Optimization for multi-objective explanation system

Yashvi Desai
Dept. of Computer Science
North Carolina State University
Email: ydesai@ncsu.edu

Jay Shah
Dept. of Computer Science
North Carolina State University
Email: jshah7@ncsu.edu

Ameya Tathavadkar
Dept. of Computer Science
North Carolina State University
Email: atathav@ncsu.edu

***Abstract*—The multi-objective semi-supervised explanation system is a machine learning approach intended to elucidate the outcomes of models trained on both labeled and unlabeled data. Its multi-objective design aims to optimize various parameters while considering the quality and accessibility of labeled data. However, in practical scenarios, the system is often constrained by a limited budget or fixed number of times it can access the values of any one example due to cost or time limitations. This necessitates an optimal utilization of the available information to generate accurate predictions and explanations for each example.**
In our research paper our main goal is to optimize the SWAY and EXPLAIN algorithms. The main goal of SWAY algorithm is to map data points to a lower dimension by minimizing distance between the points. In order to do this the whole dataset is divided into two clusters the "best" cluster and the "rest" cluster. Then the explain algorithm attempts to derive a set of rules which can differentiate between the "best" and the "rest" data.

I. Introduction

Balancing multiple objectives is a significant challenge for multi-objective semi-supervised explanation systems, particularly when some objectives are in competition or conflict with each other. For instance, optimizing for accuracy may result in a model that is less interpretable, while optimizing for interpretability may yield a model with lower accuracy. Sway by itself is also dependent on a lot of hyper-parameters. Thus our aim is to tune these hyper-parameters as well which could in turn give us a better result. In order to get a better split of the data we will also be using K-MEANS classifier which might improve the overall clustering results.Another technique that we will be exploring is the Agglomerative clustering method and study it's effects on the performance as well. After selecting the hyper-parameters which give the best possible result we will examine if our algorithm selects a better subset as compared to the "rest" of the data.

A. Structure Overview

1) Questions raised on baseline process:

- How optimal is the baseline model?
  We analyzed and experimented with the baseline model to understand its shortcoming. We aim to answer this question by develop a method which performs significantly better than the base model.
- How accurately does sway locate best from the rest?
  The *SWAY* function is a recursive algorithm that identifies the most influential data points in a given dataset. It achieves this by taking a data object and recursively splitting the rows into two groups until the group size reaches a minimum threshold, determined by the output of the *BETTER* function. The *BETTER* function, in turn, uses an exponential function of the weighted difference between the normalized values of the columns in the two rows to make a decisive rule for best and rest. We want to examine how different metrics for weighted distance calculation affect the performance of *SWAY*.
- How does the clustering method in baseline *SWAY* compare to other clustering methods?
  To evaluate the effectiveness of the clustering method used in baseline *SWAY*, we plan to explore and compare it with other clustering methods. While the baseline approach employs a straightforward cosine similarity measure for creating clusters, we aim to investigate more advanced and intricate clustering techniques. By doing so, we hope to gain insights into the strengths and limitations of different clustering methods and to determine which one provides the best results for our particular application.

2) Contribution: Optimization is a crucial and ongoing discussion in any research work. To achieve better results, it is necessary to explore various strategies and conduct extensive experimentation. In this research, we have delved into numerous areas for optimization of the explanation system. Our work highlights the significance of hyperparameter tuning in a semi-supervised learning system. We conducted experiments with different distance metrics to eliminate the underperformers and determine the best performing metric. The paper presents the results of our efforts in refining the sway method by employing different clustering techniques. Furthermore, this paper offers insights into how different upscaling strategies interact with each other and affect the system's performance. The findings of this research can be utilized by researchers and practitioners alike to improve the accuracy and efficiency of semi-supervised learning systems.
3) Caveats: The aim of the Sway is to reduce the number of evaluations and faster in execution.Sway with PCA might

work with smaller data, but with higher dimensional data, the results of PCA cannot be trusted because of the presense of redundant and noisy data, since then the identification of important features would be difficult.
Also, DBSCAN is a clustering alogrithm that groups cluster on the basis of density, however sway splits a single cluster into multiple smaller cluster. This might lead to a loss of information and may results in incorrect clustering results. DBSCAN with sway is also computationally expensive for large datasets.This can lead to slow performance or even memory errors when working with very large datasets.

II. Related Work

In case of most practical optimization problems the main objective is to optimize objectives which could be mutually conflicting at the same time. A multi objective algorithm can be represented as follows [1]:
minimize $F(x) = (f_1(x), f_2(x), ...f_m(x))^T$
OR
maximize $F(x) = (f_1(x), f_2(x), ...f_m(x))^T$
subject to: x ϵ Ω
where Ω is the decision space and x ϵ Ω is a decision vector , Ω belongs to R^n [1] Many times due to conflict among objectives, a MOP does not have a single optimal solution for all objectives and instead has a trade off solution which is known as the Pareto Optimal Solution [1]. In many cases in order to tackle such problems a category of algorithms known as Evolutionary Algorithms is used. These are heuristic based search algorithms which have proved to be effective in case of multi objective search problems[1]. Population based search and information exchange amongst individual data points is what these algorithms do best. Evolutionary Algorithms are able to obtain multiple Pareto Optimal Solutions in a single run of the simulation and do not have to be aware of derivative information or aggregate different properties thus overcoming the drawbacks of classical methods.
Since SWAY has already proved to be performing better than Evolutionary Algorithms like NSGA-II and SATIBEA which are state of the art alternatives in many such scenarios it seems to be a better benchmark than these EA algorithms [2]. SWAY itself is an ante hoc algorithm which comes up with rules from bins of data. These rules become the selection factor when we are differentiating best and rest data points. Overall SWAY has all the properties desirable for a baseline method like simplicity of implementation and quick execution time [2].
Distance-based automated semi-supervised clustering algorithm have been proved to identify the correct patterns/classes in the data space while eliminating the possible noisy-label information [3]. The problem of semi-supervised clustering in this context can be formulated as a multi-objective optimization problem, wherein the deviation of data points in clusters and the prediction error of the outcome variable can be considered as objective functions to be minimized [4]. Agglomerative clustering is a widely used technique for grouping similar instances together based on their features, which can be effectively utilized to optimize a multi-objective semi-supervised explanation system. Therefore, by identifying the patterns and relationships in the data, Agglomerative clustering can help to create more accurate and informative explanations for the system and serve as a useful tool for optimizing a multi-objective semi-supervised explanation system.

III. Methods

A. Algorithms

1) SWAY: SWAY algorithm is a semi-supervised clustering method that partitions the input dataset into two clusters by iteratively selecting a pivot row and dividing the rows into two halves based on their similarity to the pivot row. The algorithm selects a pivot row using an active learning strategy or a random selection, and then calculates the similarity of all other rows to the pivot row. The algorithm then divides the rows into two halves based on their similarity to the pivot row. This process is repeated recursively on each half until the size of the remaining rows is smaller than a threshold value.
If the similarity between the two halves is greater than a predefined threshold, the algorithm adds the rows in the worse half to the worse list and recursively calls the function on the better half. For this, the sway function uses a worker function to cluster the input dataset data and returns three values: two lists containing the best and worst clusters identified by the algorithm, and the total number of evaluations performed during the clustering process.
2) SWAY(K-means): In this case we used the crux of the SWAY algorithm but have tried to optimize it's performance by using KMEANS in order to generate clusters. Now instead of using the pivot row to evaluate the similarity and split data points into clusters we are using KMEANS in order to split data into two clusters. Based on this split the KMEANS algorithm gives us the best and rest values. For this swayKMEANS uses workersway function which is responsible for initiating the KMEANS algorithm.Additionally we have also performed hyper parameter tuning in order to try and get imrproved performance.
3) SWAY(Agglomerative): For Sway with Agglomerative clustering we use Agglomerative Clustering in order to generate the group of clusters from which we can derive the best and the rest tuples. Agglomerative clustering is a bottom up clustering algorithm which initially treats individual data point as a single cluster. It then starts to recursively merge clusters based on their similarity(in our case euclidean distance) until a threshold value is reached. This is another approach of getting the halves that comprise of the best and worst clusters. In this case as well we have done some hyper parameter tuning to extract the best possible performance.
4) EXPLAIN: The Explain algorithm is a rule-based classification algorithm that generates a set of rules with multiple conditions from a given dataset. The algorithm

aims to identify the important predictors that are relevant for classifying the target variable and generate rules based on these predictors.
The explain function takes in three input arguments: data, best, and rest. The data argument denotes the dataset that is employed to produce the rules. The best argument represents a portion of the data that performs better than a particular threshold, whereas rest signifies a portion of the data that represents the lower performing class. To generate candidate rules, the function employs the bins function to divide the data into several bins based on the values of the predictor variables. For each bin, the function determines the most informative predictor variables that best separate the positive and negative classes by utilizing the value function. This function computes a score for each variable based on its ability to distinguish between the positive and negative classes.

B. Data

For this paper 11 different datasets, which are of different categories are used. The datasets consists of two type of attributes, numeric and symbolic. The summary of all the datasets is as follows:
1. auto93 - This is a dataset about the details of cars. This dataset has 398 rows and 8 columns. The goal attempts to minimize the weight of the car, and maximize Acceleration and Miles per Gallon.
2. auto2 - This dataset is a modified version of StatLib dataset. There are 93 rows and 23 attributes, in the dataset. The multi goal involves minimizes the weight and maximizing the acceleration
3. china - This dataset has 499 instances and 19 columns.The objective includes minimizing the N_efforts.
4. coc1000 - This dataset is in the software project estimation domain. There are 1000 rows and 25 columns out of which 5 of them are dependent attributes and the rest are independent. The objective is maximize LOC and minimize application experience, platform experience, risk and effort.
5. coc10000 - This is a dataset with 10000 rows and 25 columns. The goal of this dataset is same as coc1000.
6. nasa93dem - The dataset has 94 rows with 29 attributes, in the software effort domain. This dataset aims to maximize the KLOC and minimize the effort, Defects and months.
7. healthCloseIsses12mths0001-hard - This is a dataset with 10,000 rows and 8 columns. 3 of them are dependent attributes and the rest 5 are independent attributes. The aim is to minimize the magnitude of relative error, and maximize the percentage of relative error deviation and accuracy.
8. healthCloseIsses12mths0011-easy - This dataset has the same domain of issue close as the above one, the difference is that the data is ”easier” than the above dataset.
9. pom - This dtaset has 10,000 rows and 13 attributes, in the agile project management space. The multi-goal for this dataset is to maximize the completion rate and minimize the ideal time and cost.
10. SSM - It is one of the largest dataset with 239,360 rows and 15 columns. This dataset belongs to the computational physics domain. This dataset aims to minimize the number of iterations and time to solutions.
11. SSN - The dataset has 53662 instances and 19 columns. The goal of this dataset is to minimize the PSNR and Energy.

C. Performance Measure

Prior to executing any algorithms on the data, all data attributes were used to rank the data based on the Zitzler predicate, and the resulting rankings were normalized from 1 to 100, inclusive. This ranking system proves beneficial for summarization techniques, as it allows for the identification of the ”superior” rows in the data. An experiment was then conducted to identify the efficieny of the algorithms, where the algorithms were run over 20 iterations and under a fixed budget. The distribution of the evaluation of each algorithm are summarized.

D. Summarization Methods

To get valuable outcome from the results and to see how different each algorithm performs compared to the others, statistical tests are used.
We used bootstrap test as significance test, here a large number of samples (called bootstrap samples) are drawn with replacement from the original dataset. Initally, we calculate the bootstrap statistic, which is the difference between the statistic calculated on the original data and the statistic calculated on the bootstrap samples. We then calculate the p-value. If the p-value is less than a pre-specified significance level. The significance level we used is of 0.05, we reject the null hypothesis and conclude that the parameter estimate is statistically significant.
Cliff Delta is used to measure the difference in effect size of the two groups. It ranges from -1 to 1, where a value of 0 indicates no difference between the groups, a value of 1 indicates complete difference, and a value of -1 indicates complete reversal of the difference.
The conjunction of effect size test and significance test are deployed to get the comparison table

IV. RESULTS

1) Answering raised questions:

- How optimal is the baseline model ?
  It does a good job at dividing the clusters and separating the best from the rest. But it not the most optimal method. In order to see if the baseline model is optimal or not we compare the results that we get from sway with ”all”. If the results are better then we can be sure that sway is performing as intended. In order to check how optimal our current baseline model is we will try changing the set of hyper-parameters to generate a new result ”sway2”. We will then compare the results achieved from ”sway2” with sway to determine if the baseline model was most optimal or not.

- How accurately does sway locate best from the rest? Our team experimented with various weighted distance matrices to find the optimal distance formula that results in the best top results. Amongst them were the weighted Manhattan and Euclidean distance. Ultimately, we found that the existing exponential formula yielded the best results out of the three experimented formulas.

$$EuclideanDistance = \sqrt{\sum_{i=1}^{n} (w_i(x_i - y_i))^2}$$

- How does the clustering method in baseline SWAY compare to other clustering methods? Our results show that the baseline simple, clustering function served the purpose of the method but was not the optimal solution. On trying other clustering methods, we found a better performing Explanation System. While K-means clustering was an under-performing method in most the cases, Agglomerative clustering showcased better performance in almost all cases.

In this section, we present summarized results obtained from refined SWAY clustering methods. To evaluate these methods, we have included result and comparison tables from a few test datasets. In our experimentation, we used three different versions of SWAY. The first version, called *sway*, is the original function from the baseline. The second version, *sway2*, is based on k-means clustering, while the third version, *sway3*, is based on agglomerative clustering. The *top* row of each table displays the top performing rows, which are selected using the "betters" function. This function sorts the rows in descending order based on their score and returns the top "n" rows. These tables provide insights into the performance of different SWAY clustering methods and can aid in selecting the most suitable method for a given dataset.

Our objective is to delve into the behavior of data after making adjustments to the clustering methods, employing different distance metrics and tuning hyperparameters. We aimed to compare various clustering methods with the baseline approach and optimize them to create a method that separates the best and the rest of the data rows. Through experimentation, we were able to achieve successful results, and Sway3, which uses agglomerative clustering, emerged as the best-performing method. This finding highlights the effectiveness of Sway3 in clustering data and supports the idea that selecting appropriate clustering techniques can significantly improve the performance of such explanation systems.

In Tables I through XVI, we present the results of our optimized model and analyze its behavior. In comparing the three different explanation systems, we observed only minor fluctuations in their performance ranking. Specifically, we considered the following systems: S, which utilizes the baseline clustering method in SWAY; S2, which employs K-means clustering; and S3, which utilizes agglomerative clustering. Upon ranking the worst performer, we found that S and S2 were closely matched, with over 50% of test files giving the worst rank to S2. Based on these results, we were unable to make a definitive conclusion as to whether K-means clustering is a good method for this use-case. To address this, we introduced agglomerative clustering, which produced promising results. In fact, more than 50% of the test files showed S3 as the best performing method for maximizing and minimizing all attributes. With only a few exceptions, such as Table III, we found that S3 was consistently the best performer. Even in cases where S3 was not the top performer, it ranked as the second-best performer and was never the overall worst performer.

	Lbs-	**Acc+**	**Mpg+**
all	2807	15.5	20
sway	2151.75	16.33	34
sway2	1970	16.4	30
sway3	1836	17.515	30
xpln	2226.35	15.98	30
xpln2	2807	15.5	20
xpln3	2778.95	15.51	20.5
top	1995	18.8	40

TABLE I
AUTO93 RESULTS

	Lbs-	**Acc+**	**Mpg+**
all to all	=	=	=
all to sway	$\neq$	$\neq$	$\neq$
all to sway2	$\neq$	$\neq$	$\neq$
all to sway3	$\neq$	$\neq$	$\neq$
sway to sway2	$\neq$	$\neq$	$\neq$
sway to sway3	$\neq$	$\neq$	$\neq$
sway2 to sway3	$\neq$	$\neq$	=
sway to xpln	$\neq$	$\neq$	$\neq$
sway2 to xpln2	$\neq$	$\neq$	$\neq$
sway3 to xpln3	$\neq$	$\neq$	$\neq$
sway to top	$\neq$	$\neq$	$\neq$

TABLE II
AUTO93 COMPARISON

	CityMPG+	**HighwayMPG+**	**Weight-**	**Class-**
all	21	28	3040	17.7
sway	23.5	30.7	2946	16.095
sway2	31	36	1845	8.4
sway3	23.8	31	2679	14.34
xpln	22.4	29.35	2967	16.655
xpln2	27.85	32.55	2481.5	10.68
xpln3	23	29.85	2815.25	14.965
top	33.6	41.2	2055	9.08

TABLE III
AUTO2 RESULTS

	CityMPG+	HighwayMPG+	Weight-	Class-
all to all	=	=	=	=
all to sway	≠	=	≠	≠
all to sway2	≠	≠	≠	≠
all to sway3	=	≠	≠	≠
sway to sway2	≠	≠	≠	≠
sway to sway3	≠	≠	≠	≠
sway2 to sway3	≠	≠	≠	≠
sway to xpln	≠	≠	≠	≠
sway2 to xpln2	≠	≠	≠	≠
sway3 to xpln3	≠	≠	≠	≠
sway to top	≠	≠	≠	≠

TABLE IV
AUTO2 COMPARISONS

	N_effort-
all	2098
sway	427.05
sway2	170
sway3	190.2
xpln	649.95
xpln2	613.3
xpln3	600.05
top	155

TABLE V
CHINA RESULTS

	N_effort-
all to all	=
all to sway	≠
all to sway2	≠
all to sway3	≠
sway to sway2	≠
sway to sway3	≠
sway2 to sway3	≠
sway to xpln	≠
sway2 to xpln2	≠
sway3 to xpln3	≠
sway to top	≠

TABLE VI
CHINA COMPARISONS

	LOC+	AEXP-	PLEX-	RISK-	EFFORT-
all	1059.8	3	3	5.1	19540.5
sway	1125.35	2.95	2.85	4.65	20826
sway2	1298	3	3	3	30169
sway3	1547.7	2.35	2.75	8.45	66982.9
xpln	1069.4	2.9	3	5.35	20069.4
xpln2	1066	3	3	4.9	19770
xpln3	1052.55	3	3	5.55	20600.7
top	1556.2	2	1	3.6	30079.6

TABLE VII
COC1000 RESULTS

	LOC+	AEXP-	PLEX-	RISK-	EFFORT-
all to all	=	=	=	=	=
all to sway	≠	≠	≠	≠	≠
all to sway2	≠	=	=	≠	=
all to sway3	≠	≠	≠	≠	≠
sway to sway2	≠	≠	≠	≠	≠
sway to sway3	≠	≠	≠	≠	≠
sway2 to sway3	≠	≠	≠	≠	≠
sway to xpln	≠	≠	≠	≠	≠
sway2 to xpln2	≠	≠	≠	≠	≠
sway3 to xpln3	≠	≠	≠	≠	≠
sway to top	≠	≠	≠	≠	≠

TABLE VIII
COC1000 COMPARISONS

	Kloc+	Effort-	Defects-	Months-
all	47.5	252	2007	21.4
sway	47.375	247.74	1833.45	19.905
sway2	5.5	24	188	9.9
sway3	2.2	8.4	69	6.6
xpln	32.655	158.06	1283.8	17.575
xpln2	17.74	64.2	574.45	14.215
xpln3	14.225	63.98	487.55	12.635
top	5.5	18	172	9.1

TABLE IX
NASA93DEM RESULTS

	Kloc+	Effort-	Defects-	Months-
all to all	=	=	=	=
all to sway	≠	≠	≠	≠
all to sway2	≠	≠	≠	≠
all to sway3	≠	≠	≠	≠
sway to sway2	≠	≠	≠	≠
sway to sway3	≠	≠	≠	≠
sway2 to sway3	≠	≠	≠	≠
sway to xpln	≠	≠	≠	≠
sway2 to xpln2	≠	≠	≠	≠
sway3 to xpln3	≠	≠	≠	≠
sway to top	≠	≠	≠	≠

TABLE X
NASA93DEM COMPARISON

	MRE-	ACC+	PRED40+
all	75.0325	7.1305	25
sway	75.1825	7.1385	25.625
sway2	100	0	37.5
sway3	77.753	6.4335	27.5
xpln	75.0835	7.164	25.625
xpln2	75.058	7.1665	25
xpln3	77.2435	6.525	26.25
top	64.91	11.3335	25

TABLE XI
HEALTH-HARD RESULTS

	MRE-	ACC+	PRED40+
all to all	=	=	=
all to sway	≠	≠	≠
all to sway2	≠	≠	≠
all to sway3	≠	≠	≠
sway to sway2	≠	≠	≠
sway to sway3	≠	≠	≠
sway2 to sway3	≠	≠	≠
sway to xpln	≠	≠	≠
sway2 to xpln2	≠	≠	≠
sway3 to xpln3	≠	≠	≠
sway to top	≠	≠	≠

TABLE XII
HEALTH-HARD COMPARISON

	MRE-	**ACC+**	**PRED40+**
all	75.1215	7.164	25
sway	77.94	6.3965	26.875
sway2	100	0	37.5
sway3	85.09	4.2895	30.625
xpln	77.7	6.4155	26.875
xpln2	75.0455	7.144	25
xpln3	82.2605	5.077	28.75
top	64.91	11.33	25

TABLE XIII
HEALTH-EASY RESULTS

	MRE-	ACC+	PRED40+
all to all	=	=	=
all to sway	≠	≠	≠
all to sway2	≠	≠	≠
all to sway3	≠	≠	≠
sway to sway2	≠	≠	≠
sway to sway3	≠	≠	≠
sway2 to sway3	≠	≠	≠
sway to xpln	≠	≠	≠
sway2 to xpln2	≠	≠	≠
sway3 to xpln3	≠	≠	≠
sway to top	≠	≠	≠

TABLE XIV
HEALTH-EASY COMPARISON

	Cost-	**Completion+**	**Idle-**
all	325.123	0.898	0.2415
sway	210.213	0.895	0.2375
sway2	385.81	0.92	0.3
sway3	156.024	0.892	0.2325
xpln	287.404	0.899	0.2305
xpln2	327.835	0.896	0.236
xpln3	290.282	0.899	0.239
top	138.846	1	0

TABLE XV
POM RESULTS

	Cost-	**Completion+**	**Idle-**
all to all	=	=	=
all to sway	≠	≠	≠
all to sway2	≠	≠	≠
all to sway3	≠	≠	≠
sway to sway2	≠	≠	≠
sway to sway3	≠	≠	≠
sway2 to sway3	≠	≠	≠
sway to xpln	≠	≠	≠
sway2 to xpln2	≠	≠	≠
sway3 to xpln3	≠	≠	≠
sway to top	≠	≠	≠

TABLE XVI
POM COMPARISON

V. DISCUSSION

A. *Threats To Validity:*

In our case we have used KMEANS and Agglomerative Clustering. As mentioned earlier KMEANS is highly susceptible to outliers. Thus, if the data is not cleaned correctly KMEANS will keep on giving poor performance. Also KMEANS is not a good option in case of large datasets as it stores lot of information and has to keep on sorting and calculating all the distances. In case of Agglomerative clustering it is not as sensitive to outliers as compared to KMEANS. On the other hand the time complexity of Agglomerative clustering is quite high. So it doesn't work very well when we try to scale our data.

B. *Discussion*

Since SWAY and SWAY with Agglomerative clustering provides efficient results in case of Multi Objective Search Space Problems it could also be a viable option for Single Objective Search Space Problem as bench marking algorithms.

C. *Future Work*

Based on our current progress we would say that it is possible to explore other solutions by combining various algorithms like PCA,KNN,DBSCAN with SWAY. As far as our current approaches are concerned it is possible to get a better result from KMEANS by using grid search to optimize it's parameters. In case of Agglomerative we can use various distance metrics to see how it affects the results. We can also permute through various linkage criterion which are used to decide which two clusters to merge at each step. In order to overcome the high time complexity of the algorithm for large datasets we can use a parallel processing methods like multi-processing or distributed computing.

VI. CONCLUSION

Based on the experimental results, it can be observed that the baseline SWAY model performed exceptionally well as a benchmark algorithm. However, the SWAY(KMEANS) model demonstrated inferior performance compared to the benchmark in many cases. This could be due to KMEANS's sensitivity to the scale of data and irrelevant features. Moreover, KMEANS is prone to being affected by outliers that may be present in the data. In contrast, the Agglomerative clustering approach has shown promising results. The SWAY(AGGLOMERATIVE) model outperformed the benchmark model, as observed in the results section. This could be due to the fact that AGGLOMERATIVE clustering approaches generally perform better for problems with large search spaces. Therefore, it can be concluded that the Agglomerative clustering method can be a suitable alternative for optimizing the performance of the SWAY model.

References

[1] https://iopscience.iop.org/article/10.1088/1742-6596/1288/1/012057

[2] https://arxiv.org/abs/1608.07617

[3] Gupta, A. et al. (2022) Semi-supervised cascaded clustering for classification of Noisy label Data, arXiv.org. Available at: https://arxiv.org/abs/2205.02209 (Accessed: April 20, 2023).

[4] Multi-objective Semi-supervised Clustering for Finding Predictive Clusters ¿ cs ¿ arXiv:2201.10764 Computer Science ¿ Neural and Evolutionary Computing [Submitted on 26 Jan 2022] Multi-objective Semi-supervised Clustering for Finding Predictive Clusters Zahra Ghasemi, Hadi Akbarzadeh Khorshidi, Uwe Aickelin