International Journal of Database Theory and Application
Vol.10, No.1 (2017), pp.105-118
http://dx.doi.org/10.14257/ijdta.2017.10.1.10

Kernel Credal Classification Rule – Application on Road Safety

Khawla El Bendadi*[1], Yissam Lakhdar[1] and El Hassan Sbai[2]

[1]Department of physics, Faculty of Sciences, University Moulay Ismail Meknes, Morroco
[1]k.elbendadi@gmail.com; [1]lakhdar_yissam@yahoo.fr
[2]Heigh school of Technologie, University Moulay Ismail
BP 3103 Route d'Agouray, 50006 Toulal, Meknes, Morocco
[2]sbaielhassan@yahoo.fr

Abstract

A credal partition based on belief functions has been proposed in the literature for data clustering. It allows the objects to belong; with different masses of belief; not only to the specific classes, but also to the sets of classes called meta-class which correspond to the disjunction of several specific classes. In this paper, a kernel version of the credal classification rule (CCR) is proposed to perform the classification in feature space of higher dimension. Each singleton class or meta-class is characterized by a center that can be obtained using many way. The kernels based approaches have become popular for several years to solve supervised or unsupervised learning problems. In this paper, our method is extended to the CCR. It is realized by replacing the inner product with an appropriate positive definite function, implicitly perform a nonlinear mapping of the input data into a high-dimensional feature space, and the corresponding algorithm is called kernel Credal Classification Rule(KCCR). We present in this work KCCR algorithm to powerful corresponding nonlinear form using Mercer kernels. The approach is applied for the classification of experimental data collected from a system called Vehicle-Infrastructure-Driver (VID), based on several representative trajectories observations made in a bend, to obtain adequate results with data experimentally realized based on the instructions given to drivers. The test on real experimental data shows the value of the exploratory analysis method of data. Another experiments using the generated and real data form benchmark database are presented to evaluate and compare the performance of the KCCR method with other classification approaches.

Keywords: *Classification, theory of belief functions, evidential C-means, kernel methods, vehicle trajectory.*

1. Introduction

The credal partition is a general extension of the fuzzy classification [1], which remains so far the most popular data clustering method, and it works with fuzzy partition under the probabilistic framework.

The credal partition based on the belief functions theory has been introduced in the data clustering analysis by Denoeux [2], for data clustering under the belief function framework. It allows the object not only to belong to single clusters, but also to belong to any subsets of the frame of discernment $\Omega = \{w_1, \ldots, w_2\}$ by allocating a mass of belief of each object to all elements of the power-set of Ω denoted 2^{Ω}.

In practice, the credal partition provides more sophisticated results than other partitioning techniques. An Evidential C-Means (ECM) [3] and a Noise-Clustering algorithm [4] clustering method inspired from FCM [1] have been proposed for the credal partition of object data.

ISSN: 2005-4270 IJDTA

In previous related works, Credal Classification Rule (CCR) [5] had developed to overcome the limitation of ECM by introducing an interpretation of the meta-class. In CCR approach, each specific class is characterized by the corresponding class center (*i.e.* prototype) computed using the classical way, for example, the class centers produced by FCM can be used in CCR for the unsupervised data classification. When the training data set is available in the supervised data classification problem, the average vector of the training data in each class can be considered as the class center in CCR. The center of meta-class may be calculated based on the involved specific classes' centers it should have the same, as much as possible similar, distance to each involved specific class center.

Since the original CCR uses the Euclidean distance to measure similarity between center of class and data points, it can be only effective in clustering for spherical clusters. Many algorithms may be derived from the CCR in order to cluster more general dataset. Most of those algorithms are realized by replacing the Euclidean distance with other similarity measures (metric) [6].

In this paper, kernel-based credal classification algorithm (KCCR) is proposed. KCCR adopts a kernel-induced metric to replace the original metric in CCR. By replacing the inner product with an appropriate 'kernel' function, one can implicitly perform a nonlinear mapping to a high dimensional feature space without increasing the number of parameters.

The kernel method was used in many learning approaches, such as Support Vector Machines (SVM) [7], kernel principal component analysis [8], kernel fisher discriminant analysis [9] and Kernel version of Evidential C-means [10].

Our algorithm is tested on real data, which are a set of trajectories cornering vehicles (LCPC/Nantes) [11]. Collect trajectories used is made by a slide composed of an instrumented vehicle and a given bend configuration.

The goal is to acquire a set of trajectories physically carried out by various drivers and various instructions. For this, 32 drivers crossed the turn ten times, and they have obtained 320 trajectories. These drivers were selected using the following criteria (age, gender, driving experience).

In order to perform these tests, two instructions were given to drivers: fast driving and reduced driving.

This bend can lead to study driver behavior.

After a brief presentation of credal classification rule in Section 2, we propose in Section 3 the kernel version of CCR. In Section 4, we present some classification results based on artificial and real data sets, and we compare the performances of the CCR with respect to another classification method. Conclusions are given in Section 5.

2. Brief Recall of Credal Classification Rule

In this section we present a brief recall of CCR approach. This method work with credal partition, and it led us to compute the masses of belief associated with specific class, meta-class, and outlier class. The credal partition is based on the belief function which have been introduced by shafer [12] in 1976 in his mathematical theory of evidence.

The construction of a meta-class is made by the union from at least two specific classes w_i and w_j. The basic belief of assignment (bba) is a function $m(.)$ from 2^Ω to $[0,1]$ satisfying $m(\emptyset) = 0$ and $\sum_{A \in 2^\Omega} m(A) = 1$, where Ω represent the frame of discernment. The credal partition is defined as $M = (m_1, \dots, m_n)$ where m_i represent the bba of the object x_i associated with the different elements of the power-set 2^Ω.

The implementation of the CCR algorithm requires first the determination of cluster centers (specific and meta-classes), then the construction of the bba by calculating the distance between the object to classify and each class center.

In the first step CCR determine the center of classes. For specific class, the center of each one can be obtained in several ways. For example, we can use the average of training data or the centers produced by an unsupervised clustering method as FCM or ECM.

Let us consider a data set $X = \{x_1, \dots, x_n\}$ of point in a feature space to be classified in the frame of discernment $\Omega = \{w_0, w_1, \dots, w_h\}$, where w_0 represent the outlier class. The contribution of credal classification method lies in the introduction of meta-class which is used to model imprecision of the objet. CCR propose a new method to determine the center of the meta-class. For this, the center of meta-class must keep the same distances to all the centers of the involved specific classes.

In CCR, c_U is a center of meta-class named U which is chosen at the same distance to all the centers of the specific classes included in U. So the following condition must be satisfied

$$d(c_U, c_i) = d(c_U, c_j), \forall w_i, w_i \in U, i \neq j \tag{1}$$

The computation of the meta-class centers depends on the dimension of the vector c_U. Equation (1) referred to the calculation of the distance between the center and the objet to classify. If the dimension of vector c_U is equal to 1, then it can produce only one solution. If the dimension of c_U is bigger than $|U| - 1$, we find many possible solutions for c_U. Given that the meta-class center should be close as possible to all the center of the specific class included in U, the closer will be selected using $\arg[min_{c_u} \sum_{w_j \in U}(d(c_U, c_j))]$. If the dimension of c_U is smaller than $|U| - 1$, it will become an optimization problem to seek the solution of c_U, it can be solved using any classical nonlinear optimization method as the least squares estimate method. In the second step, we should construct our bba's.

Let us consider one object x_s to be classified. The initial mass of x_s on a specific class should be a monotone decreasing function of the distance between the object and the corresponding class center, and it's denoted by:

$$\tilde{m}(w_i) = f_1(d(x_s, c_i)), \forall i = 1, \dots, h \tag{2}$$

In CCR, the mass of belief for the object on a meta-class should be calculated by the center of the specific class involved in the meta-class, also we use the distance to the center of this meta-class using a function denoted by $f_2(.)$ shown in equation (2). In the determination of the mass of belief for an object x_s, the smaller distance between x_s and the centers of singleton classes included in U means that this object belong to this meta-class.

$$\tilde{m}(U) = f_2(\Delta(x_s, c_i)) \tag{3}$$

where

$$\Delta(x, c_U) \triangleq \frac{1}{|U| + \gamma}\left[\sum_{w_i \in U} d^{\beta}(x_s, c_i) + \gamma d^{\beta}(x_s, c_U)\right] \tag{4}$$

The tuning parameter β can be fixed to a small value (1 or 2). The parameter γ is a tuning weighting factor of the distance between the object and the center of the meta-class. The exponential decreasing functions for $f_1(.)$ and $f_2(.)$ are used in [13]. The mass of belief on specific class and meta-class are given by:

$$\tilde{m}(w_i) = e^{-\lambda d^{\beta}(x_s, c_i)} \tag{5}$$

$$\tilde{m}(U) = e^{-\lambda \Delta(x, c_U)} \tag{6}$$

The outlier class contains every objet which is so far from all the classes using an outlier threshold δ as controller parameter. The mass of belief of outlier class is defined by:

$$\tilde{m}(Uw_0) = e^{-\lambda\delta} \tag{7}$$

where

$$\delta = \eta \times \arg[\max\left(d^{\beta}(c_i, c_j)\right)] \quad (8)$$

3. Kernel Credal Classification Rule

3.1. Mercer Kernel

Depending on the type of the tested data, different distances can be used in the CCR method. In [14] Euclidean distance was adopted to classify spherical data. In another work [5] mahalanobis distance was used to deal with the anisotropic data sets.

In order to overcame the limitations of both de Euclidean and mahalanobis distances, a kernel based credal classification rule (KCCR) is proposed to produce clusters in a high, possibly infinite, dimensional space based on Mercer Kernel technique.

To solve the problem of non-linearly-separable clusters, many methods have been modified incorporating kernels such as SVM [15] which performs better than other classification algorithms in many applications. The success of SVM has extended the use of kernels to other learning algorithms, kernel k-means [16], kernel fuzzy c-means [17]. These methods use positive-definite kernel, also referred to as Mercer kernel, to implicitly map data from original space called input space into a high dimensional space called feature space.

The kernel k is defined by:

$$k(x,y) = \Phi(x)^t \Phi(y) \quad (9)$$

where Φ: $X \rightarrow F$ is a mapping from the input space X to a high dimensional feature space F.

The advantage of kernel consists in the possibility of computing distance measure between observations in the feature space F without explicitly knowing Φ. The kernel induced distance measure can be computed as follows:

$$\begin{aligned} \left\|\Phi(x_i) - \Phi(x_j)\right\|^2 &= ((\Phi(x_i) - \Phi(x_j))((\Phi(x_i) - \Phi(x_j)) \\ &= \Phi(x_i)\Phi(x_i) - 2\Phi(x_i)\Phi(x_j) + \Phi(x_j)\Phi(x_j) \\ &= K_{ii} - 2K_{ij} + K_{jj} \end{aligned} \quad (10)$$

We find many type of mercer kernels such as Linear, Polynomial, Gaussian, Exponential and Sigmoid kernels. A Gaussian kernel will be used in our experiments:

$$k(x,y) = exp\left(\frac{-\|x-y\|^2}{\sigma^2}\right) \quad (11)$$

here σ represent the bandwidth of the kernel.

If we confine ourselves to the Gaussian kernel which is used almost in the literature, then $k(x,x) = 1$, and

$$\left\|\Phi(x_i) - \Phi(x_j)\right\|^2 = 1 - 2k(x_i, x_j) \quad (12)$$

3.2. Algorithm

In this part, the KCCR algorithm is introduced in detail to overcome the limitations of the use of the Euclidean distance. The algorithm KCCR as CCR provides a method to calculate the masses of beliefs for singleton class, meta-class and noise class, also it mainly consists of two steps : firstly we determine the center of the specific and meta classes, secondly we should construct the bba's based on the distance between the object and classes centers.

Let us consider a nonlinear map as $\Phi : x \rightarrow \Phi(x) \in F$, where $x \in X$. X denotes the data space, and F the transformed feature space with higher dimension. The data set X will be classified in the frame of discernment $\Omega = \{w_0, w_1, \dots, w_h\}$. The center of each specific class can be obtained using several methods, so we can use the centers produced by FCM or ECM, or the average of training data.

CCR approach proposed a new approach to calculate the centers of meta-class. The object is assigned to a meta-class when it is in approximately the same distance from the centers classes.

We consider a meta-class denoted by U, and c_U is the center of this meta-class which is chosen at the same distance to all the centers of the specific classes included in U.

For this the following conditions must be satisfied

$$\left\| c_U^{\phi} - c_i^{\phi} \right\| = \left\| c_U^{\phi} - c_j^{\phi} \right\|, \forall w_i, w_j \in U, i \neq j \tag{13}$$

and it can be proven that $\|.\|$defined in equation (13) is a metric in the original space in case that $k(x, y)$ takes as the Gaussian kernel function, which measures the similarity between the center of specific class c_i^{ϕ} and center of meta-class c_U^{ϕ}.

When the dimension of vector c_U^{ϕ} is equal to one, the equation (13) represent one constraint and it can produce only one solution of c_U^{ϕ}. If c_U^{ϕ} has a dimension bigger than one, many solutions for c_U^{ϕ} are given by $\arg[min_{c_U^{\phi}} \sum_{w_j \in U}(d^{\phi}(c_U, c_j))]$. The use of a nonlinear method for optimization should be very useful if the dimension of c_U^{ϕ} is smaller than one.

After determination of the centers of specific class and meta-class, now we construct the bba's of objects belonging to a specific class, meta-class or outlier class, based on the similarity between the object and the other classes. For credal classification we have to compute three type of masses of belief according to the involvement of the object.

If the object is committed to a specific class $w_i$which indicates that x is closer to w_i. So the initial mass of belief of x should be a monotone decreasing function denoted by $f_1(.)$, its expression is as follow:

$$\widetilde{m}(w_i) = f_1\left(d^{\phi}(x, c_i)\right), \quad \forall\, i = 1, \dots, h \tag{14}$$

with $d^{\phi}(x, c_i)= \|\Phi(x) - \Phi(c_i)\|^2 = 1 - 2K(x, c_i)$

If the object belong to a meta-class, it means that this object must be very close to these specific classes. The determination of the mass of belief of object on meta-class U should be done by both the distances to the centers of the specific classes involved in the meta-class and the distance to the center of this meta-class. Therefore, the mass of belief for x on the meta-class Ushould be a function denoted by$f_2(.)$ of both the distance to each center of the specific class involved in U and distance to center of meta-class c_U^{ϕ}.

$$\widetilde{m}(U) = f_2(\Delta^{\phi}(x, c_U)) \tag{15}$$

where

$$\Delta^{\phi}(x, c_U) \triangleq \frac{1}{|U| + \gamma} \cdot [\sum_{w_i \in U} d^{\beta n^{\phi}}(xn) + \gamma d^{\beta^{\phi}}(x, c_U)] \tag{16}$$

β is a tuning parameter that can be fixed to a small positive value (1or 2). The parameter γ is an adjustment-weighting factor of the distance between the object and the center of meta-class. $\Delta^{\phi}(x, c_U)$ is the average value of the distances between x and all the centers of the involved specific classes, and the distance between x and c_U^{ϕ}. The function $f_2(.)$should be decreasing with the increasing of $\Delta^{\phi}(x, c_U)$.

Both functions $f_1(.)$ and $f_2(.)$ are monotone decreasing function. In this work as CCR approach, we use the exponential decreasing function which is frequently used in classification problem. So the mass of belief for specific and meta-class are given by:

$$\widetilde{m}(w_i) = e^{-\lambda d^{\beta^{\phi}}(x, c_i)} \tag{17}$$

$$\tilde{m}(U) = e^{-\lambda \Delta^{\Phi}(x, c_U)} \quad (18)$$

It is worth noting that λ is a parameter, which can be determined by the average distance between each pair of centers of the specific classes.

$$\lambda = \frac{1}{\frac{1}{2}|U|(|U|-1)} \cdot \sum_{i=1}^{n} \sum_{j>i} d^{\beta^{\Phi}}(c_i, c_j) \quad (19)$$

The object belong on the outlier class, if x is so far from all the classes. The controller parameter δ is used as outlier threshold, which can be determined as follow:

$$\delta = \eta \times \arg[max_{i,j}(d^{\beta^{\Phi}}(c_i, c_j))] \quad (20)$$

η must be a tuning parameter with positive value. The mass of belief of the outlier class for x can be written as

$$\tilde{m}(w_0) = e^{-\lambda\delta} \quad (21)$$

Since the majority of work in the belief function use the normalized bba's, the cited masses that are not normalized should be normalized using the following expression for the credal classification of the object x, $\forall A \subseteq \Omega$

$$m(A) = \frac{\tilde{m}(A)}{\sum_{B \subset \Omega} \tilde{m}(B)} \quad (22)$$

In the next section, a synthetic and real data will be presented to show the effectiveness of KCCR method.

4. Experimental Results

In this section, we show several examples to illustrate the approach presented in the previous section in order to compare the obtained results with another method to show the effectiveness of the proposed method. The first example involve a synthetic data set. The second and the third experiment is based on real data, one from UCI Machine Learning Repository [18], the other is an application on road safety which data are collected from the trajectories take in a bend [11].

The credal classification has the advantage of considering the imprecise objects in the box in the imprecise classification and not as misclassified objects (misclassification). In our experiments, the concept of imprecision of the classification and misclassification is introduced to evaluate the performance of our method KCCR.

For this, classification error is calculated by: $R_e = (N_e / T) * 100\%$, where N_e is the number of misclassified objects and T is the total number of objects under test. The degree of imprecision R_{ij} is calculated by $R_{ij} = (N_{ij} / T) * 100\%$, where $N_{ij}$is the number of objects that assigned to the meta-clusters with a cardinality value j.

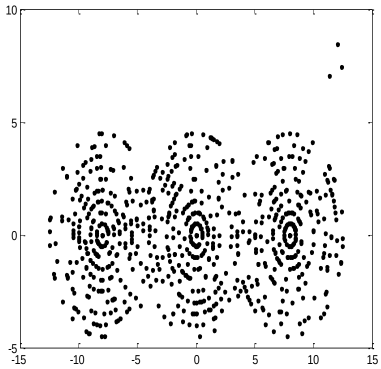
The choice of the value of the parameter γ is critical, since it acts on the number of objects in the meta-clusters. Because the right choice for this parameter influence positively on the result of the clustering, we obtain a good compromise between the degree of imprecision and classification error that induces a very good classification rate. In our experiments, several values of the parameter γ are tested.

4.1. Experiment 1 (3-Spherical Classes)

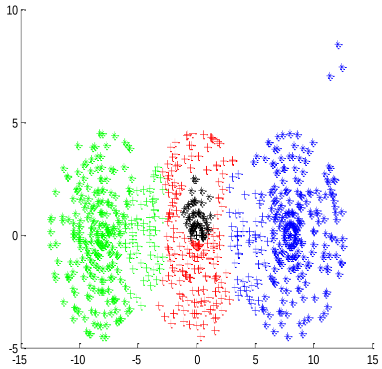
In this experiment, we consider a set of data generated as a spherical 3 classes as shown in Figure 1. This distribution of data consists of 903 data points with three noise point. The centers of the three circles are respectively $c_1(0,-8)$, $c_2(0,0)$, $c_3(0,8)$ and the radii are $r_1 = 4.5$, $r_2 = 4.5$, $r_3 = 4.5$.

KCCR algorithm is applied for the classification of this data set and it is compared with CCR and ECM. In this example, we tested both CCR and KCCR approaches with the same parameters, and we choose $\alpha = 2$ and $\eta = 0.5$.

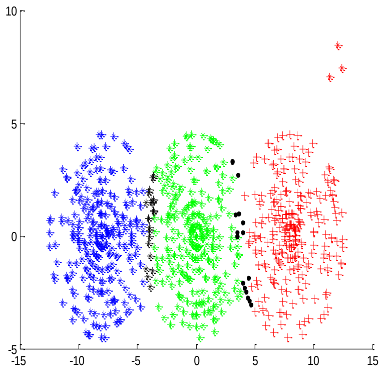
After several experiment with KCCR method, the value of the bandwidth of the kernel is fixed in $\sigma^2 = 20$. The classification results with different methods are shown in Figure1. In this work and in our experiments, we use the centers acquired by FCM. Table 1 shows the different values of error rate and imprecision rate where "NA" means Not Applicable. The outliers are well detected by ECM, CCR and KCCR.



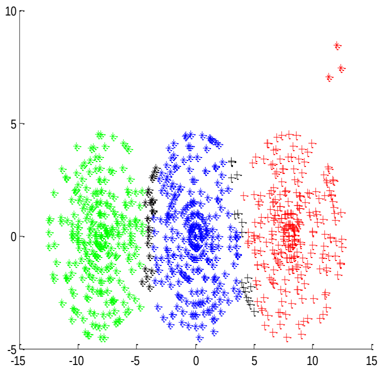
a) Original data



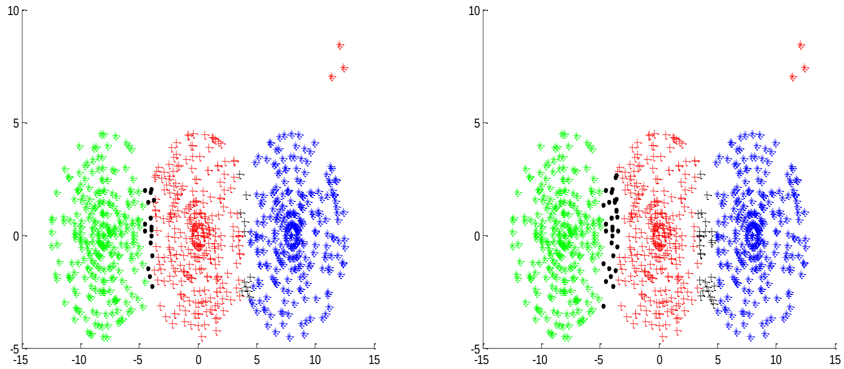
b) Classification results by ECM



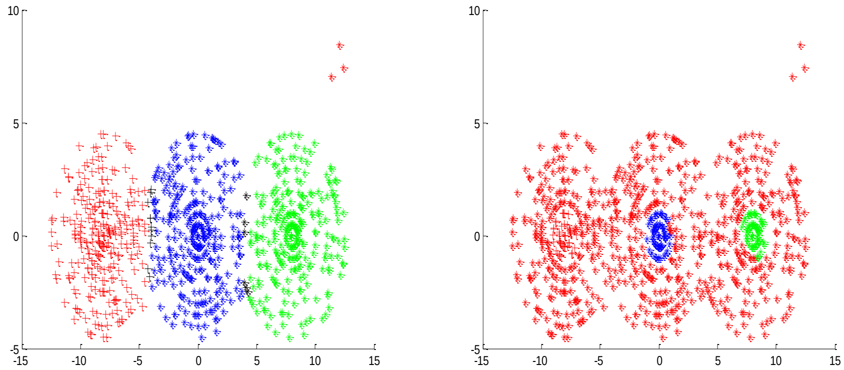
c) Classification results by CCR with $\gamma = 0.5$



d) Classification results by CCR with $\gamma = 1$



e) Classification results by KCCR with γ =0.5 f) Classification results by KCCR with γ =1



g) Classification results by KCCR with $\eta = 0.05$, $\beta = 2$ h) Classification results by CCR with $\eta = 0.05$, $\beta = 2$

Figure 1. Classification Results by different Methods for a 3-Classes Problem

In the first Figure (Figure 1.a), we find the distribution of the original data, which shows that the class in the middle is partially overlapped with the other two classes. The points belonging to the overlapped areas represent considerable ambiguity and they are hard to be correctly classified. As shown in the results, ECM, CCR and KCCR produce 3 specifics classes.

In the results obtained by ECM we can see that the class represented by green star symbol (w_1) and the other class with blue star symbol (w_3) are not close and they are totally separated, but there are still many points originally from w_2 that are wrongly committed to the meta-cluster $w_1 \cup w_2 \cup w_3$ labeled by black in (Figure 1.b). These results show the wrong behavior of ECM.

KCCR and CCR approaches produce more reasonable credal classification results in comparison with ECM method. By giving different value of γ, the overlapped area leads to more data points in meta-class as shown in (Figure 1.c-d-e-f). So the γ value should be tuned to obtain a good compromise between the imprecision degree and the misclassification rate.

In KCCR, when we take $\gamma = 1$, we have the same error rate (0%), and the imprecision rate is equal to 6.44 %, which is less than that obtained by CCR (6.56 %) with the same value of γ where $\gamma = 1$.

As shown in (Figure 1.g), by setting the parameter $\eta = 0.05$ and $\beta = 2$, our method find a good compromise between the imprecision degree and the misclassification rate contrary to CCR method, which tends to be not applicable because it produce much noisy points as mentioned in (Figure 1.h).

The imprecision rate produced by our approach is equal to 1.89 %, and the error rate have as value 1.67 %, but with the CCR method we cannot applied this approach because this method use an Euclidean distance contrary to KCCR which produce clusters in a high dimensional space based on kernel.

In this experiments, we show that the parameter β influence the results of classification with KCCR method, for example, by setting $\beta = 2$ we can get a less error rate than CCR method as shown in (Figure 1.g).

By this we can conclude that KCCR provide a good results than CCR, and we show that by introducing the kernel method in credal classification, KCCR approach can be performed to be more effectiveness and powerful .

Table 1. Error Rate and Imprecision Rate for Spherical Classes with Different Methods

	R_e	R_{i2}
ECM $\alpha = 1$	0 %	22.11 %
CCR β**=1,** γ**=0.5**	0.89 %	4.11 %
CCR β **=1;**γ **=1**	0 %	6.56 %
CCR β **=2;**η **=0.05;**γ **=0.5**	NA	NA
KCCR β **=1;**γ **=0.5**	1.33 %	3.11 %
KCCR β **=1;**γ **=1**	0 %	6.44 %
KCCR β **=2;**η **=0.05;**γ **=0.5**	1.67 %	1.89 %

4.2. Experiment 2 (Real Data Set)

In this experiment, we use two types of real data namely, Iris and Seeds data from UCI repository [18]. To test the performance of KCCR, we compare the results with those obtained by ECM and CCR, which uses the Euclidean metric.

Iris data sets consists of 3 different types of irises' (Setosa, Versicolour, and Virginica) which represents 3 classes, stored in a 150 instances and 4 attributes. All the detailed characteristics of the data sets can be found on UCI repository archive at http://archive.ics.uci.edu/ml/.

Our approach is applied in these two data sets for the unsupervised data classification. The obtained results are compared with ECM, CCR and KCCR. In KCCR, the class centers obtained by FCM are adopted. Different values of γ in CCR and KCCR are selected to show their influence on the results.

Table 2. Classification Results of Iris Data with Different Methods

	R_e	R_{i2}
ECM $\alpha = 1$	8	4.67
CCR γ **=1**	7.33	6
CCR γ **=2**	3.33	12
KCCR γ **=1**	7.33	6
KCCR γ **=2**	3.33	12

For the iris data set, the number of errors generat8ed by ECM causes a high imprecision degree of the results, which is not an efficient solution as shown in Table 2.

By using ECM the value of error rate is equal to 8%. Contrary to KCCR, which the value reach 3.33% when γ =2 .So we can see that the obtained results are encouraged.

KCCR provides a good compromise between the error and the imprecision rate as CCR, they produces smaller error rate than ECM, and the imprecision degrees of CCR and KCCR are generally smaller than with ECM.

For the seeds data set, when ECM and CCR produce similar error rates and imprecision rate with the given parameter β and γ, the imprecision degree of KCCR remains better than with ECM and CCR as shown in Table 3.

The result of seeds data mentioned in Table 3 shows that KCCR provides better performances than ECM and CCR.

Table 3. Classification Results of Seeds Data with Different Methods

	R_e	R_{i2}
ECM $\alpha = 1$	5.71	13.33
CCR γ**=0.5**	8.10	8.57
CCR γ **=1**	5.71	13.33
CCR γ **=1;** β **=2**	17.14	13.33
KCCR γ**=0.5**	8.57	3.33
KCCR γ **=1**	6.19	9.52
KCCR γ **=1;** β **=2**	7.62	5.71

We can conclude that KCCR gives a good compromise between error rate and imprecision rate (R_e =7.62 and R_{i2}=5.71) by setting the parameter $\gamma = 1$ and $\beta = 2$. In this case, CCR produces a very higher error rate (17.14%) which can explain the wrong behavior of CCR using seeds data.

4.3. Experiment3 (Vehicles Trajectories)

In this Experiment, we applied the kernel credal Classification Rule algorithm (KCCR) described in Section III to perform the classification of experimental vehicle trajectories, which are the results of the interaction of the system called Vehicle-Infrastructure-Driver (V-I-D). The trajectories used in this study were collected from an instrumented vehicle and a given bend configuration [11, 19].

The observation of the vehicle (trajectory) consists of six variables which are the positions (x_1, x_2), the speeds (v_1, v_2) and accelerations (γ_1, γ_2), longitudinal and lateral, respectively, in the Galilean reference.

The set of trajectories comprises 232 trajectories. We assume that the observations are representative trajectories practiced in this bend.

To apply KCCR algorithm we used a Gaussian kernel, but the selection of the parameter that defines the kernel (bandwidth for the Gaussian kernel) remains one of the problems of kernel methods.

The application of KCCR algorithm to the trajectories of vehicles gives four classes as results. These classes represent in the reality the behavior of four types of drivers. The first class represents the family of the slowest trajectories of safe driving. The second class is the family of fast trajectories with safe driving. The third class is the family of the slowest trajectories of sport driving. The fourth class represents the family of the fastest trajectories with sporting driving.

The classification method based on C-means adopted by Koita [11, 19] allows us to get four classes of stable trajectories. They physically correspond to four behavioral styles of each driver. Table 4 shows the distribution of the number of trajectories observed.

The use of C-means affects all trajectories to a given class. Contrary to our method, that produces trajectories belong not only to singleton clusters, but also to meta-clusters.

The use of ECM method to classify the paths from the bend [20] provide several meta-clusters and we obtain only two classes. After a kernel version of the ECM method have been tested [21], by this we could find the four classes desired but we still keeping too much trajectories in meta-clusters.

Table 4. Clustering Results of Trajectory (Number of Trajectories)

	C_1	C_2	C_3	C_4
C-means	30	49	76	77
CCR	34	42	67	70
KCCR	30	45	75	76

The use of our method produces the best results in comparison with C-means and CCR as mentioned in Table 4. By setting $\gamma = 0.5$, we find only 6 trajectories not correctly classified that are in meta-clusters, which means that these trajectories have a behavior of two types of drivers contrary to CCR that provides 19 imprecise trajectories. The data belonging to meta-clusters represents a new behavior inspired from at least two behaviors that are established by the experts in the field. These data should be examined to determine to which class belongs.

Table 5. Error Rate and Imprecision Rate for Trajectories

	R_e	R_{i2}
CCR	1.29%	8.19%
KCCR	0.43%	2.59%

Based on the C-means labels, we calculate the error rate and imprecision rate for CCR and KCCR approaches as mentioned in Table 5, which show that CCR produce a very higher imprecision rate $R_{i2} = 8.19$ contrary to KCCR that result a less imprecision rate, which is equal to $R_{i2} = 2.59$.

5. Conclusion

A new clustering method, called KCCR (Kernel Credal Classification Rule) based on the belief functions, has been proposed and evaluated in this paper. KCCR is an extension of CCR and an alternative of ECM. After analyzing the limitations of CCR with different metrics, we have proposed another version based on kernel function to solve the problems by mapping data with nonlinear relationship to appropriate features spaces.

Several experiments using the artificial and real data have been presented to evaluate the performance of KCCM with respect to other method. Our results show that KCCM provide a good credal classification result.

KCCR as CCR produce three kinds of clusters: Singleton clusters, meta-clusters and outliers clusters. By introducing meta-clusters, KCCR reduce the misclassification errors. The KCCR method as the CCR approach is able to detect the outliers in the data sets.

Acknowledgments

The authors wish to thank the professor Abdourahmane KOITA from IFSTTAR laboratory in Paris-Est University to provide us the cornering trajectory data.

References

[1] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, **(1981).**

[2] T. Denœux, M.-H. Masson, EVCLUS: EVidential CLUStering of proximity data, IEEE Trans. on Systems, Man and Cybernetics Part B, Vol. 34(1), **(2004)**, pp. 95–109.

[3] M.-H. Masson, T. Denœux, ECM: An evidential version of the fuzzy cmeans algorithm, Pattern Recognition, Vol. 41(4), **(2008)**, pp. 1384–1397.

[4] R.N. Dave, Clustering relational data containing noise and outliers, Pattern Recognition. Lett. 12 **(1991)**, pp. 657–664.

[5] Zhun ga Liu, Quan Pan, Jean Dezert, Grégoire Mercier Credal classification rule for uncertain data based on belief functions. Pattern Recognition 47(7): **(2014)**, pp. 2532-2541.

[6] K.L.Wu, M.S.Yang, Alternative c-means clustering algorithms, Pattern Recognition **(2002)** 35: 2267-2278

[7] N. Cristianini, J.S. Taylor. An Introduction to SVMs and other kernel-based learning methods. Cambridge Univ. Press, **(2000)**.

[8] B. Scholkopf, A. J. Smola, and K.-R.Muller. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. Neural Computation, 10(5), **(1998)**, pp.1299-1319.

[9] S. Mika, G. Ratsch, J. Weston, B. Scholkopfand K.-R. Muller.Fisher Discriminant Analysis with Kernels. Neural Networks for Signal Processing IX. **(1999)**, pp. 41-48.

[10] M.-H. Masson and T. Denoeux. KECM : Une version noyau de l'algorithme évidentiel des c-moyennes. In Rencontres Francophones sur la Logique Floue et ses Applications (LFA '08), **(2008)**, pp 216-223, Lens, France.

[11] Koita A., Daucher D., Fogli M., New Probabilistic Approach to Estimate Vehicle Failure Trajectories in Curve Driving. Probabilistic Engineering Mechanics.Volume 34 (**2013**), pp 73–82.

[12] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, **1976**.

[13] T. Denœux, A k-nearest neighbor classification rule based on Dempster Shafer Theory, IEEE Trans. on Systems, Man and Cybernetics, Vol. 25(05), pp. 804–813,**1995**.

[14] Zhun-ga Liua, Quan Pana, Jean Dezertb, Gregoire Mercierc Credal Classification of Uncertain Data Using Belief Functions SMC '13 Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics **(2013)**, pp.1001-1006

[15] Corinna Cortes, Vladimir Vapnik Support-Vector Networks. Machine Learning Vol 20(03), pp.2730–297,1995

[16] M Girolami Mercer kernel-based clustering in feature space .Neural Networks, IEEE Transactions on 13 (3), 780-784.

[17] Wu Z, Xie,W.X Yu J.P. "Fuzzy C-means Clustering Algorithm Based on Kernel Method" In: Proceedings of Fifth International Conference on Computational Intelligence and Multimedia Applications, (**2003)**, pp 49-56,

[18] A. Frank, A. Asuncion, UCI Machine Learning Repository, University of California, School of Information and Computer Science, Irvine, CA, USA,**(2010)** ⟨http://archive.ics.uci.edu/ml⟩.

[19] Daucher D., Koïta A. Modélisation de trajectoires de véhicules en virage par analyse de données expérimentales - application à la sécurité routière, dans XIXe Congrès français de Mécanique (CFM'09), **(2009)**, Marseille, France, 24-28 août

[20] K.El Bendadi, Y LAKHDAR, El-H Sbai, A KOITA Classification des trajectoires de vehicules en virage par C-moyenne évidentielle. Mediterranean Telecommunication Journal 4: 2. **(2014)**, pp. 62-66 October.

[21] K.El Bendadi, Y LAKHDAR, El-H Sbai, A KOITA Vehicle Trajectories Classification using Kernel Evidential C-means algorithm for Failure Trajectory Prediction In: CESA 6th Multi-conferences on Computational Engineering in Systems Applications IEEE , **(2015)** Marrakech (Morocco).

Authors

Khawla El Bendadi, She got a master degree on "computer sciences, computer graphics and imagery" from faculty of sciences, Dhar El Mehraz, university Sidi Mohamed Ben Abdallah, Fez, Morocco, in 2012. Since september 2012, she has been working on his Phd thesis at the university of My Ismaïl, Meknes, Morocco, with the research team "Control Systems and Information Processing (CSIP)" at higher school of technology of Meknes. His scientific interests include data analysis, Credal Classification, theory of evidence, cluster analysis and machine learning.

Yissam Lakhdar, received his Ph.D thesis at the university of My Ismaïl, Meknes, Morocco, now, he is attached to the research team "Control Systems and Information Processing (CSIP)" at higher school of technology of Meknes as researcher. His scientific interests include data analysis, artificial intelligence, cluster analysis, image processing and machine learning.

SBAI El Hassan, received the Doctorate degree (Doctorat de 3ème cycle) from the university of Sidi Mohammed Ben Abdellah, Fez, Morocco in 1992, and the Doctorate of sciences (Doctorat d'Etat) from the university of My Ismaïl, Meknes, Morocco in 2001. Since 1993, he is a Professor in automatics and informatics at higher school of technology of Meknes. His main research interest are cluster analysis, pattern recognition, machine learning and computer vision.