# An improved fuzzy support vector machine algorithm for highly imbalanced datasets in the co-authorship recommendation problem

Vo Duc Quang(✉)
*School of Information and Communications Technology*
*Hanoi University of Science And Technology*
Hanoi City, Vietnam
quangvd@vinhuni.edu.vn

Hoang Huu Viet
*Faculty of Information Technology*
*Vinh University*
Vinh City, Nghe An, Vietnam
viethh@vinhuni.edu.vn

Nguyen Hai Yen
*Faculty of Information Technology*
*Vinh University*
Vinh City, Nghe An, Vietnam
nguyenhaiyen1632@gmail.com

Tran Dinh Khang
*School of Information and Communications Technology*
*Hanoi University of Science And Technology*
Hanoi City, Vietnam
khangtd@soict.hust.edu.vn

***Abstract*—The co-authorship recommendation problem aims to suggest authors join research groups based on their research topics, expertise, and previous collaborations. To address this problem, we first model it as a co-authorship network, where each author is represented as a vertex, and collaborations between authors are represented as edges. This allows us to generate two-class imbalanced datasets derived from the co-authorship networks. We then propose an adaptive weight adjustment algorithm based on FSVM-CIL to classify highly imbalanced two-class datasets. To evaluate the performance of our algorithm, we conducted experiments using self-built co-authorship datasets of various sizes and imbalance ratios. Our experimental results show that our algorithm outperformed FSVM-CIL in solving the co-authorship recommendation problem.**

***Index Terms*—Co-Authorship Network, Fuzzy Support Vector Machine, Imbalanced Dataset, Recommendation Problem, Support Vector Machine, Tomek Links**

## I. Introduction

In academic social networks, the co-authorship recommendation problem is attractive since it suggests authors join research groups to collectively write scientific articles based on their research topics, expertise, and past collaborations.

The co-authorship recommendation problem is commonly addressed by representing it as a co-authorship network $G^T = (V^T, E^T, P^T, T)$, where (*i*) $T = \{t_1, t_2, \ldots, t_k\}$ denotes a set of timestamps; (*ii*) $V^T = \{v_1, v_2, \ldots, v_N\}$ denotes the set of vertices corresponding to authors of articles within the time frame $T$; (*iii*) $P^T = \{p_1, p_2, \ldots, p_M\}$ denotes the set of articles published within the time frame $T$; (*iv*) $E^T = \{v_i, v_j, p_k, t_h\}$ denotes a set of links between authors at a specific time $t_h \in T$, where two authors $(v_i, v_j) \in V^T \times V^T$ have collaborated on the same article $p_k \in P^T$. Furthermore, the set $V^T$ can be enriched with additional attributes about each author, such as nationality, affiliation, and research topics. These attributes are denoted as $A^T = \{a_1, a_2, \ldots, a_N\}$, where $a_i$ is a feature vector containing information about a specific author/vertex pair $(v_i, v_j) \in V^T \times V^T$. Given a co-authorship network $G^T$, the co-authorship recommendation problem aims to predict the likelihood of future collaboration between authors.

So far, various approaches have been suggested for addressing co-authorship networks, including similarity measurement, statistical relational learning, graph mining, and machine learning techniques [1]. Among these, machine learning has gained significant attention since it can be flexible in choosing classification algorithms and adjusting parameters during the training process to improve the machine learning models.

In practice, co-authorship networks frequently have a significant imbalance in the distribution of connections. This arises due to the tendency of authors to collaborate with only a limited number of possible partners, resulting in a much larger number of potential collaborations than actual collaborations within the network. As a result, datasets derived from co-authorship networks typically manifest as highly imbalanced two-class datasets. Therefore, when utilizing a machine learning approach, we need robust classification algorithms to effectively tackle highly imbalanced datasets.

This paper proposes an adaptive weight adjustment algorithm based on FSVM-CIL [15] to address the two-class imbalanced learning problem. Initially, our algorithm uses FSVM-CIL to find a set of fuzzy weights for training samples. Then, our algorithm iteratively performs to train a classifier model WSVM [14] on the training samples, adjust the fuzzy weights of training samples identified by Tomek Links pairs [4], and evaluate the trained model. Our algorithm returns a classification model with the maximum geometric mean. We collected author and article data from the

979-8-3503-3122-6/23/$31.00 ©2023 IEEE
DOI 10.1109/ACOMPA61072.2023.00017

TABLE I
METRICS IN CO-AUTHORSHIP NETWORKS

| |
|---|
| **a) Neighbors-based link metrics** |
| $CN(v_i, v_j) = \|\Gamma(v_i) \cap \Gamma(v_j)\|$ |
| $AA(v_i, v_j) = \sum_{v_k \in \Gamma(v_i) \cap \Gamma(v_j)} \frac{1}{log(\|\Gamma(v_k)\|)}$ |
| $JC(v_i, v_j) = \frac{\Gamma(v_i) \cap \Gamma(v_j)}{\Gamma(v_i) \cup \Gamma(v_j)}$ |
| $PA(v_i, v_j) = \|v_i\| \times \|v_j\|$ |
| $RA(v_i, v_j) = \sum_{v_k \in \Gamma(v_i) \cap \Gamma(v_j)} \frac{1}{\|\Gamma(v_k)\|}$ |
| **b) Paths-based link metrics** |
| $SH(v_i, v_j) = \frac{1}{d(v_i, v_j)}$ |
| $Katz(v_i, v_j) = \sum_{l=1}^{\infty} \beta^l \|path^l_{v_i, v_j}\|$ |
| **c) Additional author information-based link metrics** |
| $SW(v_1, v_2, \cdots, v_N) = \begin{cases} 2, \text{if } S_1(v_1) = S_1(v_2) = \cdots = S_1(v_N), \\ 1, \text{if } S_2(v_1) = S_2(v_2) = \cdots = S_2(v_N), \\ 0, \text{otherwise}, \end{cases}$ |
| $CC(v_i, v_j) = SW(v_i, v_j) + \sum_{v_k \in \Gamma(v_i) \cap \Gamma(v_j)} SW(v_k, v_i, v_j)$ |

website *www.sciencedirect.com* through ScienceDirect APIs and created two-class imbalanced datasets. Our experimental results for our self-built co-authorship datasets showed that our algorithm outperforms FSVM-CIL for the co-authorship recommendation problem.

The remainder of this paper is structured as follows: Section II provides the background related to our proposed algorithm; Section III presents our proposed algorithm; Section IV describes our experimental results and discussions; Section V concludes our work.

## II. BACKGROUND

### A. Link metrics in co-authorship networks

In this section, we recall link metrics used in our study. Given a co-authorship network $G^T = (V^T, E^T, P^T, T)$ and two connected vertices $(v_i, v_j) \in V^T \times V^T$, the popular link metrics of $(v_i, v_j)$ consist of Common Neighbors ($CN$), Adamic Adar ($AA$), Jaccard Coefficient ($JC$), Preferential Attachment ($PA$), Resource Allocation ($RA$), Shortest Path ($SH$), Katz, Similar Work ($SW$), and Common Country ($CC$), as shown in Table I [2], [3].

The link metrics are used for evaluating the similarity levels of authors/vertices in a co-authorship network. Considering a set of timestamps denoted as $T_1$, the co-authorship candidate table constructed from the co-authorship network $G^{T_1}$ has the following structure: the rows include information about author pairs $(v_i, v_j)$ and the columns include author pairs, link measures, and labels for each author pair. Assuming that $T_2$ represents a subsequent set of timestamps following $T_1$, the co-authorship candidate table constructed from the co-authorship network $G^{T_2}$ is utilized to assign labels indicating true collaboration (labeled as $+1$) or no collaboration (labeled as $-1$) for the data samples associated with candidate author pairs. Table II illustrates the link metrics in a co-authorship candidate table.

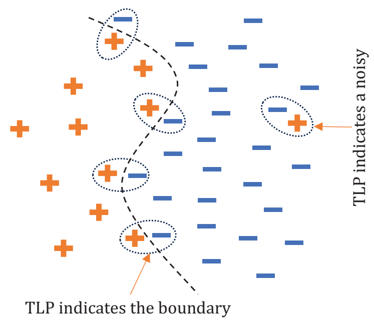

Fig. 1. An example of Tomek Links pairs

We can see that the candidate table is a set of co-authored data samples with complete attributes and class labels. Consequently, the co-authorship recommendation problem can be reformulated as a classification problem on a two-class labeled dataset, where one class is labeled $+1$ to indicate a potential future collaboration and the other class is labeled $-1$ to indicate no future collaboration. Since the number of candidate pairs can grow exponentially in large co-authorship networks. As a result, numerous samples may arise where collaboration between authors does not exist, leading to most of the label $-1$. Consequently, this situation gives rise to a highly imbalanced two-class dataset. Therefore, when approaching the co-authorship recommendation problem using a classification model, it becomes necessary to address the challenges of handling highly imbalanced two-class datasets.

### B. Tomek Links

The Tomek Links algorithm [4] is designed to identify pairs of samples from two different classes that are closest to each other in terms of distance. These pairs, determined by the algorithm, are referred to as Tomek Link pairs (TLPs). We assume that $D = \{(x_i, y_i) | i = 1, 2, \cdots, N\}$ of labeled training samples, where each sample $x_i \in \mathcal{R}^n$ is assigned to a class label $y_i \in \{-1, +1\}$. Moreover, we let $D^+$ and $D^-$ be the sets of positive- and negative-class samples in $D$, respectively, i.e. $D = D^+ \cup D^-$. The distance between $x_i \in D^+$ and $x_j \in D^-$ is denoted as $d(x_i, x_j)$. A pair $(x_i, x_j)$ is determined as a TLP if there is no sample $x_k$ such that $d(x_i, x_k) < d(x_i, x_j)$ or $d(x_j, x_k) < d(x_i, x_j)$. Figure II-B illustrates the locations of TLPs in a dataset. A TLP is either the creation of a class boundary by two samples in the pair or the possibility of one in two samples being noisy.

The Tomek Links algorithm is frequently employed in the classification problem of imbalanced datasets after applying the SMOTE algorithm [5] and its variants. This generates additional samples for the minority class, resulting in more balanced and clearer datasets. Consequently, classification algorithms exhibit improved accuracy when applied to minority samples. Recently, several improvements of the Tomek Links

TABLE II
AN ILLUSTRATED TABLE OF CO-AUTHORSHIP CANDIDATES

| Index | Candidate pairs $(v_i, v_j)$ | Linking metrics $x = \{CN(v_i,v_j), AA(v_i,v_j), JC(v_i,v_j), PA(v_i,v_j), RA(v_i,v_j), SH(v_i,v_j), Kazt(v_i,v_j), CC(v_i,v_j)\}$ | Labels $y \in \{-1,+1\}$ |
|---|---|---|---|
| 1 | $(v_1, v_2)$ | $\{CN(v_1,v_2), AA(v_1,v_2), JC(v_1,v_2), PA(v_1,v_2), RA(v_1,v_2), SH(v_1,v_2), Kazt(v_1,v_2), CC(v_1,v_2)\}$ | $+1$ |
| 2 | $(v_1, v_3)$ | $\{CN(v_1,v_3), AA(v_1,v_3), JC(v_1,v_3), PA(v_1,v_3), RA(v_1,v_3), SH(v_1,v_3), Kazt(v_1,v_3), CC(v_1,v_3)\}$ | $-1$ |
| 3 | $(v_1, v_4)$ | $\{CN(v_1,v_4), AA(v_1,v_4), JC(v_1,v_4), PA(v_1,v_4), RA(v_1,v_4), SH(v_1,v_4), Kazt(v_1,v_4), CC(v_1,v_4)\}$ | $-1$ |
| 4 | $(v_2, v_3)$ | $\{CN(v_2,v_3), AA(v_2,v_3), JC(v_2,v_3), PA(v_2,v_3), RA(v_2,v_3), SH(v_2,v_3), Kazt(v_2,v_3), CC(v_2,v_3)\}$ | $+1$ |
| ... | ... | ... | ... |

algorithm consist of OOS [6], CNN+Tomek links [7], NCL [8], SMOTE+ENN [9].

### C. WSVM and FSVM-CIL algorithms

Support Vector Machine (SVM) [10]–[12] is a well-known machine learning technique. For a two-class learning problem, we consider a set $D = \{(x_i, y_i)|i = 1, 2, \cdots, N\}$ of labeled training samples, where each sample $x_i \in \mathcal{R}^n$ is assigned to a class label $y_i \in \{-1, +1\}$. SVM aims to find a separating hyperplane defined by the pair $(w, b)$ that separates the training samples into two classes by solving the following optimization problem:

$$\begin{aligned} \text{Minimize } \Phi(w) &= \frac{1}{2}w^T w + C\sum_{i=1}^{N}\varepsilon_i \\ \text{subject to } & y_i(\langle w, \phi(x_i)\rangle + b) \geq 1 - \varepsilon_i, \\ & \varepsilon_i \geq 0, i = 1, 2, \cdots, N, \end{aligned} \quad (1)$$

where the trainning vectors $x_i$ are mapped into a higher dimentional space by the function $\Phi(x_i)(i = 1, 2, \cdots, N)$. The slack variables $\varepsilon_i > 0$ hold for misclassified samples and therefore, $\sum_{i=1}^{N}\varepsilon_i$ is the total number of misclassified samples. Besides, $C$ is a user-defined positive parameter to control the tradeoff between minimum classification error and maximum margin. The problem in Eq. (1) is a quadratic-optimization problem and can be solved by constructing a Lagrangian representation and transforming it into the a dual problem [10], [11].

SVM is a robust classification algorithm for balanced datasets. However, when SVM is applied to imbalanced datasets, it could produce a model biased toward the negative class and perform poorly on the positive class. Therefore, several enhancements to SVM have been proposed to solve the two-class imbalanced learning problem [12]–[15].

One such enhancement is the Weighted SVM (WSVM) [14]. WSVM assigns each sample $x_i$ a weight $m_i \in [0, 1]$ according to its relative importance in the class such that different sample has different contribution to forming a classification model. WSVM aims to find the solution $(w, b)$ of the following optimization problem:

$$\begin{aligned} \text{Minimize } \Phi(w) &= \frac{1}{2}w^T w + C\sum_{i=1}^{N}m_i\varepsilon_i \\ \text{subject to } & y_i(\langle w, \phi(x_i)\rangle + b) \geq 1 - \varepsilon_i, \\ & \varepsilon_i \geq 0, i = 1, 2, \cdots, N. \end{aligned} \quad (2)$$

It is important to note that higher weight values $m_i$ indicate a greater significance in accurately classifying the corresponding samples, while lower weight values $m_i$ reduce their influence on generating the optimal separating hyperplane.

FSVM-CIL [15] is an effective improvement of FSVM [12] to tackle the challenge of two-class imbalanced learning. In FSVM-CIL, higher fuzzy weight values, denoted by $m_i^+$, are assigned to positive-class samples $x_i^+$, while lower fuzzy weight values, denoted by $m_i^-$, are assigned to negative-class samples $x_i^-$, $i = 1, 2, \cdots, N$. The fuzzy membership functions are defined as follows:

$$\begin{aligned} m_i^+ &= f\left(x_i^+\right) * r^+, \\ m_i^- &= f\left(x_i^-\right) * r^-, \end{aligned} \quad (3)$$

where $f(x_i) \in (0, 1)$ represents a fuzzy membership function that reflects the importantce of a sample $x_i$ in its own class. To reflect the class imbalance, $r^+ = 1$ is assigned to positive-class samples and $r^- = r$ is assigned to negative-class samples, where $r$ is the positive-to-negative class ratio $(r < 1)$. Moreover, $f(x_i)$ is defined based on three distance measures from sample $x_i$: (*i*) the distance to its own class center ($d_i^{cen}$); (*ii*) the distance to the estimated hyperplane ($d_i^{shp}$) defined as the center of the entire dataset; and (*iii*) the distance to the actual hyperplane ($d_i^{hyp}$) formed by a basic SVM model. For each distance-based measure, FSVM-CIL constructs two fuzzy membership functions: a fuzzy linear function ($lin$) and a fuzzy exponential function ($exp$). As a result, six fuzzy membership functions for sample $x_i$ are defined as follows:

1) Based on the distance to its own class center:

$$f_{lin}^{cen}(x_i) = 1 - \frac{d_i^{cen}}{max\left(d_i^{cen}\right) + \Delta}, \quad (4)$$

$$f_{exp}^{cen}(x_i) = \frac{2}{1 + exp\left(\beta d_i^{cen}\right)}. \quad (5)$$

2) Based on the distance to the estimated hyperplane:

$$f_{lin}^{shp}(x_i) = 1 - \frac{d_i^{shp}}{max\left(d_i^{shp}\right) + \Delta}, \quad (6)$$

$$f_{exp}^{shp}(x_i) = \frac{2}{1 + exp\left(\beta d_i^{shp}\right)}. \quad (7)$$

3) Based on the distance to the actual hyperplane:

$$f_{lin}^{hyp}(x_i) = 1 - \frac{d_i^{hyp}}{max\left(d_i^{hyp}\right) + \Delta}, \tag{8}$$

$$f_{exp}^{hyp}(x_i) = \frac{2}{1 + exp\left(\beta d_i^{hyp}\right)}. \tag{9}$$

It is noted in Eqs. (4–9) that $\Delta$ is a small positive value used to avoid the cases where $f_{lin}^{cen}(x_i) = 0$, $f_{lin}^{shp}(x_i) = 0$, or $f_{lin}^{hyp}(x_i) = 0$. The parameter $\beta \in [0, 1]$ controls the slope of the exponential functions $f_{exp}^{cen}(x_i)$, $f_{exp}^{shp}(x_i)$, and $f_{exp}^{hyp}(x_i)$.

## III. Proposed Algorithm

In this section, we propose a method to adjust fuzzy weights based on tomek link pairs and an adaptive weight adjustment algorithm based on FSVM-CIL to address highly imbalanced two-class datasets.

### A. Tomek link pairs-based fuzzy weight adjustment

In dealing with a highly imbalanced two-class dataset, researchers commonly employ techniques to reduce negative-class samples [16] and generate additional positive-class samples [5] to create a more balanced dataset. Subsequently, the Tomek Links algorithm [4] eliminates noisy samples and enhances classification boundaries in the created dataset. However, removing noisy samples may discard certain positive-class samples, while they need to be kept at most. To overcome this limitation, we propose a solution that involves identifying pairs of TLPs and evaluating their positions in the distribution space of samples. Then, we adjust the weights assigned to each sample, prioritizing the importance of positive-class samples, reducing the importance of negative-class samples, and significantly decreasing the influence of noise samples while preserving the fundamental characteristics of the original datasets. To do this, we employ the Tomek Links algorithm to identify sensitive samples and design four rules for adjusting fuzzy weights to control their impact on forming a classifier model. Figure 2 illustrates the four cases in which sensitive samples are identified based on the relative positions of TLPs along with their $K$-nearest neighbors. The symbols "+" and "−" represent positive- and negative-class samples, respectively, and we chose $K = 4$. A pair of TLPs falls into one of the four cases: (1) within the positive margin, (2) outside the positive margin but classified as negative-class noise, (3) within the negative margin, or (4) outside the negative margin but classified as positive-class noise.

Our method for adjusting fuzzy weights is shown in Algorithm 1, where $h_t$ is a WSVM classifier, $h_{KNN}$ is a *KNN* classifier, $K$ is the number of nearest neighbors for a given sample $x_i$, and the set of parameters $\{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$ is used for adjusting the fuzzy weights. Specifically, the algorithm initializes a set of $TLPs = \{\}$. Then, it identifies a set TLPs of elements $(x_i, x_j)$ such that $(x_i, y_i)$ and $(x_j, y_j)$ are the nearest neightbours of each other (lines 3-10). Subsequently, for each pair $(x_i, x_j) \in$ TLPs satisfying the conditions $y_i = 1$ and $y_j = -1$, the algorithm checks and adjusts the fuzzy weights based on the four cases illustrated in Figure 2:

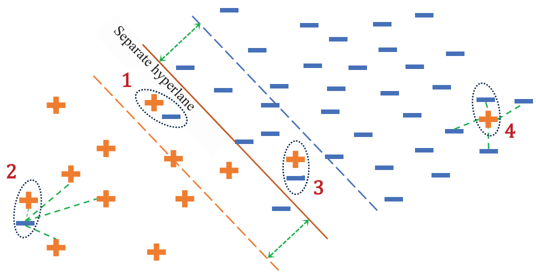


Fig. 2. An illustration of four cases for sensitive samples found by TLPs

1) If $h_t$ classified both $x_i$ and $x_j$ into the positive class, meaning $h_t(x_i) = 1$ and $h_t(x_j) = 1$, *i.e.* the pair $(x_i, x_j)$ falls within the positive margin (case 1), the algorithm follows these steps: The fuzzy weight $m_i^+$ is adjusted upward by $\sigma_1$ to increase the influence of $x_i$, while the fuzzy weight $m_j^-$ is adjusted downward by $\sigma_1$ to reduce the influence of $x_j$ (lines 13-14). However, if the $K$-nearest neighbors $x_{j_k}(j_k = 1, 2, \cdots, K)$ of $x_j$ belong to the positive class, indicating that $x_j$ was classified as a negative class noise (case 2), a significant downward adjustment of $m_j^-$ is applied to reduce the influence of $x_j$ by $\sigma_2$ (lines 15-17). Hence, to ensure appropriate adjustments, $\sigma_1$ and $\sigma_2$ are chosen such that $0 < \sigma_1 < 0.5$ and $0 < \sigma_2 < 1$.
2) If $h_t$ classified both $x_i$ and $x_j$ into the negative class, meaning $h_t(x_i) = -1$ and $h_t(x_j) = -1$, *i.e.* the pair $(x_i, x_j)$ falls within the negative margin (case 3), the algorithm follows these steps: The fuzzy weight $m_i^+$ is adjusted upward by $\sigma_3$ to increase the influence of $x_i$, while the fuzzy weight $m_j^-$ is adjusted downward by $\sigma_3$ to reduce the influence of $x_j$ (lines 20-21). However, if the $K$-nearest neighbors $x_{i_k}(i_k = 1, 2, \cdots, K)$ of $x_i$ belong to the negative class, indicating that $x_i$ was classified as a positive class noise (case 4), a significant downward adjustment of $m_i^+$ is applied to reduce the influence of $x_i$ by $\sigma_4$ (lines 22-24). Hence, to ensure appropriate adjustments, $\sigma_3$ and $\sigma_4$ are chosen such that $0 < \sigma_3 < 0.5$ and $0 < \sigma_4 < 1$.

By applying our method, we effectively increase $m_i^+$ and decrease $m_j^-$ to prioritize accurately classifying positive-class samples $x_i$. Furthermore, if $x_i$ is identified as a positive-class noise and $x_j$ is identified as a negative-class noise, our method significantly reduces both $m_i^+$ and $m_j^-$ to minimize the influence of samples $x_i$ and $x_j$ in forming a classifier model. As a result, our method yields a set of adjusted fuzzy weights, denoted as $\{m_i^+, m_i^-\}$, for all samples $x_i \in D$ $(i = 1, 2, \ldots, N)$.

### B. Adaptive weight adjustment algorithm

This section proposes an FSVM-CIL-based Adaptive Weight Adjustment (F-AWA) algorithm for the highly im-

**Algorithm 1:** TLPs-based fuzzy weight adjustment

**function** AdjFW($D, h_t, K, \sigma_1, \sigma_2, \sigma_3, \sigma_4, m_i^+, m_i^-$)
  Initialize $TLPs = \{\}$;
  **for** $i = 1$ **to** $N$ **do**
    find the nearest neighbor $(x_j, y_j)$ of $(x_i, y_i)$;
    **if** $(x_i, y_i)$ *is the nearest neighbor of* $(x_j, y_j)$ **then**
      **if** $(x_i, x_j) \notin TLPs$ *and* $(y_i \neq y_j)$ **then**
        $TLPs = TLPs \cup \{(x_i, x_j)\}$;
      **end**
    **end**
  **end**
  **for** *each* $(x_i, x_j) \in TLPs$ *that* $y_i = 1$ *and* $y_j = -1$ **do**
    **if** $h_t(x_i) = 1$ *and* $h_t(x_j) = 1$ **then**
      $m_i^+ = m_i^+ \times (1 + \sigma_1)$;
      $m_j^- = m_j^- \times (1 - \sigma_1)$;
      **if** $h_{KNN}(x_{j_k}) = 1$ **then**
        $m_j^- = m_j^- \times \sigma_2$;
      **end**
    **end**
    **if** $h_t(x_i) = -1$ *and* $h_t(x_j) = -1$ **then**
      $m_i^+ = m_i^+ \times (1 + \sigma_3)$;
      $m_j^- = m_j^- \times (1 - \sigma_3)$;
      **if** $h_{KNN}(x_{i_k}) = -1$ **then**
        $m_i^+ = m_i^+ \times \sigma_4$;
      **end**
    **end**
  **end**
  **return** $\{m_i^+, m_i^-\}, i = 1, 2, \cdots, N$;
**end function**

**Algorithm 2:** Adaptive weight adjustment algorithm

**function** F-AWA($D, h_1, K, \sigma_1, \sigma_2, \sigma_3, \sigma_4, T$)
  find $\{m_i^+, m_i^-\}$ by FSVM-CIL($D, r^+, r^-, \Delta$);
  **for** $t := 1$ **to** $T$ **do**
    fit $h_t$ using WSVM with $\{m_i^+, m_i^-\}$ for $\forall x_i \in D$;
    call AdjFW($D, h_t, K, \sigma_1, \sigma_2, \sigma_3, \sigma_4, m_i^+, m_i^-$);
    evaluate the geometric mean $g(h_t)$ of $h_t$;
  **end**
  $l := max(g(h_t)), \forall t = 1, 2 \ldots, T$;
  **return** $h_l$;
**end function**

TABLE III
DESCRIPTION OF CO-AUTHORSHIP DATASETS

| Group | Datasets | Positive sample number | Negative sample number | Total sample number | Positive percent-age |
|---|---|---|---|---|---|
| **Group I** (No. of positive samples = $33 \pm 3$) | *CoAuthor-1* | 32 | 368 | 400 | 8% |
| | *CoAuthor-2* | 36 | 564 | 600 | 6% |
| | *CoAuthor-3* | 33 | 792 | 825 | 4% |
| | *CoAuthor-4* | 35 | 1715 | 1750 | 2% |
| | *CoAuthor-5* | 31 | 1894 | 1925 | 1.61% |
| **Group II** (No. of positive samples = $70 \pm 7$) | *CoAuthor-6* | 70 | 805 | 875 | 8% |
| | *CoAuthor-7* | 75 | 1175 | 1250 | 6% |
| | *CoAuthor-8* | 71 | 1704 | 1775 | 4% |
| | *CoAuthor-9* | 68 | 3332 | 3400 | 2% |
| | *CoAuthor-10* | 63 | 3850 | 3913 | 1.61% |

balanced two-class datasets. The algorithm presented in Algorithm 2 employs the WSVM algorithm [14] as a basic classifier denoted by $h_t$ for fuzzy weight adjustment. Initially, F-AWA finds a set of fuzzy weights $\{m_i^+, m_i^-\}$ for all the samples $(x_i, y_i) \in D$ by using the FSVM-CIL algorithm [15]. Then, F-AWA runs in $T$ iterations, in which at each iteration $t = 1, 2, \ldots, T$, it follows the steps: (*i*) training a classifier model $h_t$ using WSVM with the fuzzy weights $\{m_i^+, m_i^-\}$ on all the samples $(x_i, y_i) \in D$; (*ii*) adjusting the fuzzy weights $\{m_i^+, m_i^-\}$ for all the samples $(x_i, y_i) \in D$ by using the function described in Algorithm 1; (*iii*) evaluating the geometric mean $g(h_t)$ of the model $h_t$. After completing the $T$ iterations, the algorithm outputs a classification model $h_l$ ($l \in [1, T]$) such that the geometric mean $g(h_l)$ is maximized. This indicates that $h_l$ achieves the highest accuracy for each of the two classes while maintaining a balanced performance across both classes.

## IV. Experiments

In this section, we present experiments to evaluate the efficiency of our proposed algorithm for the co-authorship recommendation problem.

### A. Datasets

To generate datasets for the co-authorship recommendation problem, we collected the article titles, publication years, content summaries, keyword lists, and authors' information from three journals consisting of *Chemical Physics Letters*, *Journal of Molecular Biology*, and *Biochemical and Biophysical Research Communications* through ScienceDirect APIs from the year 2011 to the year 2016. For each journal from the year 2011 to the year 2014, we created a co-authorship candidate table consisting of link metrics as shown in Table II. The information from the year 2015 to the year 2016 was used to determine the labels of the data samples. If two authors were co-authors on an article, the data sample was assigned a label $+1$. Otherwise, it was assigned a label $-1$. By doing so, we obtained a co-authorship dataset consisting of 477136 samples, in which the number of positive- and negative-class samples is 7683 and 469453, respectively. This means that the percentage of positive-class samples in our dataset is 1.61%.

To evaluate the performance of our algorithm for datasets with varying characteristics, we employed a bootstrap technique on our co-authorship dataset to generate 10 sub-datasets categorized into two groups of small and large sizes. Each group consisted of five datasets given in Table III. It is worth noting that datasets with a percentage of positive-class samples below 10% are considered highly imbalanced.

### B. Experimental results

Our experimental results were evaluated using several metrics, including SE (Sensitivity), SP (Specificity), GM (Geometric Mean), AUC (Area Under Curve), ACC (Accuracy), and F1S (F1-Score) [17]. However, SE, GM, and AUC are particularly crucial when evaluating the performance of algorithms for the two-class imbalanced learning problem.

We chose FSVM-CIL to compare its results with those obtained by our algorithm since FSVM-CIL outperformed both WSVM and FSVM in addressing the two-class imbalanced learning problem [15]. In our experiments, we employed a 5-fold cross-validation method for each dataset. For FSVM-CIL, we utilized optimal parameter values inherited in [15], which are as follows: $\beta \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$, $\Delta = 10^{(-6)}$, and $C = 100$. In our F-AWA algorithm, we set $K = 5$, $\sigma_2 = \sigma_4 = 0.5$, $\sigma_1 = \sigma_3 = 0.1$, $T = 20$, and the fuzzy membership functions was the Euclidean distance. Both FSVM-CIL and F-AWA algorithms were implemented using Python 3.11 on a computer with a Core i7-8550U CPU, 1.8GHz, and 16GB RAM running on Windows 11.

**Experiment 1.** In this experiment, we compared the performance of F-AWA with that of FSVM-CIL on the datasets with a small size in Group I. Our experimental results are shown in Table IV, where $l \in [1, T]$ is the iteration at which the GM achieved its maximum value. Accordingly, we observed that in 5 datasets with 30 cases using fuzzy membership functions of FSVM-CIL, our F-AWA achieved GM and AUC higher than FSVM-CIL in 23 cases. This is because we used the method of iteratively adaptive adjustment of the sample weights based on TLPs to prioritize the correct classification of positive samples. Specifically, for *CoAuthor-1* and *CoAuthor-2* (*i.e.*, the positive percentage is $8\%$ and $6\%$, respectively), F-AWA obtained the maximum GM, SE, and AUC in 3 of 6 cases using fuzzy membership functions. For *CoAuthor-3* and *CoAuthor-5*, F-AWA obtained the maximum GM, SE, and AUC in 5 of 6 cases, while for *CoAuthor-4*, F-AWA obtained the maximum GM, SE, and AUC in all 6 cases using fuzzy membership functions. This show that F-AWA is more efficient than FSVM-CIL, especially for highly imbalanced datasets.

**Experiment 2.** In this experiment, we compared the performance of F-AWA with that of FSVM-CIL on the datasets with a large size in Group II. Our experimental results are shown in Table V, where $l \in [1, T]$ is the iteration at which the GM achieved its maximum value. The experimental results show that in 5 datasets with 30 cases using fuzzy membership functions of FSVM-CIL, our F-AWA achieved GM and AUC higher than FSVM-CIL in 22 cases. Specifically, F-AWA obtained the maximum GM, SE, and AUC in 4 of 6 cases for *CoAuthor-6*, and 3 in 6 cases for *CoAuthor-7*. For datasets with lower positive percentage consisting of *CoAuthor-7*, *CoAuthor-8*, and *CoAuthor-9* (*i.e.*, the positive percentage is $4\%$, $2\%$, and $1.61\%$, respectively), F-AWA obtained GM and AUC higher than FSVM-CIL in almost cases (*i.e.*, 15 of 18 cases) using fuzzy membership functions. For three of the remaining cases, the value of GM found by FSVM-CIL and F-AWA is approximate (*i.e.*, their deviation is smaller than $1.0\%$). This shows that our F-AWA is efficient for highly imbalanced datasets of large sizes.

## V. Conclusions

In this paper, we proposed an adaptive weight adjustment algorithm based on FSVM-CIL, namely F-AWA, to address highly imbalanced two-class datasets. First, F-AWA uses FSVM-CIL to find a set of fuzzy weights for training samples. Second, F-AWA iteratively runs to train a classifier model using WSVM, adjust the fuzzy weights of training samples identified by Tomek Links pairs, and evaluate the geometric mean of the training model. Finally, F-AWA returns a classification model with the maximum geometric mean. To evaluate the performance of F-AWA for the co-authorship recommendation problem, we collected author and article data from the website *www.sciencedirect.com* through ScienceDirect APIs and created two-class imbalanced datasets. Our experimental results for our self-built co-authorship datasets showed that F-AWA outperforms the FSVM-CIL for the co-authorship recommendation problem.

## References


[1] X. Kong, Y. Shi, S. Yu, J. Liu, and F. Xia, "Academic social networks: Modeling, analysis, mining and applications," *Journal of Network and Computer Applications*, vol. 1, no. 132, pp. 86–103, 2019.

[2] H. A. Hasin and D. Hassan, "Link prediction in co-authorship networks," *Science Journal of University of Zakho*, vol. 10, no. 4, pp. 235–257, 2022.

[3] T. D. Khang, V. D. Quang, and N. D. T. Anh, "Co-authorship recommendation systems," *Hue University Journal of Science*, vol. 127, no. 2A, pp. 109–120, 2018.

[4] I. Tomek, "Two modifications of CNN," *IEEE Transactions on Systems Man and Communications*, vol. 6, no. 1, pp. 769–772, 1976.

[5] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, no. 1, p. 321–357, 2002.

[6] M. Kubat and S. Matwin, "Addressing the curse of imbalanced training sets: One-sided selection," in *Proceedings of the Fourteenth International Conference on Machine Learning (ICML 1997)*, Tennessee, United States, Jul. 1997, pp. 179–186.

[7] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 20–29, 2004.

[8] J. Laurikkala, "Improving identification of difficult small classes by balancing class distribution," in *Proceedings of the 8th Conference on Artificial Intelligence Medicine in Europe*, Cascais, Portugal, Jul. 2001, pp. 63–66.

[9] Z. Xu, D. Shen, T. Nie, and Y. Kou, "A hybrid sampling algorithm combining M-SMOTE and ENN based on random forest for medical imbalanced data," *Journal of Biomedical Informatics*, vol. 107, no. 1, pp. 1–11, 2020.

[10] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 1, p. 273–297, 1995.

[11] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 1, p. 121–167, 1998.

[12] C. F. Lin and S. D. Wang, "Fuzzy support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 464–471, 2002.

[13] R. Akbani, S. Kwek, and N. Japkowicz, "Applying support vector machines to imbalanced datasets," in *Proceedings of the 15th European Conference on Machine Learning (ECML 2004)*, Pisa, Italy, Sep. 2004, pp. 39–50.

[14] X. Yang, Q. Song, and A. Cao, "Weighted support vector machine for data classification," in *Proceedings of the 2005 IEEE International Joint Conference on Neural Networks*, Montreal, QC, Canadal, Jul. 2005, pp. 859–864.

[15] R. Batuwita and V. Palade, "FSVM-CIL: Fuzzy support vector machines for class imbalance learning," *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 3, pp. 558–571, 2010.

[16] X. Y. Liu, J. Wu, and Z. H. Zhou, "Exploratory undersampling for class-imbalance learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 2, pp. 539–550, 2008.

[17] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on knowledge and data engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

TABLE IV
CLASSIFICATION RESULTS OF FSVM-CIL AND F-AWA FOR DATASETS IN GROUP I

| Dataset | Algorithm | Fuzzy Function | Adjusted | SP (%) | SE (%) | GM (%) | F1S (%) | ACC (%) | AUC (%) |
|---|---|---|---|---|---|---|---|---|---|
| *CoAuthor-1* | FSVM-CIL | $f_{lin}^{cen}$ | None | 93.75 | 36.98 | 54.25 | 34.53 | 89.25 | 65.37 |
| | F-AWA | | $l = 12$ | 93.40 | **40.32** | **58.60** | 37.29 | 89.17 | **66.86** |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 90.59 | 44.29 | **60.19** | 34.97 | 86.92 | **67.44** |
| | F-AWA | | $l = 10$ | 89.96 | **44.44** | 60.18 | 34.18 | 86.33 | 67.20 |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 83.61 | **45.40** | **52.92** | 25.81 | 80.58 | **64.50** |
| | F-AWA | | $l = 12$ | 91.68 | 36.51 | 52.76 | 30.61 | 87.33 | 64.09 |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 89.23 | 27.14 | 42.90 | 19.65 | 84.25 | 58.19 |
| | F-AWA | | $l = 13$ | 89.23 | **33.97** | **49.75** | 25.46 | 84.83 | **61.60** |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 84.61 | 51.90 | 57.47 | 26.95 | 82.00 | 68.26 |
| | F-AWA | | $l = 5$ | 68.83 | **76.67** | **60.73** | 24.31 | 69.50 | **72.75** |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 82.60 | **53.33** | **56.43** | 28.23 | 80.33 | **67.97** |
| | F-AWA | | $l = 5$ | 93.93 | 31.27 | 48.76 | 29.16 | 89.00 | 62.60 |
| *CoAuthor-2* | FSVM-CIL | $f_{lin}^{cen}$ | None | 89.08 | **45.36** | 59.92 | 32.66 | 86.44 | 67.22 |
| | F-AWA | | $l = 1$ | 94.27 | 40.71 | **60.99** | 36.19 | 91.06 | **67.49** |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 92.67 | 40.95 | 60.48 | 32.37 | 89.56 | 66.81 |
| | F-AWA | | $l = 9$ | 91.96 | **43.57** | **62.24** | 32.95 | 89.06 | **67.77** |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 80.99 | **56.19** | **61.36** | 27.33 | 79.44 | **68.59** |
| | F-AWA | | $l = 10$ | 93.62 | 37.02 | 57.09 | 31.42 | 90.22 | 65.32 |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 82.35 | **45.24** | **52.57** | 19.98 | 80.06 | **63.79** |
| | F-AWA | | $l = 1$ | 93.38 | 29.88 | 50.52 | 25.69 | 89.56 | 61.63 |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 66.29 | 77.14 | 59.38 | 17.76 | 66.83 | 71.71 |
| | F-AWA | | $l = 1$ | 64.99 | **80.00** | **59.97** | 17.86 | 65.78 | **72.49** |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 94.38 | 35.95 | 57.32 | 32.62 | 90.89 | 65.17 |
| | F-AWA | | $l = 12$ | 94.09 | **38.81** | **59.20** | 34.85 | 90.78 | **66.45** |
| *CoAuthor-3* | FSVM-CIL | $f_{lin}^{cen}$ | None | 93.60 | 43.81 | 62.33 | 30.19 | 91.64 | 68.71 |
| | F-AWA | | $l = 1$ | 94.61 | **46.19** | **65.05** | 34.31 | 92.69 | **70.40** |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 92.30 | 38.89 | 56.91 | 25.43 | 90.18 | 65.59 |
| | F-AWA | | $l = 1$ | 92.42 | **38.89** | **58.68** | 25.50 | 90.30 | **65.66** |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 87.78 | 43.49 | 57.96 | 22.69 | 86.06 | 65.64 |
| | F-AWA | | $l = 1$ | 94.45 | **38.89** | **59.70** | 29.53 | 92.24 | **66.67** |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 93.35 | 30.16 | 50.46 | 21.92 | 90.83 | 61.76 |
| | F-AWA | | $l = 6$ | 93.18 | **34.92** | **56.39** | 24.50 | 90.87 | **64.05** |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 82.92 | 51.59 | 56.08 | 15.58 | 81.66 | 67.25 |
| | F-AWA | | $l = 1$ | 71.35 | **72.06** | **59.69** | 14.46 | 71.35 | **71.71** |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 89.04 | **45.71** | **58.25** | 25.82 | 87.35 | **67.38** |
| | F-AWA | | $l = 1$ | 95.75 | 37.14 | 56.79 | 31.40 | 93.41 | 66.45 |
| *CoAuthor-4* | FSVM-CIL | $f_{lin}^{cen}$ | None | 95.06 | 33.33 | 52.34 | 17.18 | 93.83 | 64.20 |
| | F-AWA | | $l = 15$ | 95.78 | **38.10** | **58.54** | 21.93 | 94.63 | **66.94** |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 92.75 | 40.00 | 59.24 | 16.12 | 91.70 | 66.38 |
| | F-AWA | | $l = 14$ | 93.78 | **40.00** | **59.93** | 18.06 | 92.70 | **66.89** |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 94.73 | 29.52 | 46.88 | 14.99 | 93.43 | 62.13 |
| | F-AWA | | $l = 8$ | 95.96 | **33.33** | **52.72** | 20.22 | 94.70 | **64.65** |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 93.86 | 26.67 | 42.67 | 11.65 | 92.51 | 60.26 |
| | F-AWA | | $l = 2$ | 94.71 | **26.67** | **46.54** | 13.57 | 93.35 | **60.69** |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 63.52 | 54.29 | 38.72 | 7.16 | 63.33 | 58.90 |
| | F-AWA | | $l = 1$ | 56.35 | **67.62** | **46.03** | 6.91 | 56.57 | **61.98** |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 92.63 | 31.43 | 48.32 | 13.86 | 91.41 | 62.03 |
| | F-AWA | | $l = 1$ | 92.85 | **39.05** | **55.11** | 18.97 | 91.77 | **65.95** |
| *CoAuthor-5* | FSVM-CIL | $f_{lin}^{cen}$ | None | 95.74 | 30.48 | 50.36 | 15.22 | 94.68 | 63.11 |
| | F-AWA | | $l = 4$ | 96.16 | **35.08** | **54.36** | 19.10 | 95.19 | **65.62** |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 92.54 | 39.37 | 56.55 | 13.38 | 91.69 | 65.95 |
| | F-AWA | | $l = 1$ | 92.27 | **40.32** | **57.09** | 13.09 | 91.45 | **66.30** |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 94.99 | 28.57 | 47.57 | 13.34 | 93.90 | 61.78 |
| | F-AWA | | $l = 1$ | 95.21 | **39.52** | **60.24** | 18.91 | 94.32 | **67.37** |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 94.51 | 28.89 | 46.93 | 13.53 | 93.45 | 61.70 |
| | F-AWA | | $l = 15$ | 93.65 | **32.86** | **50.05** | 12.44 | 92.68 | **63.25** |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 72.77 | **44.29** | **48.26** | 7.44 | 72.31 | **58.53** |
| | F-AWA | | $l = 1$ | 72.10 | 43.17 | 47.22 | 7.66 | 71.64 | 57.64 |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 95.95 | 25.40 | 40.96 | 11.96 | 94.81 | 60.67 |
| | F-AWA | | $l = 7$ | 95.90 | **35.56** | **53.38** | 20.16 | 94.93 | **65.73** |

TABLE V
CLASSIFICATION RESULTS OF FSVM-CIL AND F-AWA FOR DATASETS IN GROUP II

| Dataset | Algorithm | Fuzzy Function | Adjusted | SP (%) | SE (%) | GM (%) | F1S (%) | ACC (%) | AUC (%) |
|---|---|---|---|---|---|---|---|---|---|
| *CoAuthor-6* | FSVM-CIL | $f_{lin}^{cen}$ | None | 66.63 | **81.90** | **70.61** | 30.78 | 67.85 | **74.27** |
| | F-AWA | | $l = 1$ | 75.07 | 69.52 | 68.13 | 33.03 | 74.63 | 72.30 |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 91.76 | 48.57 | 66.35 | 40.14 | 88.30 | 70.17 |
| | F-AWA | | $l = 1$ | 91.22 | **51.90** | **68.49** | 41.53 | 88.08 | **71.56** |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 63.89 | **84.76** | **70.83** | 29.64 | 65.56 | **74.33** |
| | F-AWA | | $l = 5$ | 82.40 | 59.05 | 66.24 | 35.35 | 80.53 | 70.72 |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 90.14 | 44.76 | 62.60 | 34.58 | 86.51 | 67.45 |
| | F-AWA | | $l = 3$ | 89.65 | **46.19** | **63.73** | 34.79 | 86.17 | **67.92** |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 64.80 | 80.48 | 64.52 | 24.74 | 66.06 | 72.64 |
| | F-AWA | | $l = 3$ | 59.63 | **90.00** | **68.97** | 26.01 | 62.06 | **74.81** |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 95.40 | 37.14 | 58.77 | 39.75 | 90.74 | 66.27 |
| | F-AWA | | $l = 4$ | 93.46 | **46.67** | **65.23** | 41.97 | 89.71 | **70.06** |
| *CoAuthor-7* | FSVM-CIL | $f_{lin}^{cen}$ | None | 62.72 | **86.22** | **70.97** | 22.96 | 64.13 | **74.47** |
| | F-AWA | | $l = 2$ | 69.42 | 74.67 | 67.68 | 24.02 | 69.73 | 72.04 |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 93.22 | 42.22 | 61.92 | 33.68 | 90.16 | 67.72 |
| | F-AWA | | $l = 5$ | 93.65 | **44.00** | **63.44** | 36.06 | 90.67 | **68.82** |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 60.00 | **90.22** | **72.17** | 22.56 | 61.81 | **75.11** |
| | F-AWA | | $l = 1$ | 62.75 | 84.44 | 69.26 | 21.94 | 64.05 | 73.60 |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 76.31 | **61.78** | **63.54** | 25.94 | 75.44 | **69.04** |
| | F-AWA | | $l = 1$ | 81.45 | 52.89 | 60.40 | 26.25 | 79.73 | 67.17 |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 59.04 | 91.11 | 72.61 | 21.86 | 60.96 | 75.07 |
| | F-AWA | | $l = 1$ | 57.70 | **94.67** | **73.87** | 22.17 | 59.92 | **76.18** |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 93.93 | 34.22 | 54.49 | 32.23 | 90.35 | 64.08 |
| | F-AWA | | $l = 14$ | 91.86 | **43.11** | **61.32** | 34.54 | 88.93 | **67.48** |
| *CoAuthor-8* | FSVM-CIL | $f_{lin}^{cen}$ | None | 71.22 | **73.38** | 65.83 | 16.86 | 71.30 | 72.30 |
| | F-AWA | | $l = 1$ | 74.24 | 71.90 | **68.54** | 20.83 | 74.14 | **73.07** |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 93.84 | **38.71** | **59.07** | 26.92 | 91.63 | **66.28** |
| | F-AWA | | $l = 3$ | 94.39 | 38.00 | 58.46 | 27.83 | 92.14 | 66.20 |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 67.69 | **80.05** | 68.19 | 16.58 | 68.20 | **73.87** |
| | F-AWA | | $l = 1$ | 70.80 | 75.76 | **68.19** | 18.62 | 71.01 | 73.28 |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 70.66 | **73.38** | 65.95 | 16.13 | 70.76 | 72.02 |
| | F-AWA | | $l = 1$ | 77.05 | 67.57 | **67.54** | 21.35 | 76.68 | **72.31** |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 76.93 | 57.19 | 55.48 | 14.61 | 76.17 | 67.06 |
| | F-AWA | | $l = 3$ | 67.81 | **80.76** | **69.97** | 17.69 | 68.34 | **74.28** |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 95.31 | 18.52 | 29.78 | 12.82 | 92.23 | 56.92 |
| | F-AWA | | $l = 5$ | 95.13 | **37.24** | **58.67** | 29.28 | 92.82 | **66.18** |
| *CoAuthor-9* | FSVM-CIL | $f_{lin}^{cen}$ | None | 40.24 | 29.95 | 32.49 | 7.07 | 40.01 | 35.09 |
| | F-AWA | | $l = 1$ | 40.39 | **29.95** | **32.54** | 7.49 | 40.16 | **35.17** |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 46.47 | 19.34 | 29.19 | 7.85 | 45.93 | 32.91 |
| | F-AWA | | $l = 5$ | 46.83 | **19.34** | **29.33** | 8.67 | 46.28 | **33.09** |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 36.95 | **34.95** | **33.22** | 5.77 | 36.90 | **35.95** |
| | F-AWA | | $l = 1$ | 36.83 | 34.23 | 32.36 | 5.16 | 36.76 | 35.53 |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 40.06 | **27.09** | **29.29** | 5.41 | 39.78 | **33.57** |
| | F-AWA | | $l = 2$ | 46.62 | 19.40 | 28.54 | 7.80 | 46.07 | 33.01 |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 40.86 | 22.91 | 23.30 | 4.21 | 40.50 | 31.89 |
| | F-AWA | | $l = 5$ | 37.80 | **31.65** | **28.51** | 5.46 | 37.68 | **34.72** |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 45.05 | 16.43 | 20.74 | 5.47 | 44.47 | 30.74 |
| | F-AWA | | $l = 3$ | 44.21 | **25.33** | **31.95** | 11.51 | 43.82 | **34.77** |
| *CoAuthor-10* | FSVM-CIL | $f_{lin}^{cen}$ | None | 92.84 | 33.85 | 52.85 | 15.46 | 91.90 | 63.35 |
| | F-AWA | | $l = 1$ | 92.27 | **39.62** | **57.82** | 16.64 | 91.43 | **65.94** |
| | FSVM-CIL | $f_{exp}^{cen}$ | None | 93.32 | 38.14 | 58.87 | 13.96 | 92.44 | 65.73 |
| | F-AWA | | $l = 4$ | 93.61 | **39.68** | **60.23** | 15.21 | 92.74 | **66.64** |
| | FSVM-CIL | $f_{lin}^{shp}$ | None | 77.77 | **57.69** | 55.53 | 9.03 | 77.45 | 67.73 |
| | F-AWA | | $l = 2$ | 81.22 | 56.54 | **61.74** | 13.24 | 80.82 | **68.88** |
| | FSVM-CIL | $f_{exp}^{shp}$ | None | 90.01 | **37.18** | 54.57 | 10.86 | 89.17 | **63.60** |
| | F-AWA | | $l = 2$ | 93.22 | 33.40 | **55.22** | 12.27 | 92.26 | 63.31 |
| | FSVM-CIL | $f_{lin}^{hyp}$ | None | 91.40 | **21.54** | 21.84 | 3.38 | 90.28 | **56.47** |
| | F-AWA | | $l = 10$ | 97.30 | 13.72 | **26.43** | 7.91 | 95.95 | 55.51 |
| | FSVM-CIL | $f_{exp}^{hyp}$ | None | 98.47 | 13.33 | 28.88 | 10.84 | 97.10 | 55.90 |
| | F-AWA | | $l = 3$ | 92.82 | **37.88** | **56.65** | 17.69 | 91.94 | **65.35** |