ECAI 2008
M. Ghallab et al. (Eds.)
IOS Press, 2008

doi:10.3233/978-1-58603-891-5-107

Prototype-based Domain Description

Fabrizio Angiulli[1]

Abstract. In this work a novel one-class classifier, namely the Prototype-based Domain Description rule (PDD), is presented. The PDD classifier is equivalent to the NNDD rule under the infinity Minkowski metric for a suitable choice of the prototype set. The concept of PDD consistent subset is introduced and it is shown that computing a minimum size PDD consistent subset is in general not approximable within any constant factor. A logarithmic approximation factor algorithm, called the CPDD algorithm, for computing a minimum size PDD consistent subset is then introduced. The CPDD algorithm has some parameters which allow to tune the trade off between accuracy and size of the model. Experimental results show that the CPDD rule sensibly improves over the CNNDD classifier in terms of size of the subset, while guaranteeing a comparable classification quality.

1 INTRODUCTION

Domain description, or one-class classification, is a classification technique whose goal is to distinguish between objects belonging to a certain class and all the other objects of the space [11]. The Nearest Neighbor Domain Description rule (NNDD) [1] is a one-class classifier accepting test objects whose nearest neighbors distances in a reference data set, assumed to model normal behavior, lie within a certain threshold. In particular, given a data set of objects, also called reference set, and two parameters k and θ, the NNDD associates a feature vector $\delta(x) \in \mathbb{R}^k$ with each object x composed of the distances from x to its first k nearest neighbors in the reference set. The classifier accepts x if and only if $\delta(x)$ belongs to the hyper-sphere (according to one of the L_r Minkowski metric, $r \in \{1, 2, \ldots, \infty\}$) centered in the origin of $\mathbb{R}^k$ and having radius θ, i.e. if and only if $\|\delta(x)\|_r \leq \theta$. The CNNDD rule is a variant of the NNDD rule using a selected subset of the data set as the reference set [1].

In this work a novel nearest neighbor based one-class classifier, called the Prototype-based Nearest Neighbor classifier (PDD), is introduced. A *prototype set* is a set of objects x_i, also called *prototypes*, each of which is associated with a radius $R(x_i)$. Given parameter θ, an object y is accepted if it lies within distance $\theta - R(x_i)$ from some prototype x_i. It is shown that the PDD classifier is equivalent to the NNDD rule under the infinity Minkowski metric (that is for $r = \infty$) for a suitable choice of the prototype set. Then the concept of PDD consistent subset is introduced, that is a subset of the original prototype set, which, loosely speaking, accepts all the discarded prototypes. It is shown that computing a minimum size PDD consistent subset is in general not approximable within any constant factor. A logarithmic approximation factor algorithm, called the CPDD algorithm, for computing a minimum size PDD consistent subset is then introduced. The CPDD algorithm has some parameters which allow to tune the trade off between accuracy and size of the model. Experimental results show that the CPDD rule sensibly improves over the CNNDD classifier in terms of size of the subset, while guaranteeing a comparable classification quality. Moreover, comparison with the one-class SVM classifier points out that both the compression ratio and the accuracy of the CPDD are comparable to that of the one-class SVM classifier, but with some advantages for the CPDD rule.

The rest of the work is organized as follows. Section 2 defines the Prototype-based Domain Description rule (PDD) and the concept of PDD consistent subset. Section 3 investigates the computational complexity of the problem of computing a minimum size PDD consistent subset. Section 4 describes the CPDD rule. Section 5 presents experimental results. Finally, Section 6 presents conclusions and future work.

[1] DEIS - University of Calabria, Italy, email: f.angiulli@deis.unical.it

2 THE PROTOTYPE-BASED DOMAIN DESCRIPTION RULE

In the following U denotes a set of objects, d a distance metric on U, D a set of objects from U, k a positive integer number, θ a positive real number, and $r \in \{1, 2, \ldots, \infty\}$ a Minkowski metric L_r.

A *prototype set* P is a set of pairs $P = \{\langle x_1, r_1\rangle, \ldots, \langle x_n, r_n\rangle\}$, where each x_i ($1 \leq i \leq n$) is an object of U, also called *prototype*, and each r_i is a real number, also called *prototype radius*. Given a prototype x_i, the prototype radius r_i associated with x_i is also denoted by $R(x_i)$.

Next the Prototype-based Domain Description one-class classifier is defined.

Definition 2.1 *Given a prototype set P, the* Prototype-based Domain Description rule *(*PDD*) according to P,* d*, and θ, is the function* $\mathrm{PDD}_{P,\mathrm{d},\theta}$ *from U to $\{-1, +1\}$ such that*

$$\mathrm{PDD}_{P,\mathrm{d},\theta}(y) = \begin{cases} +1, & \text{if } \exists x \in P \text{ such that } \mathrm{d}(x, y) + R(x) \leq \theta \\ -1, & \text{otherwise} \end{cases}$$

The PDD rule accepts an input object y (that is returns the value $+1$) if y lies within distance $R(x_i)$ from some prototype x_i.

The PDD rule is a nearest neighbor based one-class classifier. Next the definition of another nearest neighbor based one-class classifier, namely the NNDD rule, is recalled, and then the relationship between there two rules is pointed out.

Given an object x of U, the *k-th nearest neighbor* $nn_{D,\mathrm{d},k}(x)$ of x in D according to d is the object y of D such that there exist exactly $k-1$ objects z of D with $\mathrm{d}(x, z) \leq \mathrm{d}(x, y)$. If $x \in D$, then $nn_{D,\mathrm{d},1}(x) = x$.

The *k nearest neighbors distances vector* $\delta_{D,\mathrm{d},k}(p)$ of p in D is

$$\delta_{D,\mathrm{d},k}(p) = (\mathrm{d}(p, nn_{D,\mathrm{d},1}(p)), \ldots, \mathrm{d}(p, nn_{D,\mathrm{d},k}(p))).$$

Definition 2.2 ([1]) *The* Nearest Neighbor Domain Description rule *(NNDD) according to $D, \mathrm{d}, k, \theta, r$, is the function* $\mathrm{NNDD}_{D,\mathrm{d},k,\theta,r}$

from U to $\{-1,+1\}$ such that

$$\mathrm{NNDD}_{D,\mathrm{d},k,\theta,r}(p) = \mathrm{sign}(\theta - \|\delta_{D,\mathrm{d},k}(p)\|_r),$$

where $\mathrm{sign}(x) = -1$ *if* $x < 0$*, and* $\mathrm{sign}(x) = 1$ *otherwise.*

The following definition relates the PDD rule and the NNDD rule. Given a set of objects D, the *prototype set* $P(D,\mathrm{d},k,\theta)$ *associated with* D *w.r.t.* d*,* k*, and* θ is

$$\{\langle x, \mathrm{d}(x, nn_{D,\mathrm{d},k}(x))\rangle \mid x \in D \wedge \mathrm{d}(x, nn_{D,\mathrm{d},k}(x)) \leq \theta\}.$$

Relationship between the two rules is clarified by the theorem below.

Theorem 1 *Given a set of objects D, and parameters k and θ, it holds that*

$$(\forall x \in D)(\mathrm{NNDD}_{D,\mathrm{d},k,\theta,+\infty}(x) = \mathrm{PDD}_{P(D,\mathrm{d},k,\theta),\mathrm{d},\theta}(x)).$$

Proof. Let x be a generic object of D.

If $\mathrm{d}(x, nn_{D,\mathrm{d},k}(x)) \leq \theta$, then $\mathrm{NNDD}_{D,\mathrm{d},k,\theta,+\infty}(x) = \mathrm{sign}(\theta - \|\delta_{D,\mathrm{d},k}(p)\|_{+\infty}) = \mathrm{sign}(\theta - \mathrm{d}(x, nn_{D,\mathrm{d},k}(x))) = 1$. Furthermore, the pair $\langle x, \mathrm{d}(x, nn_{D,\mathrm{d},k}(x))\rangle$ belongs to $P(D,\mathrm{d},k,\theta)$ and, hence, $\mathrm{d}(x,x) + R(x) = 0 + \mathrm{d}(x, nn_{D,\mathrm{d},k}(x)) \leq \theta$ and $\mathrm{PDD}_{P(D,\mathrm{d},k,\theta),\mathrm{d},\theta}(x) = 1$.

If $\mathrm{d}(x, nn_{D,\mathrm{d},k}(x)) > \theta$, then $\mathrm{NNDD}_{D,\mathrm{d},k,\theta,+\infty}(x) = -1$. By contradiction, assume that there exists a pair $\langle y, R(y)\rangle$ in $P(D,\mathrm{d},k,\theta)$ such that $\mathrm{d}(x,y) + R(y) \leq \theta$. Since within distance $r_y = \mathrm{d}(x,y) + \mathrm{d}(y, nn_{D,\mathrm{d},k}(y))$ from x there are at least $k+1$ objects of D, it holds that $\mathrm{d}(x, nn_{D,\mathrm{d},k}(x)) \leq r_y \leq \theta$, which contradicts the hypothesis. □

Thus, from the point of view of the objects belonging to the data set D, the prototype set $P(D,\mathrm{d},\theta,k)$ is the analogue for the PDD rule of the data set D for the NNDD rule.

When the reference set D is large, space requirements to store D and time requirements to find the nearest neighbors of an object in D increase. In the spirit of the reference set thinning problem for the k-NN-rule [9, 2], the concept of NNDD reference consistent subset was defined in [1]. In the same spirit, next it is provided the definition of PDD consistent subset.

Let P be a prototype set and let S be a subset of P. The set S is said to be a PDD *consistent subset* of P with respect to d and θ, if the following relationship hold

$$(\forall \langle x, r\rangle \in P)(\mathrm{PDD}_{P,\mathrm{d},\theta}(x) = \mathrm{PDD}_{S,\mathrm{d},\theta}(x)).$$

Importantly, it also holds that a PDD consistent subset S of the set $P(D,\mathrm{d},\theta,k)$ is the analogue for the PDD rule of the data set D for the NNDD rule. It can be finally concluded from the concept of sample compression scheme [7] and from the discussion above that replacing the prototype set P with a consistent subset S of P improves both response time and generalization.

3 COMPLEXITY ANALYSIS

In this section the computational complexity of the problem of computing a minimum size PDD consistent subset is investigated. The reader is referred to [8, 3] for basics on complexity theory, NP optimization problems, and approximation algorithms. Next it is shown that, in the general case, the problem of computing a minimum size PDD consistent subset is not in the APX complexity class, which is, loosely speaking, the class of the NP optimization problems whose optimal solution can be approximated in polynomial time within a fixed factor.

Algorithm CPDD

1. for each object x_i in D, determine the distance r_i between x_i and its k-th nearest neighbor in D
2. for each object x_i such that $r_i \leq \theta$, determine the set N_i composed of the objects y of D such that $\mathrm{d}(x_i,y) + r_i \leq \varrho\theta$
3. set P to $\{x_i \in D \mid r_i \leq \theta\}$, and set S and C to the emptyset
4. while $|C| \leq \eta|P|$ do
   (a) determine the object x_j of P such that (break ties in favor of the object such that the value r_j is minimum)
   $$|N_j - C| = \max\{|N_i - C| : x_i \in P\}$$
   (b) set S to $S \cup \{\langle x_j, r_j\rangle\}$, and C to $C \cup N_j$
5. return the set S

Figure 1. The CPDD algorithm.

Given a prototype set P, distance metric d, and a positive real number θ, the PDD *Consistent Subset Problem* $\langle P,\mathrm{d},\theta\rangle$ is defined as follows: compute a PDD consistent subset S^* of P with respect to d and θ, also said a *minimum size* PDD *consistent subset*, such that, for each PDD consistent subset S of P with respect to d and θ, $|S^*| \leq |S|$.

Given a positive integer m, the *decision version* $\langle P,\mathrm{d},\theta,m\rangle_D$ of the problem $\langle P,\mathrm{d},\theta\rangle$ is defined as follows: reply "yes" if there exists a PDD consistent subset S of P with respect to d and θ such that $|S| \leq m$, and reply "no" otherwise.

Theorem 2 *The $\langle P,\mathrm{d},\theta\rangle$ problem (1) is* NP*-hard, and (2) is not in* APX.

Proof sketch. (Point 1) Membership is immediate. As for the hardness the proof is by reduction of the *Dominating Set Problem* [8]. Let $G = (V,E)$ be an undirected graph, and let $m \leq |V|$ be a positive integer. The *Dominating Set Problem* is: is there a subset $U \subseteq V$, called *dominating set* of G, with $|U| \leq m$, such that for all $v \in (V - U)$ there exists $u \in U$ with $\{u,v\} \in E$?

Define the metric d_V on the set V of nodes of G as follows: $\mathrm{d}_V(u,v) = \theta$, if $\{u,v\} \in E$, and $\mathrm{d}_V(u,v) = 2\theta$, otherwise. Let P_V be the set $\{\langle v,0\rangle \mid v \in V\}$. It can be proved that G has a dominating set of size m if and only if $\langle P_V, \mathrm{d}_V, \theta, m\rangle_D$ is a "yes" instance.

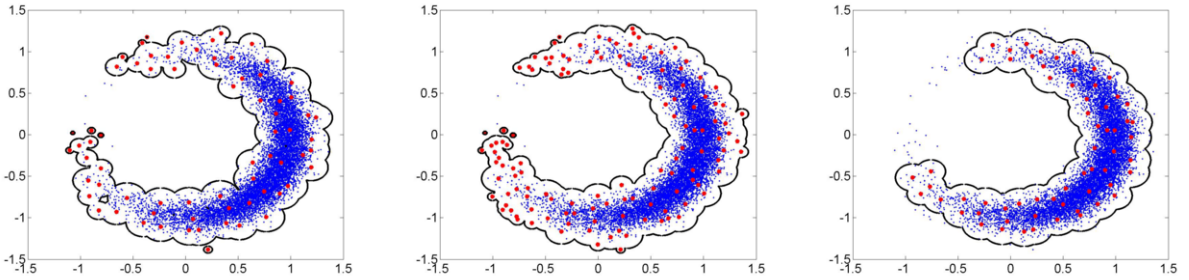
The NP-hardness of the $\langle P,\mathrm{d},\theta\rangle$ problem follows immediately from the NP-completeness of its decision version.

(Point 2) It is known that the *Minimum Dominating Set Problem*, that is the problem of determining the size of the smallest dominating set of a graph, is not in APX [4]. We note that Point 1 of this theorem defines an AP-reduction from the *Minimum Dominating Set Problem* to the *Minimum* PDD *Consistent Subset Problem* (the reader is referred to [3] for the definition of AP-reduction). As an immediate consequence of this reduction, the latter problem does not belong to APX. □

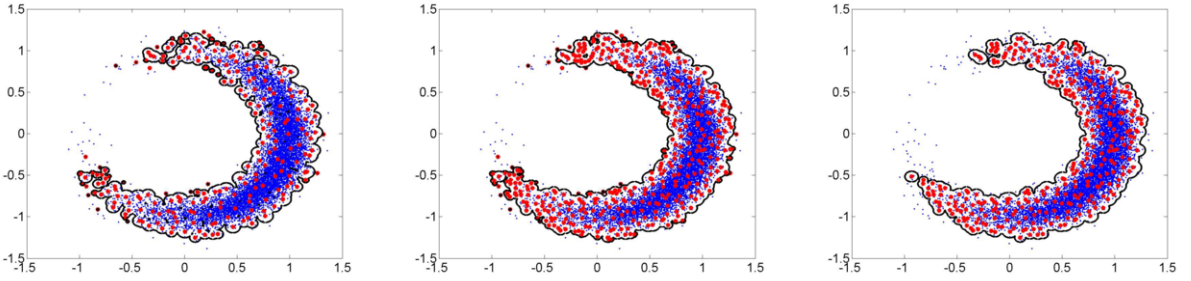
4 THE CPDD ALGORITHM

Figure 1 shows the algorithm CPDD. Given a data set D, the CPDD algorithm computes a PDD consistent subset of the prototype set $P(D,\mathrm{d},k,\theta)$ associated with D.

The algorithm receives in input a data set D, parameters k and θ, and the additional parameters $\varrho,\eta \in (0,1]$, whose use will be



(a) $\theta = 0.2$, $\varrho = 1.00$, $\eta = 1.00$ (b) $\theta = 0.2$, $\varrho = 0.75$, $\eta = 1.00$ (c) $\theta = 0.2$, $\varrho = 0.75$, $\eta = 0.99$



(d) $\theta = 0.1$, $\varrho = 1.00$, $\eta = 1.00$ (e) $\theta = 0.1$, $\varrho = 0.75$, $\eta = 1.00$ (f) $\theta = 0.1$, $\varrho = 0.75$, $\eta = 0.99$

Figure 2. Examples of PDD consistent subsets computed by the CPDD algorithm.

discussed in the following (if not otherwise specified, it is assumed that ϱ and η are both set to one).

Initially, for each object x_i of D, the algorithm determines the distance r_i to its k-th nearest neighbor (step 1) and also the set N_i of the objects of D lying within distance $\theta - r_i$ from it (step 2). The set P built in step 3 is composed of the objects occurring in the prototype set $P(D, \mathrm{d}, k, \theta)$.

Then the algorithm computes the consistent subset S following a greedy strategy (step 4). The set C consists of the objects of P which are correctly classified by the current subset S. At each step, the object x_j maximizing the number of objects in $N_j - C$ is selected and inserted in S, until C contains at least the fraction η of the objects in P (until C covers P, if $\eta = 1$).

Next theorem shows that the the size of the solution returned by the algorithm has an approximation factor.

Theorem 3 *The CPDD algorithm provides a solution having a* $1 + \ln(n)$ *approximation factor.*

Proof. Assume that the parameter ϱ is set to one. We note that the set N_i consists of precisely all the prototypes of $P(D, \mathrm{d}, \theta, k)$ which are correctly recognized through the PDD rule if x_i is included in the PDD consistent subset S.

Given a finite set $\mathcal{S}$ and a collection $\mathcal{C}$ of subsets of $\mathcal{S}$, a *set cover* for $\mathcal{S}$ is a subset $\mathcal{C}'$ of $\mathcal{C}$ such that every element in $\mathcal{S}$ belongs to at least one member of $\mathcal{C}'$. It is clear that the PDD consistent subsets of P are in one-to-one correspondence with the set covers of $\{N_i \mid x_i \in P\}$. The result hence follows by noting that step 4 of the algorithm CPDD is analogous to the greedy algorithm for the *Minimum Set Cover Problem* [6], the problem of computing a set cover of minimum size, which achieves an approximation factor of $1 + \ln(n)$, where n is the size of the set to be covered. □

Note that steps 3-5 compute a PDD consistent subset of any arbitrary prototype set.

Figure 2 reports some examples of PDD consistent subsets computed by the CPDD algorithm. The data set (blue points) is composed of ten thousands points in the plane. The parameter k was set to 5, while two distinct values for the parameters θ, ϱ and η were considered, namely 0.1 and 0.2 for θ, 0.75 and 1.0 for ϱ, and 0.99 and 1.0 for η. The Euclidean distance was employed as distance function d.

Stars (in red color) denote the prototypes belonging to the PDD consistent subset S, while the (black) curve denotes the decision boundary of the classifier $\mathrm{PDD}_{S,\mathrm{d},\theta}$. The relative size of the PDD consistent subsets reported in Figure 2 is summarized in the following table.

	$\varrho = 1.00$ $\eta = 1.00$	$\varrho = 0.75$ $\eta = 1.00$	$\varrho = 0.75$ $\eta = 0.99$
$\theta = 0.2$	70 (0.7%)	128 (1.3%)	62 (0.6%)
$\theta = 0.1$	227 (2.3%)	439 (4.4%)	337 (3.4%)

From the figure and the table above it is clear that the smaller the value of the parameter θ, the closer the class boundary to the data set shape, the greater the number of data set objects rejected by the PDD rule, and the greater the number of prototypes belonging to the consistent subset.

Moreover, the smaller the value of the parameter ϱ, the greater the number of prototypes belonging to the consistent subset, the more accurate the form of the decision boundary, and the smaller the probability of rejecting objects belonging to the class represented by the data set. For example, in Figure 2(a) ($\varrho = 1$) there is a "hole", approximately centered in $(-0.78, -0.78)$, in the lower tail of the data set (but also other smaller "holes" exist along the data set shape), while the same region is covered by the prototypes in Figure 2(b) ($\varrho = 0.75$).

Finally, the smaller the value of the parameter η, the smaller the

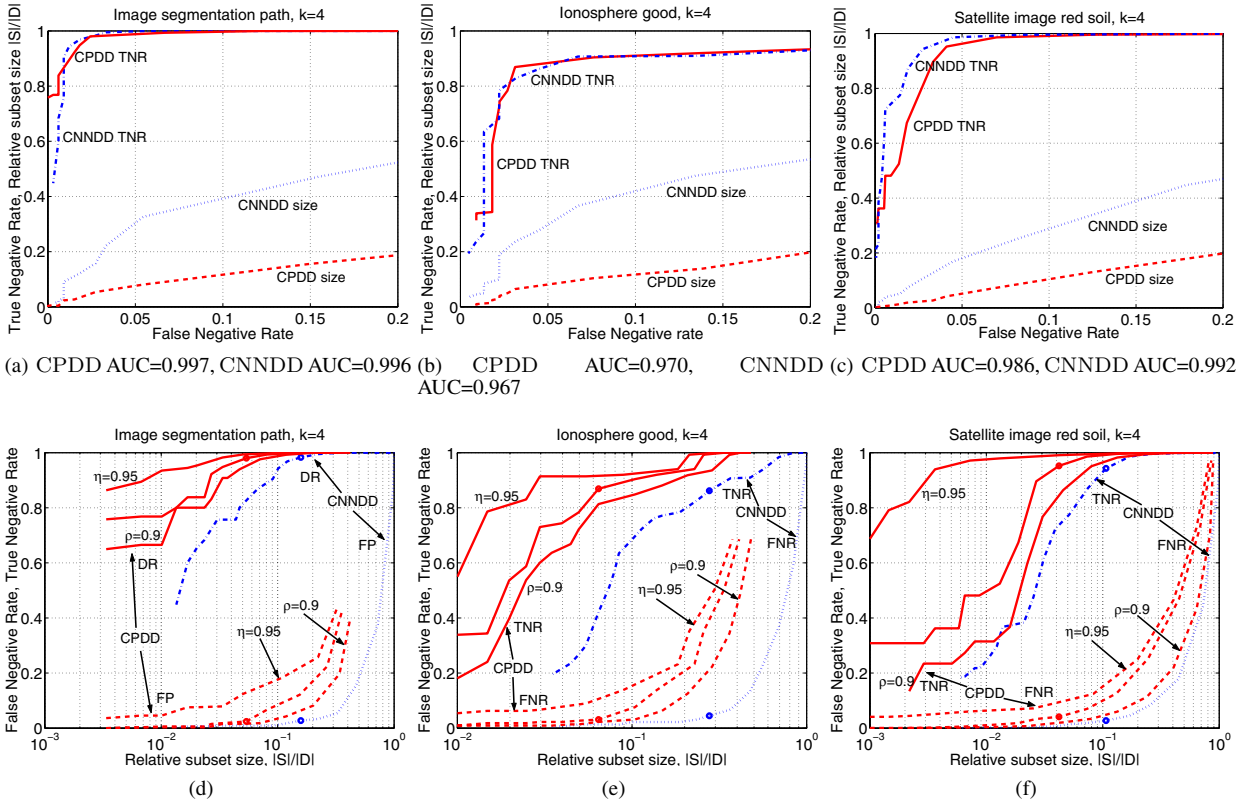


Figure 3. Comparison between the CPDD and the CNNDD rule.

number of prototypes belonging to the consistent subset, but the greater the probability of rejecting objects belonging to the class represented by the data set, since the most sparse regions of the feature space belonging to the class are left uncovered.

5 EXPERIMENTAL RESULTS

In this section, experiments involving the CPDD rule on three data sets from the UCI Machine Learning Repository, namely *Image segmentation*, *Ionosphere*, and *Satellite image*, are described.[2] In particular, for the *Image segmentation* data set (19 attributes) the *path* class (330 objects) was considered the normal one, while the remaining 1,980 objects were considered anomalies, for the *Ionosphere* data set (34 attributes) the *good* class (225 objects) was considered the normal one, while the objects of the *bad* class were considered anomalies, and for the *Satellite image* (36 attributes) the *red soil* class (1,533 objects) was considered the normal one, while the remaining 3,902 objects were considered anomalies.

Figure 3 reports comparison of the CPDD and CNNDD (for $r = +\infty$) rules on the three considered data sets. The parameter k was set to 4 in all the experiments, while the parameter θ was varied from zero to a suitable large value, and, then, the size of the subset computed, the false negative rate, and the true negative rate, were measured. If not otherwise specified, the parameters ϱ and η are set to 1. The Euclidean distance was employed as distance function d.

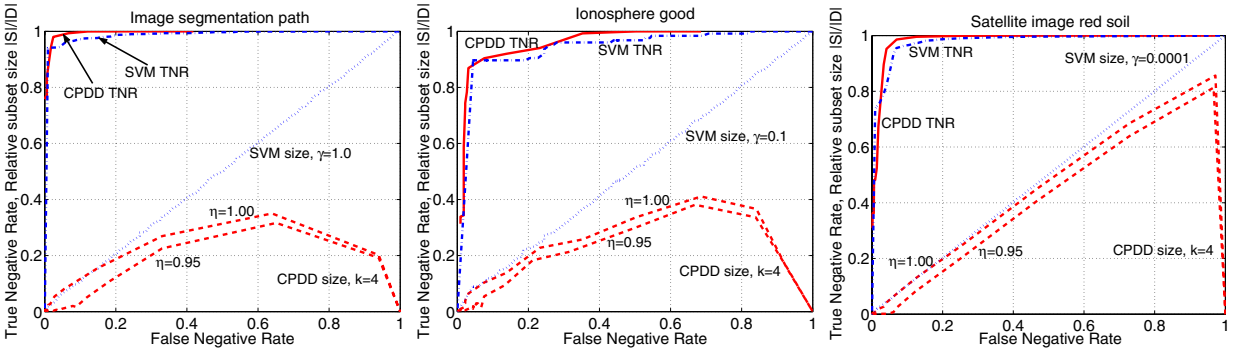
The *True Positive Rate* (TPR, for short) is the fraction of normal objects accepted by the classifier, while the *False Positive Rate* (FPR, for short) is the fraction of abnormal objects accepted by the classifier. Dually, the *False Negative Rate* (FNR, for short) is the fraction of normal objects rejected by the classifier, while the *True Negative Rate* (TNR, for short) is the fraction of abnormal objects rejected by the classifier. It holds that FNR=1-TPR and FPR=1-TNR.

Figures 3(a)-(c) compare the ROC curves of the CPDD (solid lines) and CNNDD (dash-dotted lines) methods and also the relative size $|S|/|D|$ of the corresponding consistent subsets S achieving the same value of FNR. The ROC curve is the plot of the FNR versus the TNR (or, correspondingly, TPR versus TNR), and the area under the ROC curve (AUC, for short) provides a summary to compare two classifiers. From these curves it is clear that the the CPDD consistent subset (dashed lines) is much smaller than the CNNDD subset (dotted lines) guaranteeing the same FNR. Moreover, the AUCs of the two methods are very similar.

Figures 3(d)-(f) report the TNR (solid lines) ad FNR (dashed lines) of the CPDD method, and the TNR (dash-dotted lines) and FNR (dotted lines) of the CNNDD method as a function of the relative subset size $|S|/|D|$. For the CPDD method the pair of parameters $\varrho = 1, \eta = 0.95$ (upper curve), $\varrho = 1, \eta = 1$ (middle curve), and $\varrho = 0.9, \eta = 1$ (lower curve), were considered, in order to study sensitivity to these parameters.

As far as the middle curves of the CPDD ($\varrho = 1, \eta = 1$) and the curves of the CNNDD is concerned, it can be noted that for the same value of FNR or TNR the subset of the CPDD is sensibly smaller than that of the CNNDD. As notable examples (highlighted by means of big points on the curves) compare (1)

[2] Also other data set were considered: the behavior of the method on these other data sets was analogous to what here described.



(a) CPDD AUC=0.997, SVM AUC=0.982 (b) CPDD AUC=0.970, SVM AUC=0.918 (c) CPDD AUC=0.986, SVM AUC=0.974

Figure 4. Comparison between the CPDD rule and the one-class SVM.

the CPDD subset of relative size 0.054, achieving FNR=0.024 and TNR=0.997, with the CNNDD subset of relative size 0.158, achieving FNR=0.027 and TNR=0.983, for the *Image segmentation* data set, (2) the CPDD subset of relative size 0.064, achieving FNR=0.031 and TNR=0.869, with the CNNDD subset of relative size 0.277, achieving FNR=0.045 and TNR=0.862, for the *Ionosphere* data set, and (3) the CPDD subset of relative size 0.042, achieving FNR=0.041 and TNR=0.952, with the CNNDD subset of relative size 0.106, achieving FNR=0.027 and TNR=0.943, for the *Satellite image* data set.

As far as the upper curves of the CPDD is concerned ($\varrho = 1, \eta = 0.95$), it can be noted that by decreasing the value of the parameter η, very high values of TNR are obtained in correspondence of very small subsets, but the associated FNR worsens with respect to the case $\eta = 1$. This can be explained since the smaller the parameter η, the greater the portion of the accepting region of the PDD rule which is left uncovered by the CPDD consistent subset.

As far as the lower curves of the CPDD is concerned ($\varrho = 0.9, \eta = 1$), on the contrary, it can be noted that by decreasing the value of the parameter ϱ, the FNR improves while the TNR gets worse. This can be explained since the smaller the parameter ϱ, the greater, and also the closer to each other, the number of prototypes composing the CPDD consistent subset.

Hence, by properly setting the parameters ϱ and η the user can tune the trade off between FNR and TNR and, simultaneously, between subset size and accuracy. The following table summarizes the AUCs of the CPDD for various combinations of the parameters ϱ and η.

Data set	ϱ	1.00	1.00	0.90	0.90
	η	1.00	0.95	1.00	0.95
Image segmentation		0.997	0.989	0.996	0.990
Ionosphere		0.970	0.956	0.972	0.967
Satellite image		0.986	0.981	0.989	0.987

Figure 4 compares the ROC curves of the CPDD method with that of the one-class SVM classifier [10, 5]. As for the one-class SVM, the radial basis function kernel was used, varying parameter γ between 10^{-4} and 10^2, and then the curve associated with the best AUC has been selected. As for the CPDD rule, the parameter used were $k = 4$, $\varrho = 1$, and $\eta \in \{0.95, 1.00\}$.

Interestingly, the CPDD performed better than the one-class SVM both in terms of accuracy (AUCs of the two methods are reported in figure) and in terms of size of the model. Indeed, as for the size of the model is concerned, for small FNRs, the size of the CPDD subset is practically identical to the number of support vectors, while, for greater values of FNRs, the former number is much smaller than the latter one. This can be explained by noticing that the CPDD subset does not contain the reference set outliers, which form the great majority of the reference set for large FNRs. Moreover, by setting the parameter η to 0.95, the size of the CPDD subset is further decreased, while the accuracy of the CPDD classifier remained good, as witnessed by the table above reported.

6 CONCLUSIONS AND FUTURE WORK

In this work the CPDD one-class classification algorithm has been presented and compared with the CNNDD and the one-class SVM classifiers, pointing out some advantages of the novel approach. A lot of additional questions are worth of being considered. Among them, studying the sensitivity of the method to the parameter k, comparison with the CNNDD rule under different metrics, comparison with other one-class classification methods, and using kernel functions to possibly improve size of the model and/or accuracy.

REFERENCES

[1] F. Angiulli, 'Condensed nearest neighbor data domain description', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(10), 1746–1758, (2007).

[2] F. Angiulli, 'Fast nearest neighbor condensation for large data sets classification', *IEEE Transactions on Knowledge and Data Engineering*, **19**(11), 1450–1464, (2007).

[3] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi, *Complexity and Approximation*, Springer-Verlag, Berlin, 1999.

[4] M. Bellare, S. Goldwasser, C. Lund, and A. Russeli, 'Efficient probabilistically checkable proofs and applications to approximations', in *STOC*, pp. 294–304, (1993).

[5] C.-C. Chang and C.-J. Lin, *LIBSVM: a library for support vector machines*, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[6] V. Chvátal, 'A greedy heuristic for the set-covering problem', *Mathematics of Operations Research*, **4**(3), 233–235, (1979).

[7] S. Floyd and M. Warmuth, 'Sample compression, learnability, and the vapnik-chervonenkis dimension', *Machine Learning*, **21**(3), 269–304, (1995).

[8] M.R. Garey and D.S. Johnson, *Computer and Intractability*, W. H. Freeman and Company, New York, 1979.

[9] P.E. Hart, 'The condensed nearest neighbor rule', *IEEE Trans. on Information Theory*, **14**, 515–516, (1968).

[10] B. Schlkopf, J.C. Platt, J. Shawe-Taylor, A.J. Smola, and R.C. Williamson, 'Estimating the support of a high-dimensional distribution', *Neural Computation*, **13**(7), 1443–1471, (2001).

[11] D.M.J. Tax, *One-class classification*, Ph.D. dissertation, Delft University of Technology, June 2001.