SVDD sampling with approximation guarantees for the decision boundary

Project Report

Rahul Reddy Muppidi -IIT2018103

Prathyush- IIT2018124

Mallampalli Maheswarnath-IIT2018137

Gowni Sai Pavan-IIT2018139

Instructor: Prof Om Prakash Vyas

opvyas@iiita.ac.in

Abstract— **Support vector data description (SVDD) is a machine learning technique that is used for outlier detection and classification. The main idea of SVDD is to create a spherically shaped boundary around the dataset. Despite the approach's potency, SVDD cannot perform well with large data size.In the SVDD sampling method it generally takes the subsets of the training data on which SVDD trains a decision boundary. Theoretically, a good sample should contain the boundary observations that SVDD would select as support vectors on the whole data set. However, outlying data are so important to not scrap the contiguous inlier regions and hence can avoid poor classification accuracy.**

In this paper,we proposed an efficient algorithm which does not require any hyper-parameters and which scales well with the large real word data which has better classification accuracy and less runtime.

Keywords—Classification, Outlier Detection,Hyper-parameters,Accuracy,Runtime

Introduction

The classification is the most important and esteemed area. Classification is a process of categorizing a given set of data into different classes;this is a process to check whether a test data belongs to the class or not. The classification task becomes more challenging when datasets suffer from data irregularity problems like class imbalance, class distribution skew. Normally, when a class is not well defined or absent, then the classifier may not give unbiased and correct results which is known as one-class classification. In the problem of one-class classification, there are a sufficient number of positive class objects and very few outliers, i.e. either negative data objects (a class which is of no interest) are present in small quantities. This phenomenon makes decision boundary detection a difficult task. This phenomenon makes decision boundary detection a hard task.It is also observed that like conventional classification problems such as measuring the complexity of a solution, the estimation of classification error, the generalisation of classification methods and the curse of dimensionality also appears in one-class classification.To tackle such problems SVDD can be used for sampling the dataset.

Support vector data description (SVDD) is a one-class classifier for detection of the outlier. The SVDD identifies the smallest hypersphere consisting of all the data points around the majority of inliers, to distinguish them from the outliers.One of the defining characteristics of SVDD is that only a few observations decide the decision boundary. Hence, a good sample is one for which svdd selects support vectors similar to the ones obtained on the full data set.

But fine-tuning the existing methods such that they identify boundary points the correct boundary is difficult. A reason is that the boundary they return depends significantly on the choice of hyper-parameters, and selecting suitable parameter values is not intuitive. The limitation is that including all boundary points in a sample does not guarantee SVDD training to return the original support vectors. The drawback is that selection of support vectors depends on factors such as the ratio of inliers and outliers in the sample and a sufficient number of non-boundary observations in the sample. Disregarding the few contiguous inlier regions and it yields the wrong outlier classifications after sampling. This makes the process of selection of suitable svdd sampling methods difficult.

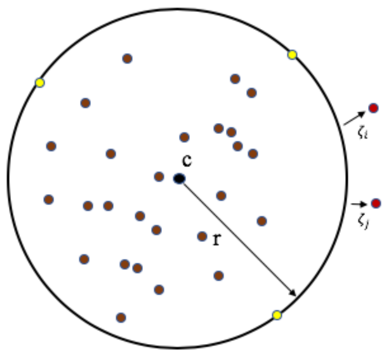


Fig [1] The hypersphere containing the target data having center c and radius r. Objects on the boundary are support vectors

I. Problem Statement

Our objective is to analyse, validate and compare the working and efficiency of various algorithms for SVDD that are focused and sample space reduction such as FDPE and RAPID. We intend to improve efficiency by pre-processing.

II. Problem Scope and Data Set

Multiple datasets will be used in this project We are also using some datasets explored in “On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study. Data Mining and Knowledge Discovery”[8]. We are also using benchmark datasets like MNIST, iris,credit card, svmguide.

Problem scope is limited to One- class classification in datasets.

III. Literature Survey

We made a brief summary of some research papers on the current topic. It contains columns of paper name, domain, methodology and results. You can see the table at below link and also in Appendix 1.

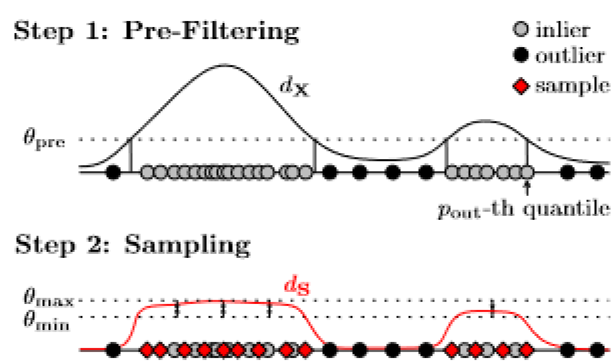
https://docs.google.com/document/d/1OEIWoB-I8WsDVZpblDVpqvOgwyKWx6EzYSwz9wlvaxU/edit?usp=sharing

IV. Methodology

The methodology is to use Density Based Sampling to scale SVDD for large data sets .We use this method so that we can exploit that an SVDD decision boundary is in fact a level-set estimate, and that inliers are a superlevel set. The idea behind our sampling method is to remove observations from a data set such that the inlier super-level set does not change.The super-level set of inliers does not change as long as not-selected observations have higher density than the minimum density of selected observations.The method has three parts they are:

1. Density-based Pre-Filtering
2. Optimal Sample Selection
3. RAPID Approximation

1. Density-based Pre-Filtering:



Fig[2]

Here we will separate the unlabeled data into outlier and inlier regions based on their empirical density.

We have to reduce the size of the data as much as possible and also maintain a good classification accuracy on the sample.

One can frame this as an optimization problem as

$$\begin{aligned} &\underset{\mathbf{S}}{\text{minimize}} && |\mathbf{S}| \\ &\text{subject to} && \mathit{diff}(f^{\mathbf{S}}, f^{\mathbf{X}}) \leq \varepsilon, \end{aligned}$$

where diff is a similarity between two decision functions and ε a tolerable deterioration in accuracy. Solving this optimization problem requires knowledge of f^X. But obtaining this knowledge is not suggested because |X| is too large to solve. However, we know that the SVDD hyperparameter C defines a lower bound on the share of observations predicted as outliers in the training data. A special case is if C = 1, since $f^X(x; C=1) = in, \forall x \in X$. In this case, diff is zero if SVDD trained on S, i.e., f^S, also includes all observations within the hypersphere.

First, we pre-filter the data based on their empirical density, such that a share of p_{out} observations are outliers. Formally, p_{out} is equivalent to choosing a threshold θ_{pre} on the empirical density, where θpre is the p_{out}-th quantile of the empirical density distribution. Using this threshold in a level-set classifier separates observations into inliers I and outliers O.

$$I = \{x \in X: g^X_{\theta pre} = in\}$$
$$O = \{x \in X: g^X_{\theta pre} = out\}$$

where g^X_θ is level-set classifier.

Second, we replace f^X with f^I and set C = 1. With this, we know that $f^I(x) = in, \forall x \in I$, without training f^I. Put differently, pre-filtering the data with an explicit threshold allows to get rid of an implicit outlier threshold C. This in turn allows to estimate the level set estimated by SVDD without actually training the classifier.

2. Optimal Sample Selection:

After pre-filtering, we can reduce Optimization Problem to a feasible optimization problem. We begin by replacing f^X with f^I.

We then frame sample selection as an optimization problem where the constraints enforce the density rule which is the sample selection step.

$$\begin{aligned} \underset{\mathbf{S}}{\text{minimize}} \quad & |\mathbf{S}| \\ \text{subject to} \quad & \mathit{diff}(f^{\mathbf{S}}, f^{\mathbf{I}}) \le \varepsilon. \end{aligned}$$

We know that both classifiers have equivalent level-set classifiers. We set $g^I_{\theta pre}$ as the equivalent level-set classifier for f^I. For f^S, there also exists a level-set classifier $g^S_{\theta'}$, but the level set θ' depends on the choice of S. Thus, we must additionally ensure that θ' indeed is the level set estimated by training SVDD on S.The modified optimization problem is

$$\begin{aligned} \underset{\mathbf{S}}{\text{minimize}} \quad & |\mathbf{S}| \\ \text{subject to} \quad & \mathit{diff}(g^{\mathbf{S}}_{\theta'}, g^{\mathbf{I}}_{\theta}) \le \varepsilon \\ & g^{\mathbf{S}}_{\theta'} \equiv f^{\mathbf{S}}, \end{aligned}$$

From theorem $g^S_{\theta'} \equiv g^I_\theta$ if d_S is uniform on I we propose to quantify the fit with a uniform distribution as the difference between the maximum density $\theta_{max} = \max_{x \in S} dS(x)$ and minimum density $\theta_{min} = \min_{x \in S} dS(x)$:

$$\Delta S_{fit} = \theta_{max} - \theta_{min}$$

One of the consequences of only approximating a uniform density is that there may be some not-selected observations $x \in I \setminus S$ with a density value $d_S(x)$ less than θ_{min}. Since the level set estimated by f^S is $L\theta_{min}$, these not-selected observations would be wrongly classified as outliers. Thus, we must also ensure that S is selected so that $d_S(x) \ge \theta_{min}, \forall x \in I \setminus S$.

3. A RAPID Approximation:

The idea of our approximation is to initialize S = I, which is an efficient and easy-to-implement algorithm to solve the optimization problem, and remove observations from S iteratively as long as S remains feasible.

As input parameters RAPID takes the data set X, the expected outlier percentage p_{out} and a kernel function k. We then iteratively select the most dense observation x_{max} in the current sample S for removal and update the densities.If $S \setminus \{x_{max}\}$ is infeasible,we terminate the process.

As required by the optimization problem, RAPID does not remove boundary points. This is because x_{max} must not be a boundary point, as long as S is not uniform, i.e., $\Delta S_{fit} > 0$. RAPID returns a small sample which has a close-to-uniform density, i.e., a small sample that still obeys the density rule, and also contains the boundary points of the original data.

V. Results

The Results for the paper are as shown in the table.

Technique	Sample percent	Time taken(secs)	Time taken (base paper)[1]
Random	50	0.2425	-
RAPID	6.02	1.84	2
FBPE	0.2	0.015	0.7

Table [1]

These results are for a normal distribution of 10000 samples with 5% outliers.
Because of pruning in pre- processing, there is speedup in time taken in case of FBPE(which is more suitable for smaller datasets) with which we are working.

VI. References

[1]Englhardt, Adrian & Trittenbach, Holger & Kottke, Daniel & Sick, Bernhard & Böhm, Klemens. (2020). Efficient SVDD Sampling with Approximation Guarantees for the Decision Boundary.

[2] Tax, David & Duin, Robert. (2004). Support Vector Data Description. Machine Learning. 54. 45-66. 10.1023/B:MACH.0000008084.60811.49.

[3] https://www.cs.toronto.edu/~kriz/cifar.html

[4] Sadeghi, Reza & Hamidzadeh, Javad. (2018). Automatic support vector data description. Soft Computing. 22. 147–158. 10.1007/s00500-016-2317-5

[5] Sohrab, Fahad & Raitoharju, Jenni & Gabbouj, Moncef & Iosifidis, Alexandros. (2018). Subspace Support Vector Data Description.

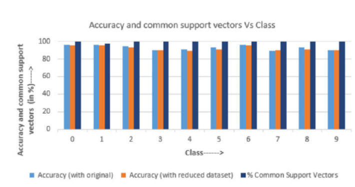
[6] Alam, Shamshe & Sonbhadra, Sanjay & Agarwal, Sonali & Nagabhushan, P. & Tanveer, M.. (2020). Sample reduction using Farthest Boundary Point Estimation (FBPE) for Support Vector Data Description (SVDD). Pattern Recognition Letters. 131. 10.1016/j.patrec.2020.01.004

[7] Sohrab, Fahad & Raitoharju, Jenni& Iosifidis, Alexandros & Gabbouj, Moncef. (2020). Multimodal subspace support vector data description. Pattern Recognition. 110. 10.1016/j.patcog.2020.107648.

[8] Campos, Guilherme & Zimek, Arthur & Sander, Joerg & Campello, Ricardo & Micenková, Barbora & Schubert, Erich & Assent, Ira & Houle, Michael. (2016). On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study. Data Mining and Knowledge Discovery. 30. 10.1007/s10618-015-0444-8.

S. No	Paper name	Domain	Description	Results
1	Englhardt, Adrian & Trittenbach, Holger & Kottke, Daniel & Sick, Bernhard & Böhm, Klemens. (2020). Efficient SVDD Sampling with Approximation Guarantees for the Decision Boundary.	SVDD	Support Vector Data Description (SVDD) is a pop-ular one-class classifiers for anomaly and novelty detection. But despite its effectiveness, SVDD does not scale well with data size.To avoid prohibitive training times, sampling methods select small subsets of the training data on which SVDD trains a decision boundary hopefully equivalent to the one obtained on the full data set. According to the literature, a good sample should therefore contain so-called boundary observations that SVDDwould select as support vectors on the full data set. However,non-boundary observations also are essential to not fragment contiguous inlier regions and avoid poor classification accuracy.Other aspects, such as selecting a sufficiently representative sample, are important as well. But existing sampling methods largely overlook them, resulting in poor classification accuracy.In this article, we study how to select a sample considering these points. Our approach is to frame SVDD sampling as an optimization problem, where constraints guarantee that sampling indeed approximates the original decision boundary. We then propose RAPID, an efficient algorithm to solve this optimization problem. RAPID does not require any tuning of parameters,is easy to implement and scales well to large data sets. We evaluate our approach on real-world and synthetic data. Ourevaluation is the most comprehensive one for SVDD sampling sofar. Our results show that RAPID outperforms its competitors in classification	Data sets with real distributions and more diverse data characteristics. The basis for our experiments are 21 standard benchmark data sets for outlier detection .Campos et al. constructed this benchmark from classification data where one of the classes is downsampled and labeled as outlier. The data sets have different sizes (80to49 534observations), dimensionality (3 to 1555dimensions) and outlier ratios (0.2 % to75.38 %, median 9.12 %).

<table>
<tr><td></td><td></td><td></td><td>accuracy, in sample size, and in runtime.Index Terms—One-class Classification, Data Reduction, Out-lier Detection, Anomaly Detection</td><td></td></tr>
<tr><td>2</td><td>Tax, David & Duin, Robert. (2004). Support Vector Data Description. Machine Learning. 54. 45-66. 10.1023/B:MACH.0000008084.60811.49.</td><td>Original SVDD proposal</td><td>Data domain description concerns the characterization of a data set. A good description covers all target data but includes no superfluous space. The boundary of a dataset can be used to detect novel data or outliers. We will present the Support Vector Data Description (SVDD) which is inspired by the Support Vector Classifier. It obtains a spherically shaped boundary around a dataset and analogous to the Support Vector Classifier it can be made flexible by using other kernel functions. The method is made robust against outliers in the training set and is capable of tightening the description by using negative examples. We show characteristics of the Support Vector Data Descriptions using artificial and real data.</td><td><table>
<tr><td></td><td colspan="6">Number of features</td></tr>
<tr><td>Method</td><td>3</td><td>5</td><td>10</td><td>20</td><td>30</td><td>64</td></tr>
<tr><td>Normal dens.</td><td>38.1</td><td>37.8</td><td>34.1</td><td>25.9</td><td>16.6</td><td>4.5</td></tr>
<tr><td>Mix.o.Gauss.</td><td>18.5</td><td>21.0</td><td>9.8</td><td>11.4</td><td>14.4</td><td>15.2</td></tr>
<tr><td>Parzen dens.</td><td>20.8</td><td>45.0</td><td>45.0</td><td>45.0</td><td>45.0</td><td>45.0</td></tr>
<tr><td>kNN</td><td>16.5</td><td>13.8</td><td>12.1</td><td>19.9</td><td>22.5</td><td>30.4</td></tr>
<tr><td>SVDD</td><td>15.5</td><td>14.3</td><td>11.6</td><td>10.9</td><td>9.9</td><td>4.9</td></tr>
</table>
Best performances from the methods are shown in bold.
fig(1)[2]
SVDD outperforms other techniques in 4 out of 6 cases.</td></tr>
<tr><td>3</td><td>Sadeghi, Reza & Hamidzadeh, Javad. (2018). Automatic support vector data description. Soft Computing. 22. 147–158. 10.1007/s00500-016-2317-5</td><td>ASVDD</td><td>Event handlers have a wide range of applications such as medical assistant systems and fire suppression systems. These systems try to provide accurate responses based on the least information. Support vector data description (SVDD) is one of the appropriate tools for such detections, which should handle lack of information. Therefore, many efforts have been done to improve SVDD. Unfortunately, the existing descriptors suffer from weak data characteristic in sparse data sets and their tuning parameters are organized improperly. These issues cause reduction of accuracy in event handlers when they are faced with data shortage. Therefore, we propose automatic support vector data description (ASVDD) based on both validation degree, which originated from fuzzy rough sets to discover data</td><td>Consequently, the results reveal that the out-come of ASVDD is meaningfully distinguishable from that of state-of-the-art methods,</td></tr>
</table>

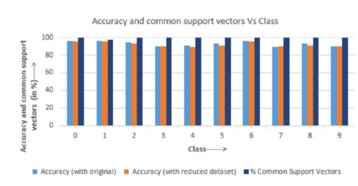
			characteristics, and assigning effective values for tuning parameters by chaotic bat algorithm. To evaluate the performance of ASVDD, several experiments have been conducted on various data sets of UCI repository. The experimental results demonstrate superiority of the proposed method over state-of-the-art ones in terms of classification accuracy and AUC. In order to prove meaningful distinction between the accuracy results of the proposed method and the leading-edge ones, the Wilcoxon statistical test has been conducted.	
4	Sohrab, Fahad & Raitoharju, Jenni & Gabbouj, Moncef & Iosifidis, Alexandros. (2018). Subspace Support Vector Data Description.	Subspace SVDD	This paper proposes a novel method for solving one-class classification problems. The proposed approach, namely Subspace Support Vector Data Description, maps the data to a subspace that is optimized for one-class classification. In that feature space, the optimal hypersphere enclosing the target class is then determined. The method iteratively optimizes the data mapping along with data description in order to define a compact class representation in a low-dimensional feature space. We provide both linear and non-linear mappings for the proposed method. Experiments on 14 publicly available datasets indicate that the proposed Subspace Support Vector Data Description provides better performance compared to baselines and other recently proposed one-class classification methods.	A tighter fitted hypersphere on the training data may possibly lead to more meaningful results,which could be achieved by restricting the range of the C values used during the cross-validation process applied on the training data for hyper-parameter selection of the proposed method. the linear version of the proposedS-SVDD clearly outperforms all other linear methods. Only For dataset 10, OC-SVM achieves a higher performance Refer Table 2 and 3 in [5] for detailed results
5	Alam, Shamshe & Sonbhadra, Sanjay & Agarwal, Sonali & Nagabhushan, P. & Tanveer, M.. (2020). Sample reduction using Farthest Boundary Point Estimation (FBPE) for Support Vector Data	Boundary estimation for SVDD	The objective of this paper is to design an algorithm to maximize the learning ability and knowledge about the target class while minimizing the number of training samples for support vector data description (SVDD). With this motivation, a novel training sample reduction algorithm is proposed in this paper that selects the most promising boundary data points	 Fig2[7] Indian pines dataset has been used to evaluate the performance of proposed

	Description (SVDD). Pattern Recognition Letters. 131. 10.1016/j.patrec.2020.01.004		as a training set. The proposed approach uses the local geometry of the distribution to estimate the farthest boundary points (also known as extreme points). The legitimacy of the proposed algorithm is verified via experiments performed on MNIST, Iris, UCI default credit card, svmguide and Indian Pines datasets.	method where 70% samples are consid-ered as training samples. It is evident that the average number of reduced training samples is 37% and the accuracy achieved with reduced training set is 0.02% less than compared to the original one[5].
6	Sohrab, Fahad & Raitoharju, Jenni & Iosifidis, Alexandros & Gabbouj, Moncef. (2020). Multimodal subspace support vector data description. Pattern Recognition. 110. 10.1016/j.patcog.2020.107648.	Multimodal subspace support vector data	In this paper, we propose a novel method for projecting data from multiple modalities to a new subspace optimized for one-class classification. The proposed method iteratively transforms the data from the original feature space of each modality to a new common feature space along with finding a joint compact description of data coming from all the modalities. For data in each modality, we define a separate transformation to map the data from the corresponding feature space to the new optimized subspace by exploiting the available information from the class of interest only. We also propose different regularization strategies for the proposed method and provide both linear and non-linear formulations. The proposed Multimodal Subspace Support Vector Data Description outperforms all the competing methods using data from a single modality or fusing data from all modalities in four out of five datasets.	We compare the results for different variant of MS-SVDD in Tables 1–5 of the supplementary material. Overall in all datasets, NPT is found to be more robust than the kernel version. Linear MS-SVDD is found to perform best over 2 datasets, similar to the NPT version, which performs best on two datasets as well. The kernel MS-SVDD performs best on one out of five datasets as compared to linear and NPT version of MS-SVDD

S. No	Paper name	Domain	Description	Results
1	Englhardt, Adrian & Trittenbach, Holger & Kottke, Daniel & Sick, Bernhard & Böhm, Klemens. (2020). Efficient SVDD Sampling with Approximation Guarantees for the	SVDD	Support Vector Data Description (SVDD) is a pop-ular one-class classifiers for anomaly and novelty detection. But despite its effectiveness, SVDD does not scale well with data size.To avoid prohibitive training times, sampling methods select small subsets of the training data on which SVDD trains a	Data sets with real distributions and more diverse data characteristics. The basis for our experiments are 21 standard benchmark data sets for outlier detection .Campos et al. constructed this benchmark from

<table>
<tr>
<td></td>
<td>Decision Boundary.</td>
<td></td>
<td>decision boundary hopefully equivalent to the one obtained on the full data set. According to the literature, a good sample should therefore contain so-called boundary observations that SVDDwould select as support vectors on the full data set. However,non-boundary observations also are essential to not fragment contiguous inlier regions and avoid poor classification accuracy.Other aspects, such as selecting a sufficiently representative sample, are important as well. But existing sampling methods largely overlook them, resulting in poor classification accuracy.In this article, we study how to select a sample considering these points. Our approach is to frame SVDD sampling as an optimization problem, where constraints guarantee that sampling indeed approximates the original decision boundary. We then propose RAPID, an efficient algorithm to solve this optimization problem. RAPID does not require any tuning of parameters,is easy to implement and scales well to large data sets. We evaluate our approach on real-world and synthetic data. Ourevaluation is the most comprehensive one for SVDD sampling sofar. Our results show that RAPID outperforms its competitors in classification accuracy, in sample size, and in runtime.Index Terms—One-class Classification, Data Reduction, Out-lier Detection, Anomaly Detection</td>
<td>classification data where one of the classes is downsampled and labeled as outlier. The data sets have different sizes (80to49 534observations), dimensionality (3 to 1555dimensions) and outlier ratios (0.2 % to75.38 %, median 9.12 %).</td>
</tr>
<tr>
<td>2</td>
<td>Tax, David & Duin, Robert. (2004). Support Vector Data Description. Machine Learning. 54. 45-66. 10.1023/B:MACH.0000008084.60811.49.</td>
<td>Original SVDD proposal</td>
<td>Data domain description concerns the characterization of a data set. A good description covers all target data but includes no superfluous space. The boundary of a dataset can be used to detect novel data or outliers. We will present the Support Vector Data</td>
<td>
<table>
<tr><th></th><th>Number of features</th><th></th><th></th><th></th><th></th><th></th></tr>
<tr><th>Method</th><th>3</th><th>5</th><th>10</th><th>20</th><th>30</th><th>64</th></tr>
<tr><td>Normal dens.</td><td>38.1</td><td>37.8</td><td>34.1</td><td>25.9</td><td>16.6</td><td>4.5</td></tr>
<tr><td>Mix.o.Gauss.</td><td>18.5</td><td>21.0</td><td>9.8</td><td>11.4</td><td>14.4</td><td>15.2</td></tr>
<tr><td>Parzen dens.</td><td>20.8</td><td>45.0</td><td>45.0</td><td>45.0</td><td>45.0</td><td>45.0</td></tr>
<tr><td>kNN</td><td>16.5</td><td>13.8</td><td>12.1</td><td>19.9</td><td>22.5</td><td>30.4</td></tr>
<tr><td>SVDD</td><td>15.5</td><td>14.3</td><td>11.6</td><td>10.9</td><td>9.9</td><td>4.9</td></tr>
</table>
Best performances from the methods are shown in bold.
fig(1)[2]
</td>
</tr>
</table>

			Description (SVDD) which is inspired by the Support Vector Classifier. It obtains a spherically shaped boundary around a dataset and analogous to the Support Vector Classifier it can be made flexible by using other kernel functions. The method is made robust against outliers in the training set and is capable of tightening the description by using negative examples. We show characteristics of the Support Vector Data Descriptions using artificial and real data.	SVDD outperforms other techniques in 4 out of 6 cases.
3	Sadeghi, Reza & Hamidzadeh, Javad. (2018). Automatic support vector data description. Soft Computing. 22. 147–158. 10.1007/s00500-016-2317-5	ASVDD	Event handlers have a wide range of applications such as medical assistant systems and fire suppression systems. These systems try to provide accurate responses based on the least information. Support vector data description (SVDD) is one of the appropriate tools for such detections, which should handle lack of information. Therefore, many efforts have been done to improve SVDD. Unfortunately, the existing descriptors suffer from weak data characteristic in sparse data sets and their tuning parameters are organized improperly. These issues cause reduction of accuracy in event handlers when they are faced with data shortage. Therefore, we propose automatic support vector data description (ASVDD) based on both validation degree, which originated from fuzzy rough sets to discover data characteristics, and assigning effective values for tuning parameters by chaotic bat algorithm. To evaluate the performance of ASVDD, several experiments have been conducted on various data sets of UCI repository. The experimental results demonstrate superiority of the proposed method over state-of-the-art ones in terms of classification accuracy and AUC. In order to prove meaningful distinction between the accuracy results of the proposed method and the	Consequently, the results reveal that the out-come of ASVDD is meaningfully distinguishable from that of state-of-the-art methods,

			leading-edge ones, the Wilcoxon statistical test has been conducted.	
4	Sohrab, Fahad & Raitoharju, Jenni & Gabbouj, Moncef & Iosifidis, Alexandros. (2018). Subspace Support Vector Data Description.	Subspace SVDD	This paper proposes a novel method for solving one-class classification problems. The proposed approach, namely Subspace Support Vector Data Description, maps the data to a subspace that is optimized for one-class classification. In that feature space, the optimal hypersphere enclosing the target class is then determined. The method iteratively optimizes the data mapping along with data description in order to define a compact class representation in a low-dimensional feature space. We provide both linear and non-linear mappings for the proposed method. Experiments on 14 publicly available datasets indicate that the proposed Subspace Support Vector Data Description provides better performance compared to baselines and other recently proposed one-class classification methods.	A tighter fitted hypersphere on the training data may possibly lead to more meaningful results,which could be achieved by restricting the range of the C values used during the cross-validation process applied on the training data for hyper-parameter selection of the proposed method. the linear version of the proposedS-SVDD clearly outperforms all other linear methods. Only For dataset 10, OC-SVM achieves a higher performance Refer Table 2 and 3 in [5] for detailed results
5	Alam, Shamshe & Sonbhadra, Sanjay & Agarwal, Sonali & Nagabhushan, P. & Tanveer, M.. (2020). Sample reduction using Farthest Boundary Point Estimation (FBPE) for Support Vector Data Description (SVDD). Pattern Recognition Letters. 131. 10.1016/j.patrec.2020.01.004	Boundary estimation for SVDD	The objective of this paper is to design an algorithm to maximize the learning ability and knowledge about the target class while minimizing the number of training samples for support vector data description (SVDD). With this motivation, a novel training sample reduction algorithm is proposed in this paper that selects the most promising boundary data points as a training set. The proposed approach uses the local geometry of the distribution to estimate the farthest boundary points (also known as extreme points). The legitimacy of the proposed algorithm is verified via experiments performed on MNIST, Iris, UCI default credit card, svmguide and Indian Pines datasets.	 Fig2[7] Indian pines dataset has been used to evaluate the performance of proposed method where 70% samples are consid-ered as training samples. It is evident that the average number of reduced training samples is 37% and the accuracy achieved with reduced training set is 0.02% less than compared to the original one[5].
6	Sohrab, Fahad & Raitoharju, Jenni &	Multimodal subspace	In this paper, we propose a novel method for projecting data from	We compare the results for different variant of

	Iosifidis, Alexandros & Gabbouj, Moncef. (2020). Multimodal subspace support vector data description. Pattern Recognition. 110. 10.1016/j.patcog.2020.107648.	support vector data	multiple modalities to a new subspace optimized for one-class classification. The proposed method iteratively transforms the data from the original feature space of each modality to a new common feature space along with finding a joint compact description of data coming from all the modalities. For data in each modality, we define a separate transformation to map the data from the corresponding feature space to the new optimized subspace by exploiting the available information from the class of interest only. We also propose different regularization strategies for the proposed method and provide both linear and non-linear formulations. The proposed Multimodal Subspace Support Vector Data Description outperforms all the competing methods using data from a single modality or fusing data from all modalities in four out of five datasets.	MS-SVDD in Tables 1–5 of the supplementary material. Overall in all datasets, NPT is found to be more robust than the kernel version. Linear MS-SVDD is found to perform best over 2 datasets, similar to the NPT version, which performs best on two datasets as well. The kernel MS-SVDD performs best on one out of five datasets as compared to linear and NPT version of MS-SVDD