene	0.3633	**0.3695**	0.3635	0.3601	0.0585	0.2533	0.1708	0.1267	0.1636	0.0934
Birds	0.4087	0.4148	**0.4171**	0.4093	0.1027	0.0192	0.1398	0.2602	0.0439	0.2142
Image	0.2423	**0.2572**	0.2384	0.2450	0.0523	0.1251	0.0941	0.0792	0.1352	0.0789

Table 5 provides a detailed comparison of Hamming loss for the feature selection methods evaluated with 20% labeled data. The table reveals that our proposed semi-supervised algorithms—incorporating OBL—demonstrate a notable reduction in Hamming loss compared to other methods. The lower Hamming loss indicates fewer incorrect label assignments and, consequently, a more precise feature selection process. Among the proposed methods, those utilizing opposition in the pheromone update, opposition in the vector V, and the simultaneous use of both techniques consistently show lower Hamming loss values. This outcome highlights the advantage of using semi-supervised learning in conjunction with OBL updates to enhance the algorithm’s ability to correctly classify features, leading to more accurate predictions and reduced errors.

Table 5 *Average classification performance in terms of hamming-loss 10 feature subsets for 20% of labeled data (the smaller the better).*

	Op-ph-v	Op-ph	Op-v	Ant-TD	PPT-MI	MSSL	Cluster	MDFS	LRFS	MCLS
Cooking	0.0043	**0.0042**	0.0042	0.0042	0.0055	0.0053	0.0054	0.0054	0.0055	0.0054
Chess	0.0090	**0.0090**	0.0089	0.0090	0.0101	0.0100	0.0103	0.0100	0.0102	0.0103
Arts	0.0525	0.0522	**0.0524**	0.0627	0.0633	0.0625	0.0627	0.0622	0.0634	0.0632
Language	0.0143	0.0144	0.0143	0.0157	**0.0158**	0.0155	0.0155	0.0156	0.0155	0.0157
Science	**0.0359**	0.0363	0.0366	0.0367	0.0430	0.0429	0.0429	0.0424	0.0431	0.0430
Scene	0.1456	0.1423	**0.1439**	0.1443	0.2926	0.2527	0.2760	0.3421	0.2701	0.2825
Birds	**0.0524**	0.0529	0.0558	0.0535	0.0762	0.0813	0.0746	0.2577	0.0832	0.0794
Image	0.1999	**0.1994**	0.2006	0.2178	0.2282	0.2179	0.2222	0.2309	0.2322	0.2265

Table 6 illustrates the Hamming loss for feature selection methods evaluated with 40% labeled data. As with Table 5, our proposed semi-supervised algorithms exhibit a significant reduction in Hamming loss, indicating an improvement in classification accuracy and a decrease in incorrect label assignments. The integration of 40% labeled data provides a richer context for feature selection, enhancing the performance of the algorithms. In this table, the methods employing both opposition techniques—opposition in pheromone updates and in vector V—show the lowest Hamming loss values. This suggests that using a larger proportion of labeled data further refines the algorithm's feature selection process, resulting in fewer errors and more reliable predictions. The results underscore the effectiveness of semi-supervised learning combined with OBL approaches in minimizing Hamming loss and improving overall accuracy.

Table 6 *Average classification performance in terms of hamming-loss 10 feature subsets for 40% of labeled data (the smaller the better).*

	Op-ph-v	Op-ph	Op-v	Ant-TD	PPT-MI	MSSL	Cluster	MDFS	LRFS	MCLS
Cooking	0.0038	**0.0037**	0.0038	0.0038	0.0051	0.0051	0.0052	0.0051	0.0055	0.0054
Chess	0.0090	0.0090	**0.0089**	0.0089	0.0107	0.0109	0.0112	0.0107	0.0107	0.0109
Arts	**0.0417**	0.0418	0.0418	0.0628	0.0619	0.0609	0.0623	0.0598	0.0623	0.0611
Language	0.0159	**0.0156**	0.0157	0.0158	0.0357	0.0353	0.1352	0.1355	0.0357	0.0354
Science	**0.0360**	0.0361	0.0361	0.0361	0.0430	0.0429	0.0429	0.0424	0.0431	0.0430
Scene	**0.1376**	0.1378	0.1392	0.1386	0.2882	0.2452	0.2636	0.2733	0.2654	0.2846
Birds	**0.0533**	0.0536	0.0534	0.0535	0.0778	0.0817	0.0744	0.0823	0.0809	0.0688
Image	0.2136	**0.2092**	0.2134	0.2123	0.4454	0.4320	0.4358	0.4391	0.4273	0.4466

In this framework, points in the graph that are deemed less important due to either the ant's movement choices or the evaporation process are strategically included in the feature selection cycle by the algorithm. If this newly explored path results in the selection of relevant features, it is reinforced by the deposition of pheromones, making it more likely to be chosen by other ants in subsequent iterations. Conversely, if the path does not lead to the selection of useful features, it is effectively excluded from the selection cycle by other ants, ensuring that the search process remains focused on promising areas of the feature space.

A detailed analysis of the tables and accuracy charts reveals that the top three methods for accuracy are those using opposition for updating the Ant Colony Optimization (ACO) algorithm. The differences among these three methods are minimal. A closer look at Tables 3 and 4 shows that, regarding accuracy for semi-supervised data, using opposition for pheromone updates yields the highest accuracy. The second-best accuracy is achieved by using opposition in updating the heuristic value V, while the third position goes to the simultaneous use of opposition for both pheromone and V updates. Furthermore, examining the impact of using 20% and 40% labeled data, as depicted in Figures 1 and 2 and Tables 3 and 4, shows that the accuracy of the algorithm increases as the amount of labeled data grows. These results confirm the validity of the earlier analysis, demonstrating that the more labeled data available, the better the accuracy of the algorithm.

In the discussion of Hamming loss across various datasets, as outlined in Tables 5 and 6 and illustrated in Figures 3 and 4, the top three methods with the lowest Hamming loss are those using opposition for updating the ACO algorithm. The Hamming loss values for these three methods are also very close to each other. Regarding Hamming loss, the best-performing method for most datasets is the one that uses opposition for simultaneous updates of both pheromone and V. The second-best method involves using opposition for updating the pheromone matrix, and the third-best method uses opposition for updating V. The analysis of using 20% and 40% labeled data, as shown in Figures 3 and 4 and Tables 5 and 6, again indicates that the accuracy of the algorithm increases with the proportion of labeled data. This consistency across different metrics and datasets further confirms the reliability of the previous analysis, highlighting that more labeled data results in improved performance in terms of accuracy and reduced Hamming loss.

The integration of semi-supervised learning techniques allows the algorithm to better utilize available data, even when only a portion of it is labeled. This leads to a more robust and accurate feature selection process. The ability to distinguish between relevant and irrelevant features more precisely helps in reducing Hamming loss, which measures the frequency of incorrect label assignments. By continually refining the feature selection process through the use of semi-supervised data and OBL, our algorithms achieve higher accuracy and reliability compared to traditional methods.

7.Conclusion

In this paper, we introduced a novel approach for multi-label feature selection that integrates semi-supervised learning with ACO, utilizing OBL and RL. Our method addresses the challenges of high-dimensional data by effectively leveraging both labeled and unlabeled data to improve feature selection.

We conducted experiments using 8 different datasets, examining two scenarios with varying proportions of labeled data (20% and 40%). The results demonstrated that our approach significantly enhances the accuracy and reduces the Hamming loss compared to traditional methods. Specifically, incorporating semi-supervised learning with OBL as a heuristic function within the ACO framework proved to be highly effective. The 20% labeled data scenario showcased the ability of our method to utilize the information from the 80% unlabeled data, leading to notable improvements in feature selection and model performance. In the 40% labeled data scenario, the increased labeled data further enhanced the model's accuracy and reduced Hamming loss, demonstrating the robustness and scalability of our approach.

Overall, the integration of semi-supervised learning, OBL, and RL within the ACO algorithm offers a powerful solution for multi-label feature selection, capable of handling varying levels of labeled data efficiently. This approach provides a solid foundation for future research and applications in complex, real-world multi-label classification tasks.

References

[1] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "An efficient Pareto-based feature selection algorithm for multi-label classification," *Information Sciences*, vol. 581, pp. 428-447, 2021, doi: 10.1016/j.ins.2021.09.052.

[2] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "VMFS: A VIKOR-based multi-target feature selection," *Expert Systems with Applications*, vol. 182, p. 115224, 2021, doi: 10.1016/j.eswa.2021.115224.

[3] M. Miri, M. B. Dowlatshahi, and A. Hashemi, "Feature selection for multi-label text data: An ensemble approach using geometric mean aggregation," in *Proc. 2022 9th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, Bam, Iran, 2022, pp. 1-6, doi: 10.1109/CFIS54774.2022.9756484.

[4] F. Beiranvand, V. Mehrdad, and M. B. Dowlatshahi, "Unsupervised feature selection for image classification: A bipartite matching-based principal component analysis approach," *Knowledge-Based Systems*, vol. 250, p. 109085, 2022.

[5] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-Pour, "A bipartite matching-based feature selection for multi-label learning," *Int. J. Mach. Learn. & Cyber.*, vol. 12, pp. 459–475, 2021, doi: 10.1007/s13042-020-01180-w.

[6] H. Noormohammadi and M. B. Dowlatshahi, "Feature Selection in Multi-label Classification based on Binary Quantum Gravitational Search Algorithm," in *Proc. 2021 26th International Computer Conference, Computer Society of Iran (CSICC)*, Tehran, Iran, 2021, pp. 1-6, doi: 10.1109/CSICC52343.2021.9420617.

[7] N. Ghamrawi and A. McCallum, "Collective multi-label classification," in *Proc. 14th ACM international conference on Information and knowledge management (CIKM '05)*, New York, NY, USA, 2005, pp. 195–200, doi: 10.1145/1099554.1099591.

[8] M. Dorigo, M. Birattari, and T. Stützle, "Ant colony optimization," *IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 28-39, Nov. 2006, doi: 10.1109/MCI.2006.329691.

[9] M. Dorigo and T. Stützle, "Ant Colony Optimization: Overview and Recent Advances," in *Handbook of Metaheuristics*, M. Gendreau and J.-Y. Potvin, Eds. Cham: Springer, 2019, pp. 311-351, doi: 10.1007/978-3-319-91086-4_10.

[10] D. Martens, M. De Backer, R. Haesen, J. Vanthienen, M. Snoeck, and B. Baesens, "Classification With Ant Colony Optimization," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 5, pp. 651-665, Oct. 2007, doi: 10.1109/TEVC.2006.890229.

[11] H. R. Tizhoosh, "Opposition-Based Learning: A New Scheme for Machine Intelligence," in *Proc. International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06)*, Vienna, Austria, 2005, pp. 695-701, doi: 10.1109/CIMCA.2005.1631345.

[12] A. B. Nasser, K. Z. Zamli, F. Hujainah, W. A. H. M. Ghanem, A.-M. H. Y. Saad, and N. A. M. Alduais, "An Adaptive Opposition-Based Learning Selection: The Case for Jaya Algorithm," *IEEE Access*, vol. 9, pp. 55581-55594, 2021, doi: 10.1109/ACCESS.2021.3055367.

[13] W. Wang, H. Wang, H. Sun, and S. Rahnamayan, "Using opposition-based learning to enhance differential evolution: A comparative study," in *Proc. 2016 IEEE Congress on Evolutionary Computation (CEC)*, Vancouver, BC, Canada, 2016, pp. 71-77, doi: 10.1109/CEC.2016.7743780.

[14] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An Introduction to Deep Reinforcement Learning," *Foundations and Trends® in Machine Learning*, vol. 11, no. 3-4, pp. 219-354, 2018, doi: 10.1561/2200000071.

[15] E. Ernst, D. Damien, and A. Louette, "Introduction to reinforcement learning," in *Generative AI*, S. Feuerriegel, J. Hartmann, C. Janiesch, and P. Zschech, Eds. *Business & Information Systems Engineering*, vol. 66, no. 1, pp. 111-126, 2024.

[16] J. C. Ang, F. H. Mirzal, H. Haron, and H. M. Hashim, "Supervised, unsupervised, and semi-supervised feature selection: a review on gene selection," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 13, no. 5, pp. 971-989, Sept.-Oct. 2015.

[17] J. Ren, D. Cai, Z. Li, and X. He, "Forward semi-supervised feature selection," in *Advances in Knowledge Discovery and Data Mining: Proc. 12th Pacific-Asia Conf. PAKDD 2008*, Osaka, Japan, 2008, pp. 970-976.

[18] M. Paniri, M.B. Dowlatshahi, and H. Nezamabadi-pour, "MLACO: A multi-label feature selection algorithm based on ant colony optimization," *Knowledge-Based Systems*, vol. 192, p. 105285, 2020, ISSN 0950-7051, https://doi.org/10.1016/j.knosys.2019.105285.

[19] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "MFS-MCDM: Multilabel feature selection using multi-criteria decision making," *Knowledge-Based Systems*, vol. 206, p. 106365, 2020.

[20] A. Hashemi, M.B. Dowlatshahi, and H. Nezamabadi-pour, "Ensemble of feature selection algorithms: a multi-criteria decision-making approach," *International Journal of Machine Learning and Cybernetics*, vol. 13, pp. 49–69, 2022, https://doi.org/10.1007/s13042-021-01347-z.

[21] F. Karimi, M.B. Dowlatshahi, and A. Hashemi, "SemiACO: A semi-supervised feature selection based on ant colony optimization," *Expert Systems with Applications*, vol. 214, p. 119130, 2023, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2022.119130.

[22] A. Halder, S. Ghosh, and A. Ghosh, "Ant based semi-supervised classification," in *Proc. Int. Conf. Swarm Intelligence*, Springer, Berlin, Heidelberg, 2010, pp. 490-497.

[23] J. Albinati, H. S. Lopes, A. Freitas, and A. Gosling, "An ant colony-based semi-supervised approach for learning classification rules," *Swarm Intelligence*, vol. 9, pp. 315-341, 2015.

[24] M. Paniri, M.B. Dowlatshahi, and H. Nezamabadi-pour, "Ant-TD: Ant colony optimization plus temporal difference reinforcement learning for multi-label feature selection," *Swarm and Evolutionary Computation*, vol. 64, p. 100892, 2021, ISSN 2210-6502, https://doi.org/10.1016/j.swevo.2021.100892.

[253] A. Hashemi and M.B. Dowlatshahi, "Exploring Ant Colony Optimization for Feature Selection: A Comprehensive Review," in N. Dey (Ed.), *Applications of Ant Colony Optimization and its Variants*, Springer Tracts in Nature-Inspired Computing, Springer, Singapore, 2024, https://doi.org/10.1007/978-981-99-7227-2_3.

[26] A. Hashemi, M. Joodaki, N.Z. Joodaki, and M.B. Dowlatshahi, "Ant colony optimization equipped with an ensemble of heuristics through multi-criteria decision making: A case study in ensemble feature selection," *Applied Soft Computing*, vol. 124, p. 109046, 2022, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2022.109046.

[27] J. Li, L. Gao, and X. Li, "Multi-operator opposition-based learning with the neighborhood structure for numerical optimization problems and its applications," *Swarm and Evolutionary Computation*, vol. 84, p. 101457, 2024. [Online]. Available: https://doi.org/10.1016/j.swevo.2023.101457

[28] P. Huang, Y. Zhou, W. Deng, H. Zhao, Q. Luo, and Y. Wei, "Orthogonal opposition-based learning honey badger algorithm with differential evolution for global optimization and engineering design problems," *Alexandria Engineering Journal*, vol. 91, pp. 348-367, 2024. [Online]. Available: https://doi.org/10.1016/j.aej.2024.02.024

[29] W. Li and G. G. Wang, "Improved elephant herding optimization using opposition-based learning and K-means clustering to solve numerical optimization problems," *J Ambient Intell Human Comput*, vol. 14, pp. 1753–1784, 2023. [Online]. Available: https://doi.org/10.1007/s12652-021-03391-7

[30] S. Farahmand-Tabar and S. Shirgir, "Opposed Pheromone Ant Colony Optimization for Property Identification of Nonlinear Structures," in *Applications of Ant Colony Optimization and its Variants*, Springer Tracts in Nature-Inspired Computing. Springer, Singapore, 2024. [Online]. Available: https://doi.org/10.1007/978-981-99-7227-2_5

[31] Alice R. Malisia and Hamid R. Tizhoosh, "Applying Opposition-Based Ideas to the Ant Colony System," presented at the 2007 IEEE Swarm Intelligence Symposium, Honolulu, HI, USA, Apr. 2007, pp. 182-189. http://doi.org/10.1109/SIS.2007.368044. Available: IEEE Xplore.

[32] M. Dorigo and G. Di Caro, "Ant colony optimization: a new meta-heuristic," in *Proc. 1999 Congress on Evolutionary Computation-CEC99*, Washington, DC, USA, 1999, pp. 1470-1477 Vol. 2, doi: 10.1109/CEC.1999.782657.

[33] M. Dorigo and C. Blum, "Ant colony optimization theory: A survey," *Theoretical Computer Science*, vol. 344, no. 2–3, pp. 243-278, Nov. 2005, doi: 10.1016/j.tcs.2005.05.020.

[34] B. Chandra Mohan and R. Baskaran, "A survey: Ant Colony Optimization based recent research and implementation on several engineering domain," *Expert Systems with Applications*, vol. 39, no. 4, pp. 4618-4627, 2012, doi: 10.1016/j.eswa.2011.09.076.

[35] K. Socha and M. Dorigo, "Ant colony optimization for continuous domains," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1155-1173, Mar. 2008, doi: 10.1016/j.ejor.2006.06.046.

[36] D. Merkle and M. Middendorf, "Modeling the Dynamics of Ant Colony Optimization," in *Evolutionary Computation*, vol. 10, no. 3, pp. 235-262, Sept. 2002, doi: 10.1162/106365602760234090.

[37] C. Blum and M. Dorigo, "The hyper-cube framework for ant colony optimization," in *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 2, pp. 1161-1172, April 2004, doi: 10.1109/TSMCB.2003.821450.

[38] T. Stutzle and M. Dorigo, "A short convergence proof for a class of ant colony optimization algorithms," in *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 4, pp. 358-365, Aug. 2002, doi: 10.1109/TEVC.2002.802444.

[39] M. den Besten, T. Stützle, and M. Dorigo, "Ant Colony Optimization for the Total Weighted Tardiness Problem," in *Parallel Problem Solving from Nature PPSN VI. PPSN 2000. Lecture Notes in Computer Science*, vol. 1917. Springer, Berlin, Heidelberg, 2000, doi: 10.1007/3-540-45356-3_60.

[40] Q. Yang et al., "Adaptive Multimodal Continuous Ant Colony Optimization," in *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 2, pp. 191-205, April 2017, doi: 10.1109/TEVC.2016.2591064.

[41] X. Liu, H. Yi, and Z. Ni, "Application of ant colony optimization algorithm in process planning optimization," *J. Intell. Manuf.*, vol. 24, pp. 1–13, 2013, doi: 10.1007/s10845-010-0407-2.

[42] C. Liu, L. Wu, W. Xiao, G. Li, D. Xu, J. Guo, and W. Li, "An improved heuristic mechanism ant colony optimization algorithm for solving path planning," *Knowledge-Based Systems*, vol. 271, 2023, 110540, ISSN 0950-7051, doi: 10.1016/j.knosys.2023.110540.

[43] M. Kelidari and J. Hamidzadeh, "Feature selection by using chaotic cuckoo optimization algorithm with levy flight, opposition-based learning and disruption operator," *Soft Comput.*, vol. 25, pp. 2911–2933, 2021, doi: 10.1007/s00500-020-05349-x.

[44] H. Liu, Z. Wu, H. Li, H. Wang, S. Rahnamayan, and C. Deng, "Rotation-Based Learning: A Novel Extension of Opposition-Based Learning," in *PRICAI 2014: Trends in Artificial Intelligence*, D. N. Pham and S. B. Park, Eds. Cham: Springer, 2014, pp. 365-376, doi: 10.1007/978-3-319-13560-1_41.

[45] M. Tubishat, N. Idris, L. Shuib, M. A. M. Abushariah, and S. Mirjalili, "Improved Salp Swarm Algorithm based on opposition based learning and novel local search algorithm for feature selection," *Expert Systems with Applications*, vol. 145, p. 113122, 2020, doi: 10.1016/j.eswa.2019.113122.

[46] H. Wang, Z. Wu, S. Rahnamayan, Y. Liu, and M. Ventresca, "Enhancing particle swarm optimization using generalized opposition-based learning," *Information Sciences*, vol. 181, no. 20, pp. 4699-4714, Oct. 2011, doi: 10.1016/j.ins.2011.03.016.

[47] S. Mohapatra and P. Mohapatra, "Fast random opposition-based learning Golden Jackal Optimization algorithm," *Knowledge-Based Systems*, vol. 275, 2023, 110679, ISSN 0950-7051, doi: 10.1016/j.knosys.2023.110679.

[48] M. A. Wiering and H. Van Hasselt, "Two novel on-policy reinforcement learning algorithms based on TD (λ)-methods," in *2007 IEEE International Symposium on Approximate Dynamic Programming and Reinforcement Learning*, 2007, doi: 10.1109/ADPRL.2007.368199.

[49] P. Cichosz, "Truncating temporal differences: On the efficient implementation of TD (lambda) for reinforcement learning," *Journal of Artificial Intelligence Research*, vol. 2, pp. 287-318, 1994.

[50] H. H. Awan and W. Shahzad, "Semi-supervised associative classification using ant colony optimization algorithm," *PeerJ Computer Science*, vol. 7, p. e676, Apr. 2021.

[51] D. Berthelot, N. Carlini, I. Goodfellow, N. Papernot, A. Oliver, and C. A. Raffel, "Mixmatch: A holistic approach to semi-supervised learning," *Advances in Neural Information Processing Systems*, vol. 32, pp. 5049-5059, 2019.

[52] J. Chen, L. Zhao, Z. Ren, Z. Yan, W. Yin, and J. Gao, "Open-Set Semi-Supervised Text Classification via Adversarial Disagreement Maximization," in *Proc. 62nd Annu. Meet. Assoc. Comput. Linguistics*, 2024, pp. 574-586.

[53] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Physical Review E*, vol. 69, no. 6, 2004, 066138, doi: 10.1103/PhysRevE.69.066138.

[54] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986-1004, 2003, doi: 10.1109/TMI.2003.815867.

[55] F. Rahutomo, T. Kitasuka, and M. Aritsugi, "Semantic cosine similarity," in *The 7th International Student Conference on Advanced Science and Technology (ICAST)*, South Korea: University of Seoul, 2012.

[56] P. Xia, L. Zhang, and F. Li, "Learning similarity with cosine similarity ensemble," *Information Sciences*, vol. 307, pp. 39-52, 2015, doi: 10.1016/j.ins.2015.02.016.

[57] L. M. Gambardella and M. Dorigo, "Ant-Q: A reinforcement learning approach to the traveling salesman problem," in *Machine Learning Proceedings 1995*, Morgan Kaufmann, 1995, pp. 252-260, doi: 10.1016/B978-1-55860-377-6.50039-3.

[58] Y. Liu and B. Cao, "A novel ant colony optimization algorithm with Levy flight," *IEEE Access*, vol. 8, pp. 67205-67213, 2020, doi: 10.1109/ACCESS.2020.2986932.

[59] G. Tsoumakas, E. Spyromitros-Xioufis, J. Vilcek, and I. Vlahavas, "Mulan: A Java Library for Multi-Label Learning," *Journal of Machine Learning Research*, vol. 12, pp. 2411-2414, 2011.G. Tsoumakas, I. Katakis, and I. Vlahavas, "Mining multi-label data," Data Min. Knowl. Discov. Handb., 2009, doi: 10.1007/978-0-387

[60] G. Tsoumakas, I. Katakis, and I. Vlahavas, "Mining multi-label data," Data Min. Knowl. Discov. Handb., 2009, doi: 10.1007/978-0-387

[61] M.-L. Zhang and Z.-H. Zhou, "ML-KNN: A lazy learning approach to multi-label learning," *Pattern Recognition*, vol. 40, no. 7, pp. 2038-2048, 2007, doi: 10.1016/j.patcog.2006.12.019.

[62] J. A. Swets, "Measuring the accuracy of diagnostic systems," *Science*, vol. 240, no. 4857, pp. 1285-1293, 1988, doi: 10.1126/science.3287615.

[63] M. Norouzi, D. J. Fleet, and R. R. Salakhutdinov, "Hamming distance metric learning," in *Advances in Neural Information Processing Systems 25*, 2012.