RAIRO-Oper. Res. 51 (2017) 639–667
DOI: 10.1051/ro/2016045

RAIRO Operations Research
www.rairo-ro.org

PREDICTIVE DECISION MAKING UNDER RISK AND UNCERTAINTY: A SUPPORT VECTOR MACHINES MODEL

Wasfi A. Al-Rawabdeh[1] and Doraid Dalalah[2]

Abstract. In this paper, a decision making model using support vector machine (SVM) approach is presented. Here, human attitude towards risk and uncertainty is identified *via* optimizing SVM certainty classification model. In particular, individuals are given different pairs of gambles in order to reveal their preference. Unlike traditional methods used to estimate the utility function through direct inquiry of the certainty equivalents, pair-wise comparisons are used here in the training process to predict human preferences and to compute the utility parameters. The presented study is characterized by first, the use of SVM in the field of decision making to classify individuals' choices, second, it uses such model to search for the optimal utility parameters, third, the model can be used to guide the decision makers towards better decisions. In contrast to existing utility models, the SVM utility approach is characterized by its tolerance to misclassification in the training and testing data sets which makes it cope with the existing violations such as the common consequence, common ratio and violation of betweenness in the utility theory. To demonstrate the merits of the model, different data sets were used from well known literature studies and new conducted surveys that elicit individual preferences. The data is split into training and testing sets. The results demonstrated a notable consistency in the computed utility parameters and remarkable predictions without the need to strict certainty equivalent estimation. The model can be beneficial in predictive decision making under risk and uncertainty.

Mathematics Subject Classification. 90Cxx, 62Cxx, 80Mxx, 60Gxx.

Received December 23, 2015. Accepted June 20, 2016.

1. Introduction

In many aspects of life, we must often make decisions which may be vulnerable to risk and uncertainty. When the outcome of the selected act is not certain, individuals tend to choose rationally, an attitude which can be explained by the expected utility theory to a certain extent.

The expected utility theory is normative theory – which tackles how people should make their decisions. It can be a descriptive or sometimes a predictive theory which tends to correctly predict people's choices without worrying about the psychological mechanisms.

It is reported that the expected utility theory makes faulty predictions about people's decisions in many real-life choice situations [9,27]; hence, a question that needs to be answered is whether people should make decisions

Keywords. Decision making, expected utility theory, support vector machines, optimization.

[1] Faculty of Economics and Administrative Sciences, The Hashemite University, Al-Zarqa, Jordan. rawabdeh@hu.edu.jo

[2] Industrial Engineering Department, Faculty of Engineering, Jordan University of Science and Technology, P.O. Box 3030, 22110 Irbid, Jordan. doraid@just.edu.jo

Article published by EDP Sciences

based on expected utility considerations. Although expected utility theory could find a wide acceptance in the decision making area, yet there exist well theorized counter examples which the expected utility fails to explain. Such examples include:

Transitivity and completeness

That is, preferences must be transitive, for example if the lottery X is preferred to Y and Y is preferred to Z, then Y must be preferred to Z. However, in reality there are some situations where rationality entails and sometimes requires the failure of transitivity and completeness. The reader may be referred to the well known Quinn's puzzle of the self-torturer for further details. In fact, this can be stated more formally in the so called Violation of Betweenness, where if for instance X is preferred to Z, and Y is a linear combination of X and Z, then based on the expected utility theory, individuals should respectively prefer the pattern X, Y, Z or Z, Y, X, however, violations of such choice patterns have been numerously reported in several studies [2,5]. The violation of betweenness emerges when some of the prospects used in the experiment are located at the edges of the probability triangle (Machina, 1982).

Independence

Which is a property that refers to the parallel utility contours in Machina triangle, while in reality, such contours are not necessarily parallel as found by Allais paradox – a prominent violation to the expected utility theory.

The first serious challenge to the expected utility theory was presented by Maurice Allais [1]. It has become known as the Allais Paradox. Machina [20] showed that the Allais paradox is a special case of a more general problem, known as the common consequence effect.

Common consequence effect in decision theory is a violation of linearity in probabilities and outcomes. It is simply described by a prospect of three outcomes, where $x_1 < x_2 < x_3$ and the probabilities $1 > p > q > 0$. Denoting a prospect by its outcomes followed by the probabilities, the majority will prefer a sure outcome x_1 denoted by $l_1 = (x_2,\, 1)$ to $l_2 = (x_1,\, p - q;\, x_2,\, 1 - p;\, x_3, q)$ conforming to the EUT, while at the same time, individuals will prefer $l_4 = (x_1,\, 1 - q;\, x_3,\, q)$ to $l_3 = (x_1,\, 1 - p;\, x_2,\, p)$ which presents a clear violation to the EUT [3]. In fact the EUT implies that the choice must be either l_1 and l_3, or l_2 and l_4. In this instance, a risk aversion behavior appears in the first pair and risk seeking in the second.

On the other hand, the common ratio effect is an empirical observation which shows that there exist outcomes $x_1 < x_2 < x_3$ and a probability $0 < p < 1$ such that $l_1 = (x_2, 1)$ is preferred to $l_2 = (x_1, 1-p; x_3, p)$ and at the same time $l_4 = (x_1,\, 1 - pq;\, x_3,\, pq)$ is preferred to $l_3 = (x_1,\, 1 - q;\, x_2,\, q)$ when the probability q is close to zero. The choice of the riskier lotteries l_2 andl_4 becomes more probable when the probabilities of the middle and the highest outcome are scaled down in the same proportion. Common consequence and common ratio violate the EUT in the sense that acts with higher expected utilities are not always preferred.

Events of "0" probability

A concept that argues non-conformity to the expected utility when multiple events have a probability of zero where indifference is expected when a tie in the expected utilities exists.

Unbounded utility

Here, some decision problems may end up with infinite expected utility of some acts. For example, if we consider St. Petersburg game, we realize that the game expected utility is positive infinite, while we find individuals pay no more than $3 for such a gamble.

In fact, the values of these infinitary games are not well described by their expected utilities. Many research proposal tackled the above violations such as Thalos and Richardson [26], Fine [13], Colyvan [6,7] and Easwaran [10] who proposed theories that still conform with the expected utility theory and provide rational evaluations of the above listed violations such as those found in St. Petersburg, Pasadena and Altadena games.

1.1. The fourfold pattern of risk

The fourfold pattern was the first descriptive challenge in the EUT. It states that human choice patterns are unpredictable due to human variations, that is, it is possible to find individuals who are risk seeking over unlikely gains, risk seeking over likely losses, and at the same time, risk averse over likely gains and risk averse over likely losses.

An example of the fourfold pattern of risk attitude is a simultaneous purchase of insurance policies and lottery tickets. The expected utility theory cannot explain why an individual exhibits such an attitude [16].

Many studies have reported the fourfold pattern of risk. Kahneman and Tversky [16] presented surveys for choices between lotteries of large hypothetical gains and losses. They presented 25 graduate students to choose between simple lotteries with small hypothetical gains and losses. They found a strong support for the fourfold pattern. Holt and Laury [15] compared between decisions of lotteries of hypothetical and real gains. They found that people are risk-seeking with hypothetical rather than real lotteries. Laury and Holt [17] introduced also hypothetical and real lotteries comparison over gains and losses. They found that in the hypothetical lotteries, people are risk-averse over gains and risk-seeking over losses. In the real lotteries they are risk-averse for gains but risk-neutral for losses.

In relation to the discussion of violations presented above, our objective in this paper is to present a model that admits the above violations, can predict human preferences and can guide people to the right choices without strict need to utility. In many literature experiments, the surveyed subjects are provided with gambles and are requested to present the certainty equivalent that makes them indifferent from playing the gamble.

In general, building predictive models using training data sets is the main goal in supervised learning. Supervised learning includes different tasks, among which we find the task of classification plays an important role in theory and applications. Classification can be found in various applications ranging from engineering to medicine and bioinformatics. Supervised learning and classification will be used in this study to build a predictive model of human decisions under risk and uncertainty.

In fact, literature is overstuffed with research that tackles the classification problem in machine learning. Among such classification techniques is the Support Vector Machines (SVM) which was developed and then improved by Boser *et al.* [4] and Cortes and Vapnik [8], respectively. A comprehensive survey on different learning techniques can be found in Hastie *et al.* [14]. Proposals on the improvement of the well known learning techniques are still ongoing [24,28,32]. For example, the multi-category versions of SVM were developed recently by Lee *et al.* [18] and Zhu *et al.* [31] Other machine learning and classifications includes ψ-learning which was first proposed by Shen *et al.* [23], IVM by Zhu and Hastie [30], distance weighted discrimination (DWD) by Marron *et al.* [21] and Qiao *et al.* [22].

Although SVM is mainly introduced as a technique for solving pattern recognition problems, we find that it can remarkably perform in classifying decision choices and in predicting human attitude in uncertain or risky situations. SVM characteristics make them a good choice for predicting human preferences between prospects of uncertain outcomes. SVM's have the capability to classify inseparable data presented by the violations found in decision-making problems such as the common consequence, common ratio and the violation of betweeness.

The main contributions of the proposed model comes in three: *first*, to predict the human risk attitudes through identifying the utility function of individuals. This is achieved by optimizing an SVM model that will specify the utility parameter, *second*, predicting individual preference between uncertain outcomes via support vector machines without the need for explicit utilities. *Third*, steer the decision maker to better decisions.

This rest of the paper is organized as follows: Section 2 presents the certainty equivalents and classifications, followed by the SVM classes in Section 3. The kernel trick is illustrated in Section 4, followed by separable and inseparable data examples respectively in Sections 5 and 6. Next, the effect of the model parameters is illustrated in Section 7 followed by testing and validation experiments in Section 8. Finally, the conclusions are presented in Section 9.

2. The certainty equivalents and classification

Certainty equivalent is defined as the amount of money that a person is willing to accept instead of the prospect itself [12]). Since decision making is usually normative, certainty equivalents are difficult to be elicited via direct questionnaires. Given the uncertainty that exists in decision problems, it is difficult for individuals to be certain about the equivalent of a gamble that makes them feel indifferent.

In this research, we believe that individuals are more comfortable in declaring their preference between two lotteries as compared to providing a strict estimate of the certainty equivalent. Reported in literature, elicited certainty equivalents have a good deal of randomness and can have wide variations that condense trust in the collected data. Instead, we try to build a model by means of SVM to predict the risk attitudes in individuals and then to predict their preference among different sets of safe and risky lotteries using pair-wise comparisons.

Here, the subjects are asked to present their preference between a safe (S) and a risky lottery (R), where ($S \succ R$) means that S is preferred to R. A lottery L is denoted by its rewards and probabilities as $L(x_1, p_1; x_2, p_2; \ldots; x_n, p_n)$. The utility of a reward x is estimated by the following widely cited utility function:

$$u(x) = x^{\beta}. \tag{2.1}$$

The expected value (EV) and the expected utility (EU) of L are given by the following two expressions, respectively:

$$EV_L = \sum\nolimits_1^n x_i p_i, \tag{2.2}$$

$$EU_L = \sum\nolimits_1^n u(x_i) p_i. \tag{2.3}$$

Furthermore, it is already established in literature that the certainty equivalent of a gamble can be found by the inverse of the expected utility, that is:

$$CE_L = EU_L^{1/\beta}. \tag{2.4}$$

If a subject is requested to estimate the certainty equivalent of a lottery, then the subjects' elicited certainty equivalent is supposed to coincide with the model certainty equivalent in (2.4). Figure 1 depicts the idea where a risk neutral person ($\beta = 1$) will have a certainty equivalent which is exactly the same as the expected value. Due to risk attitudes, the subject certainty equivalents will differ from the expected values with an amount called risk premium (RP).

Accordingly, if we sketch the points of certainty equivalents of pairs of lotteries, say (S and R), on a *CE*-*CE* diagram, we expect the pairs having ($EV_S > EV_R$) to be below the iso-expected line, and those having

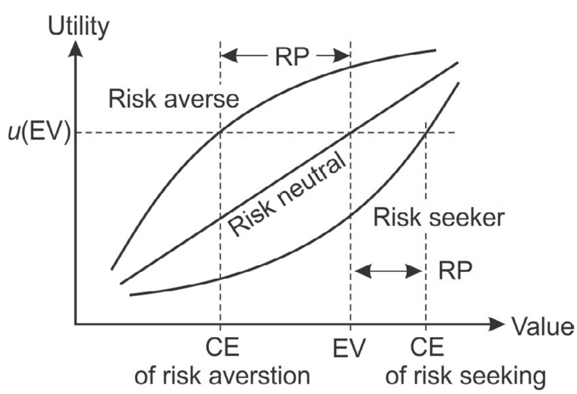

Figure 1. Risk aversion, risk seeking attitudes and the certainty equivalent illustrated.

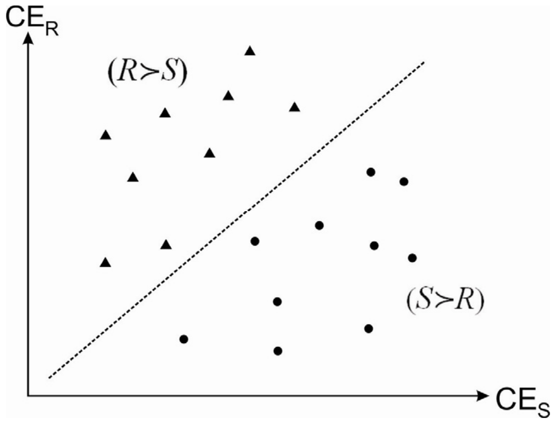

FIGURE 2. Two classes of points marked by squares and triangles around the iso-expected line.

TABLE 1. Two pairs of lotteries of equal certainty equivalents.

Lottery pair	S (\$)		Probability		R (\$)		Probability	
	x_1	x_2	p_1	p_2	x_1	x_2	p_1	p_2
1	177.78	0	0.75	0.25	1000	0	0.3162	0.6838
2	400	0	0.5	0.5	200	0	0.7071	0.2929

($EV_S < EV_R$) to be above such a line for a risk neutral person as shown in Figure 2. However, estimating the certainty equivalents is a matter of different risk attitudes, hence, different evaluations and pair mapping will result in reality. Obviously, the subjects are considered to be more consistent when they classify such pairs with more care. That is, no points overlap between the two classes (above and below the line) is found. That said, a subject whose evaluations can separate the data well has clearer risk attitude. However, real evaluations may result in linearly inseparable data (see Fig. 4). In this case, a nonlinear separator is to be found.

In fact, inseparable data results from the violations found in the EVT as well as the EUT. Finding the linear/nonlinear hyper-planes that separate the data points while taking the utility into consideration is not a trivial task. In our analysis, we realized that SVM is a very interesting and mathematically tractable approach that can admit such violations and can classify such data points via optimizing the utility function at the same time.

Consider the four lotteries in Table 1. For a utility function of the class x^{β}, when $\beta = 0.5$ all lotteries will have the same certainty equivalents of 100. Accordingly, for an individual who exhibits such utility function, he/she should be indifferent between the four lotteries.

However, in reality, many will prefer the second lottery in the first pair and the first lottery in the second pair. Figure 3 shows the perception of certainty equivalent by direct estimation of 10 subjects. Here, the subjects were asked to give their estimate of the worth of each lottery without any further knowledge. Now, if for instance the parameter β is increased, the exact certainty equivalents of both pairs will follow the shown locus. This shows that higher β values may present larger gap between the two pair points. However, due to errors in human certainty estimates and tendency to avoid risk, apparent separation between the pairs may not be present, hence, an algorithm like SVM may present the separation line by finding the best β that presents the optimal gap between the data points.

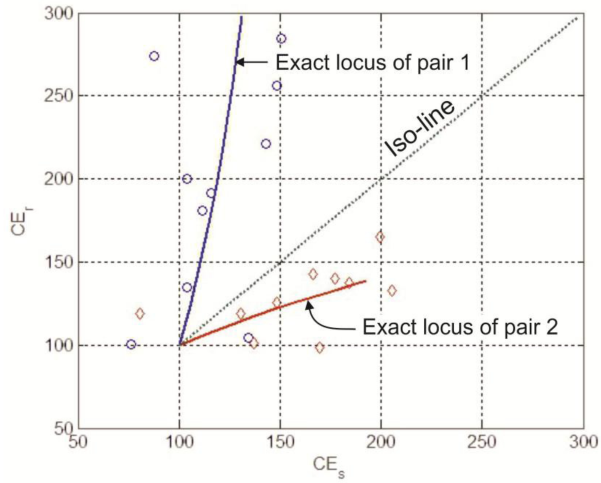

FIGURE 3. Certainty equivalents as a function of β for two lottery pairs. Diamond and round points represent the estimated certainty equivalents by direct questionnaires for 10 subjects. For instance, the elicited certainty of the first pair of lotteries of the first subject is \$75 and \$100.

3. The CE classes in SVM

Certainty equivalents usually result in inseparable data points on the *CE-CE* diagram. As a result, it will be a motivating act to find a model that copes with the existing randomness (*i.e.*, misclassification) and accepts the reported violations. Although there is no well-established notion of safe and risky prospects, in general, people tend to pick what they feel a better lottery. Accordingly, in this study, individuals will experience pairs of lotteries with different risk profiles.

When two lotteries S and R are presented to an individual, we request the individual to show his preference of one lottery over another, for instance, if someone prefers R over S, (*i.e.*, $R \succ S$) then the pair is expected to be above the iso-expected line on the *CE-CE* diagram. Similarly, if $S \succ R$ then the pair is supposed to be located below the iso-expected line. To prepare the data for SVM operation, we denote the choice by "-1" if

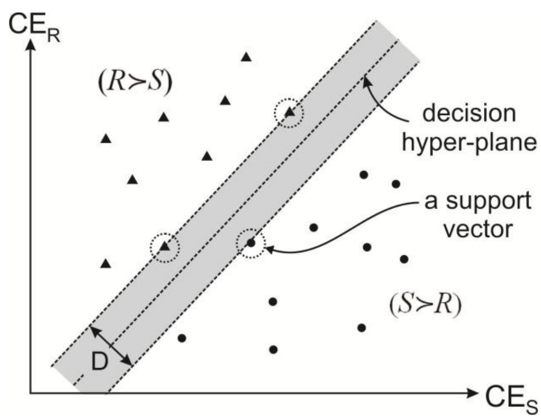

FIGURE 4. *CE-CE* diagram with the support vectors circled.

an individual prefers S over R, and by "1" if the individual preference is the other way around. Figure 4 shows a sample data, where the x and y axes represent CE_S and CE_R of the two lotteries in every pair. Note that, in the classic EVT, the 45° line is expected to separate the pairs. However, the EVT is not a superior choice anyway, nor the utility theory, due to the aforementioned violations.

Since the model certainty equivalents depend on the utility function, different values of the parameter β result in different pair locations in the *CE-CE* diagram. It will be appealing if we can find the β value that best map the data on such diagram and best separate the data into two different groups according to the choices. To find the separator line that best classifies the data and optimizes the parameter β, we need a classification approach that takes into consideration the following points:

- Acceptance of the stated violations: some lotteries with low expected utility may be selected over high expected utility and vice versa. The acceptance of such pattern includes the violation of betweenness and common consequence effects.
- If the utility of the lottery X is higher than Y and Y is higher than Z, then the model should accept to predict the preference pattern ($Z \succ Y \succ X$, *i.e.*, transitivity and completeness).
- A model that accepts soft margin between the clusters (data misclassification).
- A model that takes the expected utility parameter into consideration.
- A classification model that presents higher dimensionality in clusters (separable and inseparable data).
- No computational complications as those found in the cumulative prospect theory (CPT), the stochastic utility theory (StEUT) and the Beta stochastic utility (β-SU) models.

Although many classification algorithms exists in literature, however, it can be easily apprehended that SVM can satisfy all above conditions which fit individuals' choices and accepts the different violation patterns with soft margin clustering.

SVM is a classification approach which depends on few supporting points that best split the data into classes. In fact, SVM tries to find the band that separates the two classes and does so by maximizing the distance D between the two edges resembling the band borders as shown in Figure 4. This band represents the maximum void between data points of the two classes.

By observing Figure 4, we realize that the decision surface will be fully specified by a small subset of the data points which represent the support vectors. Clearly, the other data points have no effect on determining the decision hyper-plane. Similar approach can be applied for nonlinear case as shown in Figure 5.

Suppose that we have a set of N pairs consisting of the two lotteries S and R. The assessor will classify the pairs by "-1" or "1" where y_i $\forall i = 1, \ldots, N$ denotes such classification. Accordingly, in our model, the classification pattern will be used to find the best hyper-plane that separates the pairs and optimizes the utility

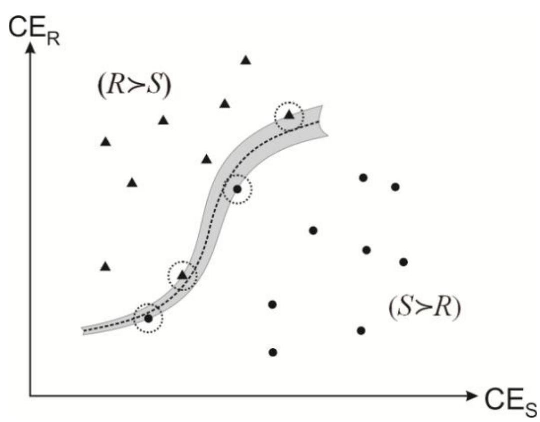

FIGURE 5. Support vectors circled on a *CE-CE* diagram of nonlinear separator.

function given in (2.1). For all pairs located below the iso-expected line, a risk neutral person will prefer $S \succ R$ as $CE_S > CE_R$, while we will find that R is preferred to S when the pairs fall above the line. However, real experiments of lottery preferences were found noticeably different.

Each data point $\boldsymbol{CE}_i$ is identified by $[CE_S\ CE_R]$ where the two components CE_S and CE_R belong to the lotteries S and R and CE is found by (2.4). Consequently, the training data set is represented by the form: $\{\boldsymbol{CE}_i, y_i\}$, $\boldsymbol{CE}_i \in \Re^2$, $y_i \in \{-1, 1\}$, $\forall i = 1, \ldots, N$.

We would like to emphasize that the location of the pairs on the *CE-CE* diagram depends on β and the separator hyper-planes will differ consequently. For a linearly separable data, a hyper-plane can be described by $\vec{w} \cdot \boldsymbol{CE} + b = 0$, where $\vec{w}$ is a coefficient vector of n elements ($n = 2$ for binary classification) and b is an intercept.

As we want to maximize the gap between the classes of $(S \succ R)$ and those of $(R \succ S)$, we can minimize $\|\vec{w}\|$ as performed in standard SVM approach. For a linearly separable data set, our model can be summarized in a primal form as follows:

$$\left.\begin{array}{l} \min \frac{1}{2} \|\vec{w}\|^2 \\ s.t. \\ y_i\left(\vec{w}.\boldsymbol{CE}_i + b\right) \geqslant 1 \quad \forall i = 1, \ldots, N \end{array}\right\}. \tag{3.1}$$

Now given a new instance $\boldsymbol{z}$, the classifier function will predict the class of this data point as "-1" if $\vec{w}\boldsymbol{z} + b \leqslant$ and as "1" if $\geqslant \vec{w} \cdot \boldsymbol{z} + b$, alternatively, this can be put across as $f\left(\boldsymbol{z}\right) = \operatorname{sign}(\vec{w} \cdot \boldsymbol{z} + b)$

The above primal optimization model in (3.1) has two decision variables, particularly, the components of $\vec{w}$ and the scalar b. The model works well when the data is linearly separable and the training certainty equivalents are available. However, the components of the point $\boldsymbol{CE}_i$ are determined by the inverse of the expected utility which is estimated *via* the parameter β.

Although the primal model can be solved by any standard optimization solver. The model can be recast in a dual form, yielding a quadratic problem of N dual variables, *i.e.*, $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_N\}$.

$$\left.\begin{array}{c} \max \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j \boldsymbol{CE}_i \cdot \boldsymbol{CE}_j \\ s.t. \\ \alpha_i \geqslant 0 \\ \sum\limits_{i=1}^{N} \alpha_i y_i = 0 \end{array}\right\}. \tag{3.2}$$

The data points are presented inside a dot product and the coefficient vector $\vec{w}$ is defined in terms of α_i as $\vec{w} = \sum_{i=1}^{N} \alpha_i y_i \boldsymbol{CE}_i$

The classifier function of any point, say $\boldsymbol{z}$, becomes:

$$f\left(\boldsymbol{z}\right) = \operatorname{sign}\left(\sum_{i=1}^{N} \alpha_i y_i \boldsymbol{CE}_i \cdot \boldsymbol{z} + b\right). \tag{3.3}$$

By implementing Lagrange relaxation and KKT conditions, the model in (3.2) can be solved by:

$$\left.\begin{array}{l} \sum\limits_{i=1}^{N} \alpha_i y_i = 0 \\ \vec{w} = \sum\limits_{i=1}^{N} \alpha_i y_i \boldsymbol{CE}_i = 0 \end{array}\right\}. \tag{3.4}$$

Hence, solving the equations in (3.4) will yield the analytical optimal solution. However, for large number of data points, optimization solvers will easily find the solution of (3.2). Now if the points are inseparable in the case of noisy data or misclassification, a slack variable ξ_i is assigned to each instance in the data. The parameter ξ_i can be thought of as the distance from the data points to the separating hyper-plane as depicted in Figure 6.

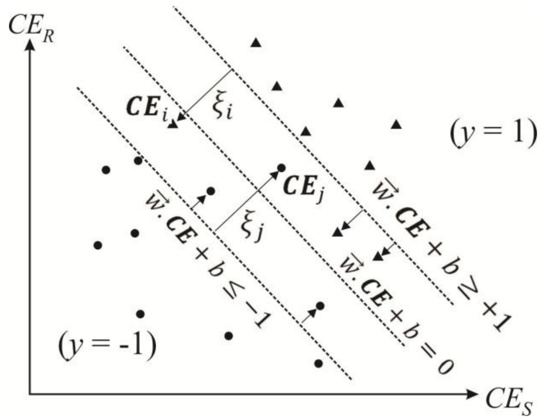

FIGURE 6. *CE*-*CE* diagram with SVM classifier of soft margins.

Mathematically, this can be represented in a dual form as:

$$\left.\begin{array}{c} \max \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1,j=1}^{N} \alpha_i \alpha_j y_i y_j \boldsymbol{CE}_i \cdot \boldsymbol{CE}_j \\ s.t. \\ C \geqslant \alpha_i \geqslant 0, \quad \forall i = 1, \ldots, N \\ \sum_{i=1}^{N} \alpha_i y_i = 0 \end{array}\right\} \tag{3.5}$$

where C is a positive tradeoff parameter. Large C values call for sharp threshold and hard margins (no misclassified data exists) which is the same as in the optimization problem in (3.2). Smaller values of C mean softer margin is selected allowing for misclassifications.

4. Inseparable data and the kernel trick

In inseparable data, the data points can be transformed into a higher dimensional space which is named as *feature space.* This operation is performed using a transformation function $\phi(\mathbf{x}_i)$ for a data set $\mathbf{x}$. By so doing, the nonlinear operation in the input space will be equivalent to linear operation in the feature space.

For data points denoted by $\mathbf{x}_i$, $\forall i = 1, \ldots, N$, a kernel function is defined as $\boldsymbol{K}(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi\ (\mathbf{x}_j)$. This is usually called the kernel trick. Power kernels are given by:

$$\boldsymbol{K}(\mathbf{x}_i, \mathbf{x}_j) = -||\mathbf{x}_i - \mathbf{x}_j||^{\beta}. \tag{4.1}$$

In our model, we will implement the power kernel given by (4.1) since it can result in an open and closed separator contours. It also provides wide variety of separation hyper-planes as compared to the radial basis kernel which tends to produce clusters bounded by closed loops. Consequently, by implementing the kernel trick, our model will look like:

$$\left.\begin{array}{c} \max \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1,j=1}^{N} \alpha_i \alpha_j y_i y_j \boldsymbol{K}(\boldsymbol{CE}_i, \boldsymbol{CE}_j) \\ C \geqslant \alpha_i \geqslant 0 \\ \sum_{i=1}^{N} \alpha_i y_i = 0 \end{array}\right\}. \tag{4.2}$$

For a test object $\boldsymbol{z}$, the discriminant function is essentially a weighted sum of the similarity between $\boldsymbol{z}$ and the support vectors. The classification function is given by:

$$\left.\begin{aligned} \boldsymbol{f} &= (w, \varphi(\boldsymbol{z})) + b = \sum_{m=1}^{s} \hat{\alpha}_m \hat{y}_m \boldsymbol{K}\left(\widehat{CE}_m, \boldsymbol{z}\right) + b \\ \boldsymbol{w} &= \sum_{m=1}^{s} \hat{\alpha}_m \hat{y}_m \varphi(\widehat{CE}_m) \end{aligned}\right\}, \tag{4.3}$$

where $\widehat{CE}_m, \hat{y}_m$ and $\hat{\alpha}_m$ are the data, the classification and the dual variables associated with the support vectors which are listed in S. If a point is not a support vector, it is associated α value will be equal to 0. The set S includes only the support vectors.

5. Linearly separable example

Although we will not use linear separators in this study, however, for the demonstration purposes, we present linearly separable instance as a first example. The model is programmed in MATLAB and the optimization is conducted using fmincon function.

Suppose the following 14 pairs of lotteries in Table 2 were presented to a subject. The subject is requested to show the preferred lotteries. The elicited classification is shown in the table. For simplicity, let $x^{0.5}$ represent the utility function. The location of the pairs is depicted in Figure 7. By applying a simple linear SVM using the model in (3.2), we find that the pairs #1, 9 and 13 are the support vectors where,

$$S = \begin{pmatrix} 198.83 & 119.98 \\ 30.45 & 52.27 \\ 16 & 8 \end{pmatrix}$$

and the corresponding solution is given by: $w_1 = -0.03458$, $w_2 = 0.05646$, $b = -0.8959$, $\alpha =$ [0.0003623 0 0 0 0 0 0 0 0.002191 0 0 0 0 0.001829] and hence, $\hat{\alpha}$ =[0.0003623 0.002191 0.001829]. Any testing pair can be classified

Table 2. Pairs of lotteries classified according to the subject. The columns show the lottery entries, the excepted value, the expected utility and the model certainty equivalent of each lottery, respectively.

	Lottery S							Lottery R							
#	x_1	x_2	p_1	p_2	EV_S	EU_S	CE_S	x_1	x_2	p_1	p_2	EV_R	EU_R	CE_R	Class
1	100	200	0.01	0.99	199	14.1	198.83	50	150	0.25	0.75	125	10.95	119.98	−1
2	100	150	0.25	0.75	137.5	11.69	136.55	0	150	0.99	0.01	1.5	0.122	0.015	−1
3	0	25	0.9	0.1	2.5	0.5	0.25	50	75	0.99	0.01	50.25	7.087	50.225	1
4	150	200	0.99	0.01	150.5	12.27	150.46	0	75	0.5	0.5	37.5	4.33	18.75	−1
5	150	200	0.25	0.75	187.5	13.67	186.83	100	150	0.99	0.01	100.5	10.02	100.45	−1
6	100	150	0.9	0.1	105	10.22	104.55	150	200	0.6	0.4	170	13.01	169.14	1
7	0	75	0.05	0.95	71.25	8.227	67.688	50	100	0.4	0.6	80	8.828	77.941	1
8	50	150	0.5	0.5	100	9.659	93.301	0	50	0.01	0.99	49.5	7	49.005	−1
9	25	50	0.75	0.25	31.25	5.518	30.446	50	75	0.9	0.1	52.5	7.23	52.273	1
10	25	50	0.9	0.1	27.5	5.207	27.114	100	150	0.9	0.1	105	10.22	104.55	1
11	25	50	0.5	0.5	37.5	6.036	36.428	0	75	0.95	0.05	3.75	0.433	0.1875	−1
12	100	150	0.99	0.01	100.5	10.02	100.45	0	75	0.99	0.01	0.75	0.087	0.0075	−1
13	0	100	0.6	0.4	40	4	16	0	50	0.6	0.4	20	2.828	8	−1
14	50	100	0.99	0.01	50.5	7.1	50.415	100	150	0.75	0.25	112.5	10.56	111.55	1

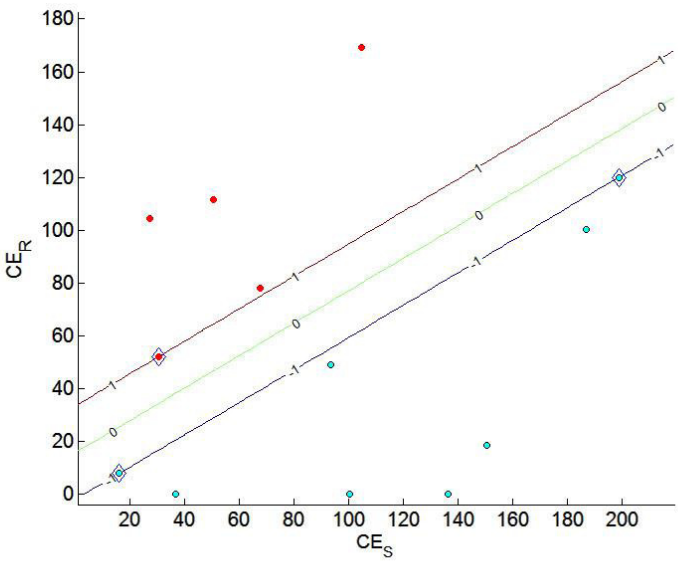

FIGURE 7. The separating line and the data points of the first example. The support vectors are marked by diamonds. The two classes are marked by different colors.

by the resulting function

$$f(\mathbf{z}) = \sum_{i=1}^{N} \alpha_i y_i \boldsymbol{CE}_i \cdot \boldsymbol{z} + b = 0.0003623\,(-1)\,(\langle 198.83 \quad 119.98\rangle \cdot \langle \boldsymbol{z}\rangle) \\ + 0.002191\,(1)\,(\langle 30.45 \quad 52.27\rangle \cdot \langle \boldsymbol{z}\rangle) + 0.001829\,(1)\,(\langle 16 \quad 8\rangle \cdot \langle \boldsymbol{z}\rangle) - 0.8959,$$

where if $f > 0$, then the pair belongs to the class denoted by "1" (*i.e.*, the subject prefers S to R) otherwise it will be in the "-1" class. For instance, the lottery pair, $S = (0, 0.25; 200, 0.75)$ and $R = (50, 0.75; 150, 0.25)$ will be classified as "-1", which means the subject is predicted to prefer S over R. Recall that, different β values result in different separators of different objective function values. In fact, our model will also search for the optimal β parameter that best classifies the data.

6. Non-linear example

Linear separators are the easiest to identify, followed by nonlinear separators of polynomial kernels. To demonstrate, consider the data in Table 2 once again. By implementing a polynomial kernel of the second degree, the following separating strip will result as shown in Figure 8. The classifier is a 3D surface in binary classifications with the "0" contour representing the decision line. Note that the border lines have to cross the support vectors which are located at the contour elevations of -1 and 1 with the separator line at the elevation of "0".

The resulting support vectors and dual variables for a β value of 0.5 and a polynomial of the second degree are:

$$S = \begin{pmatrix} 198.83 & 119.98 \\ 0.25 & 50.23 \\ 30.45 & 52.27 \\ 16 & 8 \end{pmatrix}$$

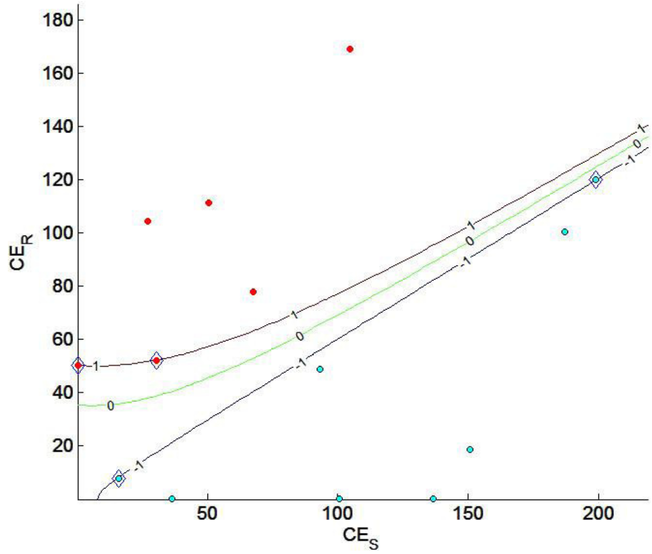

FIGURE 8. Separating hyper-plane and the separating borders. The support vectors are marked by diamond.

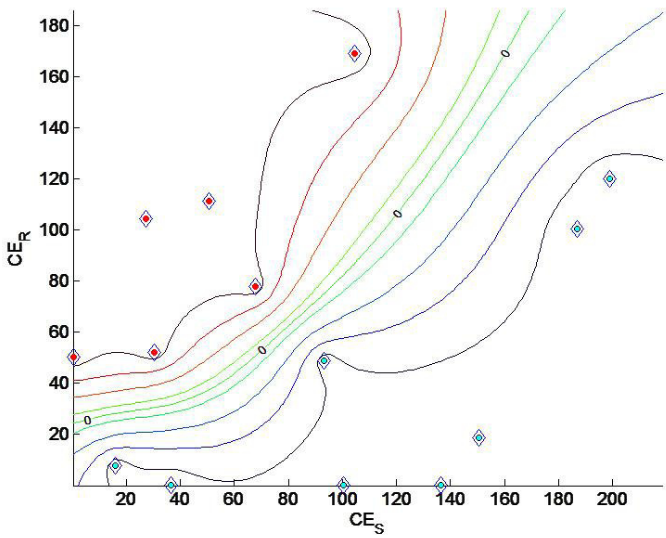

FIGURE 9. Results using the power kernel. The separating contour at elevation of "0" and other contour elevations are demonstrated.

and $b = -0.9750$, $\boldsymbol{\alpha} = [0.00012 \quad 0 \quad 0.00129 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0.00239 \quad 0 \quad 0 \quad 0 \quad 0.00356 \quad 0]$, hence $\hat{\boldsymbol{\alpha}} = [0.00012 \quad 0.00129 \quad 0.00239 \quad 0.00356]$. For example, the lottery pair, $S = (0, 0.25; 200, 0.75)$ and $R = (50, 0.75; 150, 0.25)$ will be classified as "-1".

Although polynomial kernels can provide a variety of nonlinear separators, yet, there are some more powerful kernels that can predict more efficiently, such as the power kernel given in (4.1). By considering the same data in Table 2 using a kernel of the power 0.5, we get the contours shown in Figure 9 with the "0" contour as the

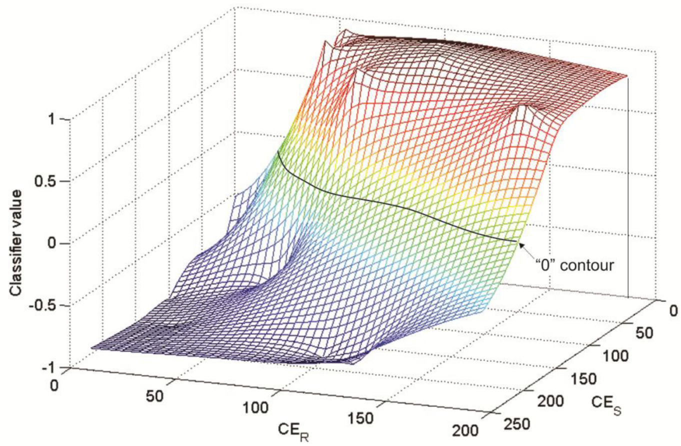

FIGURE 10. The classifier function and the separating contour at "0" elevation.

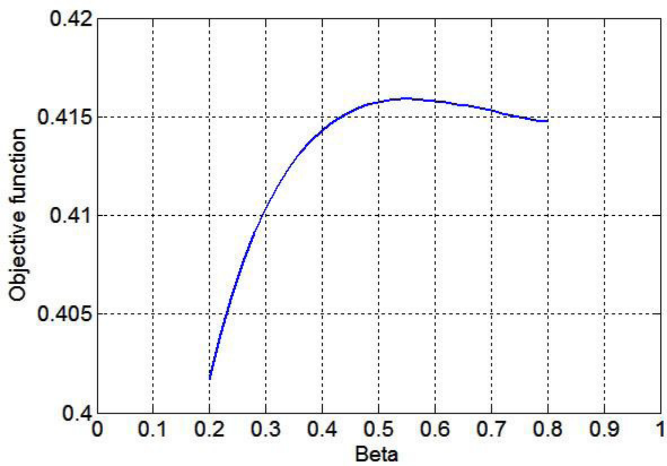

FIGURE 11. The objective function value *vs.* β using a power kernel.

separator line. Note that, for instances with larger number of lottery pairs, this kernel can be a good choice. The power kernel tends to increase the number of support vectors. Note, all the data points are support vectors (it will not be the case for well clustered points). In this example, the borders of the contours -1 and 1 are tightly bounding the support vectors. If at least one point was not a support vector, we will see such borders clearly.

The resulting parameters of the above case are: $b = -0.2404$ where S includes all the data points here, accordingly, $\hat{\boldsymbol{\alpha}} = [0.0493\ 0.0021\ 0.0909\ 0.0103\ 0.0380\ 0.0770\ 0.1139\ 0.1244\ 0.1034\ 0.0087\ 0.0568\ 0.0166\ 0.1182\ 0.0219]$ which consists of 14 elements. Figure 10 shows a 3D plot of the classifier function and the separator plane at "0" elevation.

For the above case, the objective function of (4.3) equals to (+0.4158). However, such separation process will result in different separator for different utility functions. Figure 11 shows the objective function for different β values. The optimal utility parameter β for this case is around 0.56.

TABLE 3. A set of classified lottery pairs.

	Lottery S				Lottery R				
#	x_1	x_2	p_1	p_2	x_1	x_2	p_1	p_2	Class
	100	200	0.01	0.99	50	150	0.25	0.75	-1
	100	150	0.25	0.75	0	150	0.99	0.01	-1
	150	200	0.99	0.01	0	75	0.5	0.5	-1
	150	200	0.25	0.75	100	150	0.99	0.01	-1
	30	100	0.95	0.05	0	50	0.01	0.99	-1
	25	50	0.5	0.5	0	75	0.95	0.05	-1
	100	150	0.99	0.01	0	75	0.99	0.01	-1
	0	100	0.6	0.4	0	50	0.6	0.4	-1
	0	25	0.9	0.1	50	75	0.99	0.01	1
	100	150	0.9	0.1	150	200	0.6	0.4	1
	0	75	0.05	0.95	50	100	0.4	0.6	1
	25	50	0.75	0.25	20	75	0.9	0.1	1
	25	50	0.9	0.1	100	150	0.9	0.1	1
	50	100	0.99	0.01	100	150	0.75	0.25	1

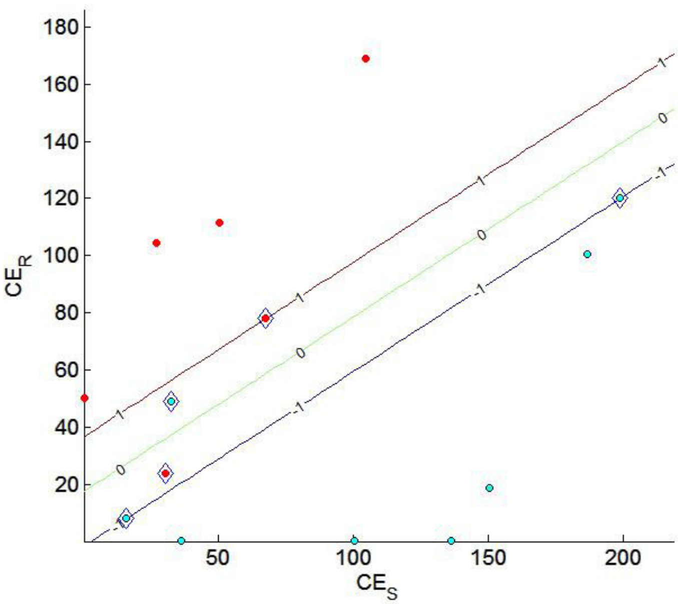

FIGURE 12. A soft margin classifier of Table 3 ($C = 1$).

7. THE EFFECT OF THE PARAMETERS ON THE MODEL PREDICTION

Essentially, the constant C in the SVM model of (4.3) is a measure of the separation softness. Low C values dictates softer margins, however, in decision making, this may be important since humans commit errors when they estimate the certainty equivalents or perform pair-wise comparisons. In fact, it is reported that almost 25% percent of individuals switch their preferences when they are surveyed twice for the same pairs of lotteries.

To demonstrate the effect of the parameter C consider Table 3 which shows a set of lottery pairs that have been classified by an individual. For $\beta = 0.5$ and $C = 1$, we get the results shown in Figure 12, where all data points within the contours (1 and -1) are found to be supporting vectors.

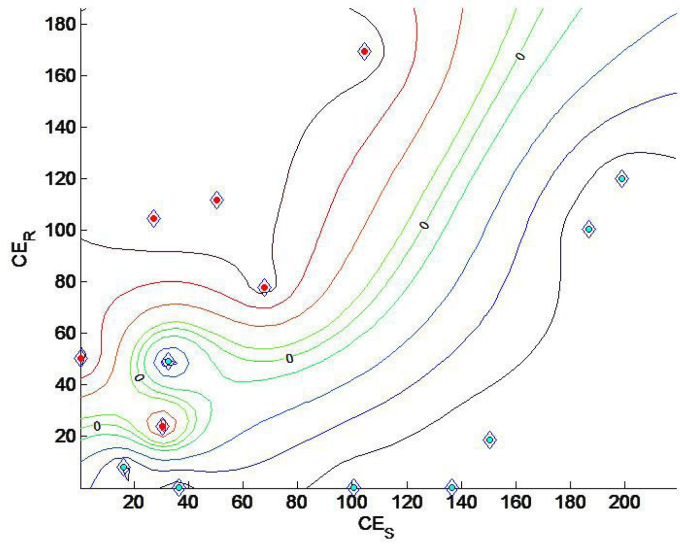

FIGURE 13. A hard margin classifier of Table 3 ($C = \infty$).

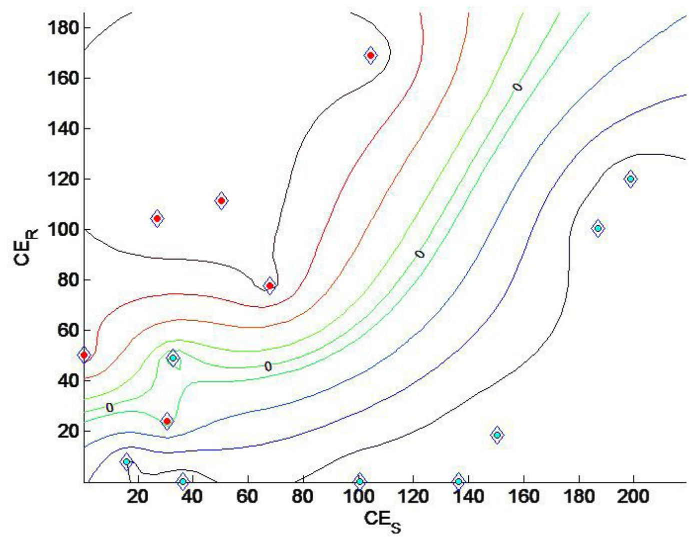

FIGURE 14. A classifier of Table 3 ($C = 0.1$). The support vector includes all the data points, $\beta = 0.5$, $b = -0.54378$ and $\hat{\boldsymbol{\alpha}} = [0.0513\ 0.0102\ 0.0237\ 0.0445\ 0.1000\ 0.0724\ 0.0366\ 0.1000\ 0.1000\ 0.0745\ 0.1000\ 0.1000\ 0.0354\ 0.0286]$.

To demonstrate the effect of the parameter C using a power kernel let us start with unbounded $\boldsymbol{\alpha}$ values (*i.e.*, $C = \infty$). The resulting separating contours are shown in Figure 13. The "0" contour represents the separating hyper-plane. Indeed, an infinite value of C results in perfect classification.

Now, by changing the value of C, say to 0.1, the resulting separator hyper-plane tends to be smoother as shown in Figure 14. In this case, misclassification is allowed in the training set and hence it is expected in the testing sets.

The test data can be examined by the separator function given in (4.3). Suppose now we want to use the same data points as a testing set. When $C = \infty$, none is predict mistakenly, while we find one data point is misclassified when $C = 0.1$, which is just fine in human preferences as there are no hard margins in reality.

8. Testing and validation of the model

In this study, we implement the idea of cross validation to test the prediction power of the model. However, misclassification in training data is expected; hence, the percent of correct predictions is not insistently the sole measure of goodness when the data is inseparable. Anyway, the SVM model is powerful in predicting the utility parameters based on pair-wise comparison as compared to surveys of strict certainty equivalent estimation.

To validate the model, we use different collections of data sets, each will be used for training once, while the others for testing. Figure 15 shows 6 data sets, each time, a set is used in training in some trials and for testing in others. Accordingly, we will record the correct prediction, the percent correct and the grand average of each scenario. Test data sets are classified according to the classifier in (4.3).

8.1. Experimental results

In this section, different data sets are used for testing and validation, some are presented in the literature, others are of our own. First, we start with a survey we conducted to collect individual preferences, the subjects were provided by pairs of lotteries to show their preference. The lotteries were designed as shown in Figure 16, where the subject has to check a preferred lottery among each pair. The lottery pairs and the subjects' preferences are shown in Tables 4–9 each for one subject.

The SVM model is applied to the different data sets shown in Tables 4–9. For each data set, the power kernel is implemented as a similarity measure while the objective function value is used to search for the optimal utility parameter. The optimal values of β are listed in the first column of Table 10. Figures 17a–17f show the contours

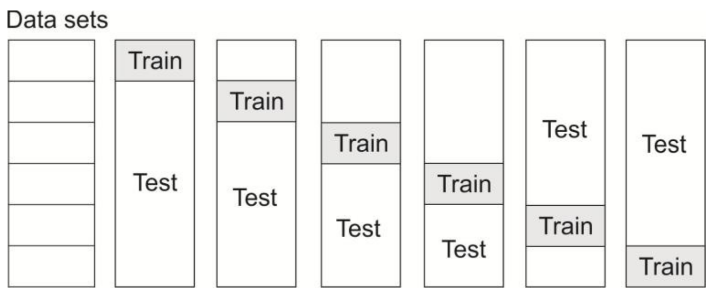

Figure 15. Cross validation of training and testing sets.

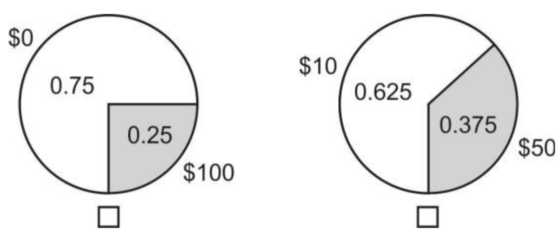

Figure 16. The questionnaire style. The subject is requested to check the preferred lottery.

TABLE 4. First conducted survey of different lotteries (data set 1).

#	S				R				Subject class
	x_1	x_2	p_1	p_2	x_1	x_2	p_1	p_2	
1	100	150	0.05	0.95	25	50	0.75	0.25	−1
2	0	100	0.1	0.9	25	50	0.05	0.95	−1
3	50	150	0.95	0.05	0	150	0.9	0.1	−1
4	50	75	0.01	0.99	50	75	0.5	0.5	−1
5	0	150	0.95	0.05	100	200	0.01	0.99	1
6	100	200	0.1	0.9	0	400	0.5	0.5	−1
7	0	200	0.5	0.5	50	150	0.5	0.5	1
8	0	800	0.05	0.95	50	75	0.01	0.99	−1
9	50	100	0.6	0.4	50	75	0.75	0.25	−1
10	100	200	0.4	0.6	100	150	0.1	0.9	−1
11	0	800	0.25	0.75	0	150	0.01	0.99	−1
12	25	50	0.01	0.99	0	400	0.75	0.25	1
13	0	400	0.95	0.05	100	150	0.4	0.6	1
14	50	100	0.4	0.6	0	100	0.05	0.95	−1
15	0	200	0.6	0.4	0	25	0.6	0.4	−1
16	0	100	0.95	0.05	0	800	0.99	0.01	1
17	50	100	0.25	0.75	0	400	0.95	0.05	−1
18	0	400	0.75	0.25	0	400	0.99	0.01	−1
19	0	75	0.25	0.75	0	75	0.6	0.4	−1
20	0	50	0.1	0.9	25	50	0.25	0.75	1
21	0	400	0.01	0.99	150	200	0.99	0.01	−1
22	50	150	0.1	0.9	150	200	0.75	0.25	1
23	0	200	0.4	0.6	0	25	0.25	0.75	−1
24	0	800	0.95	0.05	0	50	0.25	0.75	−1
25	150	200	0.4	0.6	0	150	0.95	0.05	−1
26	100	200	0.5	0.5	50	150	0.05	0.95	−1
27	25	50	0.6	0.4	0	75	0.1	0.9	−1
28	100	150	0.95	0.05	0	200	0.75	0.25	−1
29	100	150	0.75	0.25	150	200	0.5	0.5	1
30	100	200	0.6	0.4	100	200	0.1	0.9	1

and separation surfaces. Note that the power of the kernel does not affect the output as long as the power is less than 1 for a positive definite kernel.

For testing, each data set from the Tables 4–9 will be used to test the data of the others as illustrated in Figure 15. Here, we have six model instances, hence, each model will be tested 5 times. For instance, we will use the prediction model of Table 4 (data set #1) to predict the other data sets. Table 10 shows the correct predictions. For example 29 cases out of 30 were predicted correctly when we use model 2 to predict the data set 3. The percent correct of each data set is shown in Table 11 with the average in the last row. Although the main purpose of the presented SVM model is to identify the utility curve, we can also reveal remarkable predictions.

TABLE 5. Second conducted survey of different lotteries (data set 2).

	S				*R*				
#	x_1	x_2	p_1	p_2	x_1	x_2	p_1	p_2	Class
1	50	100	0.01	0.99	0	150	0.5	0.5	−1
2	0	75	0.01	0.99	100	200	0.9	0.1	1
3	100	200	0.99	0.01	0	400	0.05	0.95	1
4	50	100	0.95	0.05	100	150	0.25	0.75	1
5	0	150	0.25	0.75	0	200	0.99	0.01	−1
6	150	200	0.95	0.05	50	100	0.05	0.95	−1
7	0	800	0.6	0.4	0	75	0.9	0.1	−1
8	0	400	0.4	0.6	0	100	0.95	0.05	−1
9	50	150	0.01	0.99	0	150	0.4	0.6	−1
10	100	150	0.1	0.9	0	150	0.25	0.75	−1
11	0	400	0.6	0.4	100	150	0.6	0.4	−1
12	0	150	0.9	0.1	0	150	0.6	0.4	1
13	0	200	0.25	0.75	50	150	0.75	0.25	1
14	100	200	0.95	0.05	0	200	0.4	0.6	−1
15	50	75	0.25	0.75	50	150	0.9	0.1	1
16	0	200	0.95	0.05	0	200	0.5	0.5	1
17	50	150	0.25	0.75	50	100	0.6	0.4	−1
18	50	75	0.95	0.05	150	200	0.01	0.99	1
19	50	150	0.4	0.6	0	100	0.99	0.01	−1
20	150	200	0.75	0.25	0	50	0.9	0.1	−1
21	50	75	0.1	0.9	50	150	0.95	0.05	−1
22	50	75	0.6	0.4	50	100	0.1	0.9	1
23	0	50	0.99	0.01	0	200	0.01	0.99	1
24	50	75	0.05	0.95	0	800	0.4	0.6	1
25	50	150	0.75	0.25	0	25	0.4	0.6	−1
26	0	50	0.25	0.75	0	100	0.6	0.4	1
27	50	100	0.9	0.1	0	200	0.6	0.4	−1
28	50	75	0.75	0.25	0	100	0.75	0.25	−1
29	150	200	0.6	0.4	25	50	0.1	0.9	−1
30	0	400	0.5	0.5	0	400	0.9	0.1	−1

Clearly, we can notice the high prediction percents. The training data of each model are all predicted correctly. The grand average of the prediction percent of the 6 sets is around 87%. In all the data sets, it was noticed that the utility parameter ranges between 0.52 to 0.85.

In the next experiment, we use a well known set from Loomes and Sugden [19], here, 46 subjects were asked to present their preferences in 45 pairs of different lotteries. The payoffs of all the lotteries were a combination of £0, £10, £20 with different probabilities as shown in Table 12. Table 13 shows the actual classifications.

For each column in the responses, the SVM model is applied and tested against the remaining data sets, this results in $46 \times 46 = 2116$ tests. Table 14 shows the counts of correct predictions, and most importantly, the utility parameter β of each data set in the last column. Surprisingly, in all data sets, the optimal β value did range between 0.58 to 0.74, a result that reveals risk aversion in all subjects. The total percent of correct

TABLE 6. Third conducted survey of different lotteries (data set 3).

	S				R				
	x_1	x_2	p_1	p_2	x_1	x_2	p_1	p_2	Class
1	100	200	0.01	0.99	50	150	0.25	0.75	−1
2	0	75	0.4	0.6	100	150	0.01	0.99	1
3	0	100	0.5	0.5	50	100	0.75	0.25	1
4	100	150	0.25	0.75	0	150	0.99	0.01	−1
5	0	25	0.9	0.1	50	75	0.99	0.01	1
6	150	200	0.99	0.01	0	75	0.5	0.5	−1
7	150	200	0.25	0.75	100	150	0.99	0.01	−1
8	100	150	0.9	0.1	150	200	0.6	0.4	1
9	0	75	0.05	0.95	50	100	0.4	0.6	1
10	0	50	0.4	0.6	50	75	0.05	0.95	1
11	0	800	0.99	0.01	150	200	0.1	0.9	1
12	0	100	0.75	0.25	100	200	0.75	0.25	1
13	0	800	0.9	0.1	0	50	0.05	0.95	−1
14	0	100	0.9	0.1	0	800	0.25	0.75	1
15	0	25	0.75	0.25	50	100	0.01	0.99	1
16	0	800	0.01	0.99	0	100	0.9	0.1	−1
17	0	75	0.95	0.05	0	200	0.05	0.95	1
18	0	25	0.1	0.9	0	75	0.05	0.95	1
19	50	150	0.5	0.5	0	50	0.01	0.99	−1
20	25	50	0.75	0.25	50	75	0.9	0.1	1
21	25	50	0.9	0.1	100	150	0.9	0.1	1
22	0	400	0.99	0.01	0	200	0.25	0.75	1
23	0	50	0.01	0.99	0	400	0.01	0.99	1
24	0	150	0.6	0.4	50	150	0.01	0.99	1
25	25	50	0.5	0.5	0	75	0.95	0.05	−1
26	0	150	0.75	0.25	50	150	0.6	0.4	1
27	100	150	0.99	0.01	0	75	0.99	0.01	−1
28	0	100	0.6	0.4	0	50	0.6	0.4	−1
29	50	100	0.99	0.01	100	150	0.75	0.25	1
30	0	75	0.6	0.4	50	150	0.99	0.01	1

predictions is 73% which strongly agrees with the literature results where almost 25% of subjects tend to switch their choices for the same lotteries when repeated in the experiment.

For further demonstration, we arbitrarily picked a case to show the prediction contours of the model. Figures 18 and 19 show the model of the selected case #10. The 3D plot of the classifier shows the points classified by "1" in the top half.

In our last experiment, we use the data set of Tversky and Kahneman [27] for lotteries of positive outcomes as shown in Table 15. The resulting separator is shown in Figure 20. Note that all the data points are found to be support vectors with the optimal β value of 0.675 as shown in Figure 21.

TABLE 7. Fourth conducted survey of lotteries (data set 4).

	S				R				
	x_1	x_2	p_1	p_2	x_1	x_2	p_1	p_2	Class
1	50	100	0.1	0.9	50	100	0.5	0.5	−1
2	0	25	0.05	0.95	0	25	0.9	0.1	−1
3	0	150	0.99	0.01	50	150	0.1	0.9	1
4	100	150	0.5	0.5	100	150	0.95	0.05	−1
5	0	200	0.9	0.1	0	50	0.4	0.6	−1
6	0	75	0.1	0.9	0	50	0.99	0.01	−1
7	25	50	0.99	0.01	0	400	0.6	0.4	1
8	0	150	0.1	0.9	0	800	0.9	0.1	−1
9	0	150	0.4	0.6	100	200	0.25	0.75	1
10	0	50	0.6	0.4	0	25	0.95	0.05	−1
11	0	200	0.75	0.25	100	150	0.5	0.5	1
12	50	75	0.5	0.5	0	25	0.1	0.9	−1
13	0	200	0.01	0.99	0	75	0.01	0.99	−1
14	0	400	0.9	0.1	50	100	0.99	0.01	1
15	25	50	0.4	0.6	0	50	0.75	0.25	−1
16	0	100	0.05	0.95	0	25	0.05	0.95	−1
17	0	25	0.95	0.05	0	100	0.01	0.99	1
18	100	200	0.75	0.25	0	200	0.9	0.1	−1
19	50	150	0.9	0.1	25	50	0.9	0.1	−1
20	150	200	0.01	0.99	100	200	0.4	0.6	−1
21	0	400	0.1	0.9	100	200	0.95	0.05	−1
22	0	150	0.01	0.99	0	150	0.75	0.25	−1
23	0	75	0.75	0.25	0	25	0.01	0.99	1
24	0	25	0.01	0.99	150	200	0.95	0.05	1
25	50	150	0.99	0.01	25	50	0.99	0.01	−1
26	50	100	0.5	0.5	25	50	0.01	0.99	−1
27	0	25	0.99	0.01	100	200	0.99	0.01	1
28	150	200	0.05	0.95	50	150	0.4	0.6	−1
29	0	50	0.95	0.05	0	800	0.5	0.5	1

For the validation purposes, we have used the found model in this experiment and tested the data of the conducted survey in Table 4–9. The number of test data points and predicted counts are shown in Table 16.

Once again, the β value is found to be less than 1, which reveals risk aversion attitude. The percent of correct prediction is 93% for the used data sets.

Finally, we would like to emphasize the contribution of this study characterized by introducing the SVM to the field of decision analysis to perform supervised learning of human decision making. The approach can be used to guide individuals to make better decisions. It predicts and optimizes the utility without the need to elicit firm certainty equivalents.

TABLE 8. Fifth conducted survey (data set 5).

	S				R				
	x_1	x_2	p_1	p_2	x_1	x_2	p_1	p_2	Class
1	0	100	0.4	0.6	0	50	0.1	0.9	−1
2	0	25	0.25	0.75	0	100	0.25	0.75	1
3	0	150	0.05	0.95	100	200	0.05	0.95	1
4	50	150	0.05	0.95	150	200	0.05	0.95	1
5	0	75	0.99	0.01	50	75	0.95	0.05	1
6	0	75	0.9	0.1	0	400	0.25	0.75	1
7	0	800	0.5	0.5	25	50	0.95	0.05	−1
8	0	100	0.01	0.99	0	25	0.5	0.5	−1
9	0	50	0.05	0.95	0	800	0.75	0.25	−1
10	100	200	0.05	0.95	0	50	0.95	0.05	−1
11	50	150	0.6	0.4	100	200	0.5	0.5	1
12	0	800	0.1	0.9	0	150	0.1	0.9	−1
13	0	50	0.9	0.1	100	150	0.05	0.95	1
14	50	100	0.05	0.95	0	25	0.99	0.01	−1
15	50	75	0.9	0.1	150	200	0.25	0.75	1
16	100	200	0.25	0.75	0	75	0.4	0.6	−1
17	0	75	0.5	0.5	0	75	0.25	0.75	1
18	0	100	0.99	0.01	50	100	0.95	0.05	1
19	0	400	0.05	0.95	100	200	0.6	0.4	−1
20	50	75	0.99	0.01	0	100	0.4	0.6	−1
21	50	100	0.75	0.25	0	75	0.75	0.25	−1
22	150	200	0.1	0.9	50	100	0.25	0.75	−1
23	0	50	0.75	0.25	50	75	0.25	0.75	1
24	25	50	0.1	0.9	50	75	0.4	0.6	1
25	100	150	0.6	0.4	150	200	0.9	0.1	1
26	25	50	0.25	0.75	150	200	0.4	0.6	1

TABLE 9. Sixth conducted survey (data set 6).

	S				R				
#	x_1	x_2	p_1	p_2	x_1	x_2	p_1	p_2	Class
1	0	200	0.05	0.95	0	800	0.6	0.4	−1
2	0	200	0.1	0.9	0	800	0.6	0.4	1
3	0	25	0.5	0.5	0	25	0.75	0.25	−1
4	0	400	0.25	0.75	0	200	0.95	0.05	−1
5	25	50	0.95	0.05	0	800	0.01	0.99	1
6	100	150	0.4	0.6	50	100	0.9	0.1	−1
7	25	50	0.05	0.95	50	75	0.6	0.4	1
8	100	150	0.01	0.99	0	50	0.5	0.5	−1
9	0	50	0.5	0.5	25	50	0.6	0.4	1
10	50	75	0.4	0.6	0	100	0.5	0.5	−1
11	150	200	0.9	0.1	0	150	0.05	0.95	−1
12	0	25	0.6	0.4	50	75	0.1	0.9	1
13	0	200	0.99	0.01	0	200	0.1	0.9	1
14	0	25	0.4	0.6	25	50	0.4	0.6	1
15	0	100	0.25	0.75	0	100	0.1	0.9	1
16	0	200	0.1	0.9	0	800	0.95	0.05	−1
17	0	150	0.5	0.5	0	800	0.1	0.9	1
18	150	200	0.5	0.5	0	400	0.4	0.6	−1

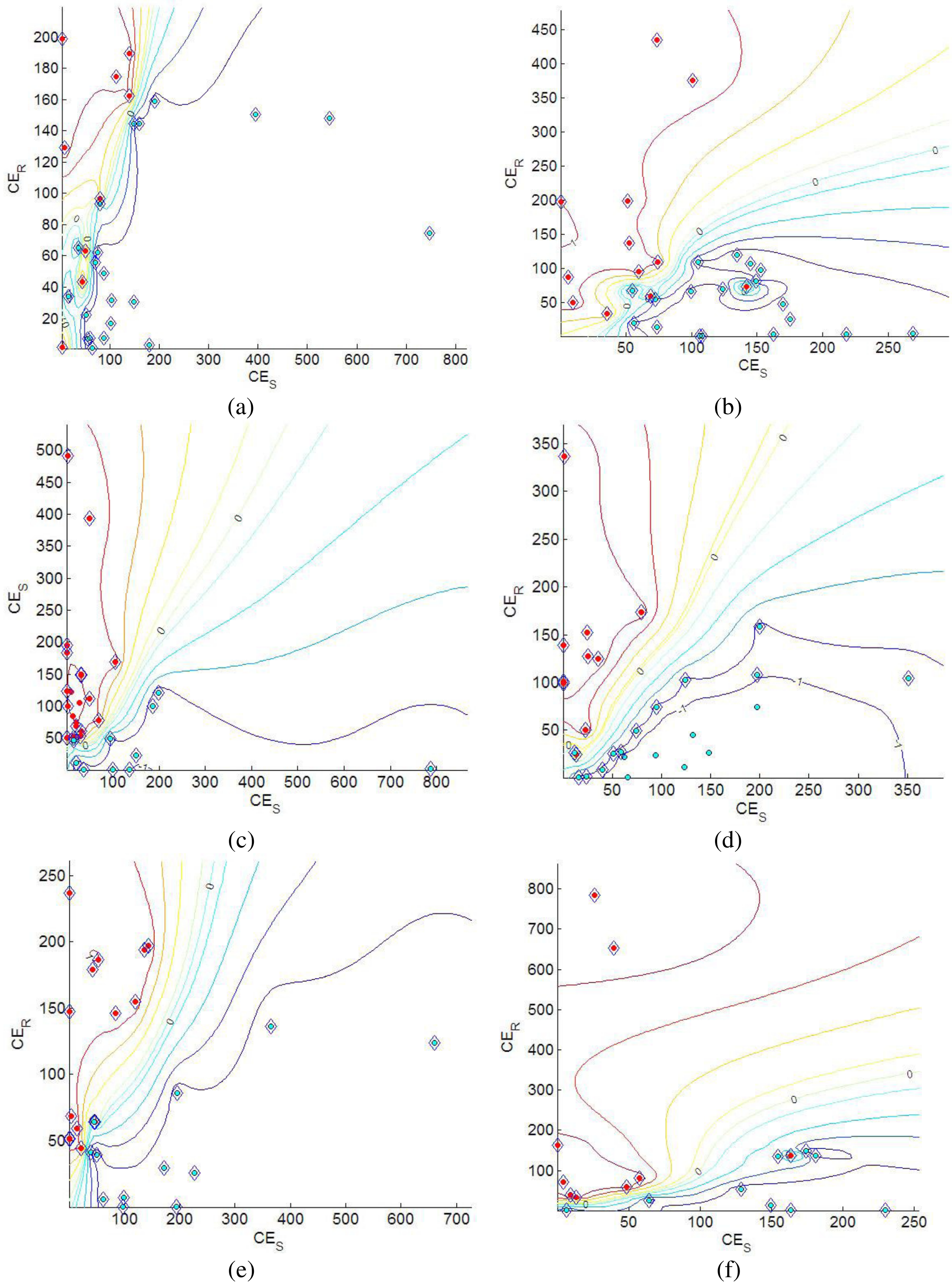

FIGURE 17. The resulting classifiers of the 6 data sets.

Table 10. The optimal parameter β and the correctly predicted counts.

Data set	β	1	2	3	4	5	6
1	0.75	30	25	22	26	18	11
2	0.84	25	30	29	26	23	14
3	0.59	23	26	30	28	23	15
4	0.8	23	27	29	29	21	15
5	0.55	23	26	29	27	26	14
6	0.52	21	26	29	28	20	18

Table 11. Percent correct and the averages.

	Data sets						
Data sets	1	2	3	4	5	6	average
1	1	0.83	0.73	0.9	0.69	0.61	0.79
2	0.83	1	0.97	0.9	0.88	0.78	0.89
3	0.77	0.87	1	0.97	0.88	0.83	0.89
4	0.77	0.9	0.97	1	0.81	0.83	0.88
5	0.77	0.87	0.97	0.93	1	0.78	0.88
6	0.7	0.87	0.97	0.97	0.77	1	0.88
					Grand average	=	0.87

Table 12. The 45 pair of lotteries of the rewards (£0, £10, £20).

#	Safe lottery S	Risky lottery R	#	Safe lottery S	Risky lottery R	#	Safe lottery S	Risky lottery R
1	(10,0.25;20,0.75)	(0,0.15;20,0.85)	16	(0,0.8;10,0.2)	(0,0.9;20,0.1)	31	(0,0.4;10,0.6)	(0,0.5;10,0.3;20,0.2)
2	(0,0.15;10,0.25;20,0.6)	(0,0.3;20,0.7)	17	(10,0.25;20,0.75)	(0,0.1;20,0.9)	32	(0,0.7;10,0.3)	(0,0.8;20,0.2)
3	(10,0.5;20,0.5)	(0,0.3;20,0.7)	18	(0,0.1;10,0.75;20,0.15)	(0,0.4;20,0.6)	33	(10,0.5;20,0.5)	(0,0.1;20,0.9)
4	(10,0.5;20,0.5)	(0,0.15;10,0.25;20,0.6)	19	(10,1)	(0,0.4;20,0.6)	34	(0,0.1;10,0.5;20,0.4)	(0,0.2;20,0.8)
5	(10,1)	(0,0.15;10,0.75;20,0.1)	20	(10,1)	(0,0.1;10,0.75;20,0.15)	35	(10,1)	(0,0.2;20,0.8)
6	(10,1)	(0,0.6;20,0.4)	21	(0,0.6;10,0.25;20,0.15)	(0,0.7;20,0.3)	36	(10,1)	(0,0.1;10,0.5;20,0.4)
7	(0,0.15;10,0.75;20,0.1)	(0,0.6;20,0.4)	22	(0,0.5;10,0.5)	(0,0.7;20,0.3)	37	(0,0.25;10,0.75)	(0,0.35;10,0.25;20,0.4)
8	(0,0.75;10,0.25)	(0,0.9;20,0.1)	23	(0,0.5;10,0.5)	(0,0.6;10,0.25;20,0.15)	38	(0,0.25;10,0.75)	(0,0.4;20,0.6)
9	(10,0.2;20,0.8)	(0,0.1;20,0.9)	24	(0,0.75;10,0.25)	(0,0.85;20,0.15)	39	(0,0.35;10,0.25;20,0.4)	(0,0.4;20,0.6)
10	(0,0.1;10,0.8;20,0.1)	(0,0.5;20,0.5)	25	(10,0.3;20,0.7)	(0,0.1;20,0.9)	40	(0,0.5;10,0.5)	(0,0.6;20,0.4)
11	(10,1)	(0,0.5;20,0.5)	26	(0,0.2;10,0.6;20,0.2)	(0,0.4;20,0.6)	41	(10,0.25;20,0.75)	(10,0.3;20,0.7)
12	(10,1)	(0,0.1;10,0.8;20,0.1)	27	(0,0.1;10,0.9)	(0,0.4;20,0.6)	42	(0,0.55;10,0.2;20,0.25)	(0,0.65;10,0.15;20,0.2)
13	(0,0.5;10,0.4;20,0.1)	(0,0.7;20,0.3)	28	(0,0.1;10,0.9)	(0,0.2;10,0.6;20,0.2)	43	(0,0.8;20,0.2)	(0,0.85;20,0.15)
14	(0,0.4;10,0.6)	(0,0.7;20,0.3)	29	(0,0.5;10,0.3;20,0.2)	(0,0.6;20,0.4)	44	(0,0.1;10,0.75;20,0.15)	(0,0.15;10,0.75;20,0.1)
15	(0,0.4;10,0.6)	(0,0.5;10,0.4;20,0.1)	30	(0,0.4;10,0.6)	(0,0.6;20,0.4)	45	(0,0.7;10,0.3)	(0,0.75;10,0.25)

TABLE 13. Subjects' responses to the 45 problems. "1" means the lottery "R" has been selected.

Lottery pair	Responses of the the 46 subjects	Lottery pair
1	−1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1	1
2	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	2
3	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	3
4	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 1 −1 −1	4
5	−1 1 −1 −1 1 1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1	5
6	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	6
7	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	7
8	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	8
9	−1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	9
10	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	10
11	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1	11
12	−1 1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 1 −1 1 −1 −1 1 −1 −1 −1 −1 −1 1 1 −1 1 −1 −1 1 1 −1 1 −1 −1 −1 −1 −1 −1 −1	12
13	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	13
14	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	14
15	1 −1 1 −1 1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 1 −1 −1 1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 1 1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 1	15
16	−1 −1 1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 1 −1 1 −1 −1 1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1	16
17	−1 1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 −1 −1 1 −1 −1 1 1 1 1 −1 −1 −1 1 −1 −1 −1 −1 −1 1 −1 −1 −1	17
18	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1	18
19	−1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	19
20	−1 1 −1 −1 1 1 −1 1 −1 −1 −1 1 −1 −1 −1 1 1 −1 −1 1 1 1 −1 1 −1 1 −1 −1 −1 1 1 1 1 1 −1 1 1 1 1 −1 −1 −1 −1 −1 −1 −1	20
21	−1 −1 −1 −1 1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 1 −1 −1 −1 1 −1 −1 −1 −1 1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1	21
22	−1 −1 1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 1 −1	22
23	1 −1 1 1 1 1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 1 −1 1 1 −1 1 −1 1 1 1 −1 1 1 −1 1 −1 −1 −1 −1 1 −1 −1 −1	23
24	1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 1 1 −1 −1 −1 1 −1 1 −1 1 1 −1 −1 −1 1 −1 1 −1 −1 −1 −1 1 −1 −1 1	24
25	−1 −1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 1 −1 1 1 −1 −1 −1 −1 1 −1 −1 −1 −1 1 −1 −1 −1	25
26	−1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 1 −1	26
27	−1 −1 −1 −1 1 −1 −1 1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 1 −1 1 −1 −1 −1 −1 −1 −1 1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1	27
28	1 1 1 1 1 1 −1 1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 1 −1 1 −1 −1 1 −1 1 1 1 −1 1 1 −1 1 1 −1 1 −1 1 −1 −1 −1 1 1	28
29	1 −1 1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 1 1 −1 −1 −1 −1 −1 1 −1 1 1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 1 −1 1 −1	29
30	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 1 −1	30
31	−1 −1 1 1 1 1 −1 1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 1 1 −1 −1 −1 1 −1 1 1 1 1 1 1 1 1 1 −1 −1 −1 −1 −1 1 −1 1 1	31
32	1 −1 1 1 1 1 −1 1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 −1 −1 1 −1 1 −1 1 1 −1 1 −1 1 1 1 1 1 −1 −1 1 −1 −1 −1 −1 1 −1 1 1	32
33	1 1 −1 1 1 1 1 1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 1 1 −1 −1 1 −1 1 1 1 1 1 1 −1 −1 −1 1 −1 −1 −1 −1 −1 1 −1 1 −1	33
34	1 1 −1 −1 1 1 −1 −1 −1 1 −1 −1 −1 −1 −1 1 1 −1 1 1 1 1 1 1 −1 1 −1 1 1 −1 1 1 1 −1 1 1 1 −1 −1 −1 −1 −1 1 −1 −1 1	34
35	−1 1 −1 1 1 −1 1 1 −1 1 −1 −1 −1 −1 −1 −1 1 −1 1 1 −1 1 1 1 −1 1 −1 −1 1 1 1 1 1 1 −1 1 1 −1 1 −1 −1 −1 1 −1 −1 −1	35
36	1 1 −1 1 1 1 1 −1 −1 1 −1 1 1 −1 −1 −1 1 −1 −1 1 1 1 1 1 −1 1 −1 −1 1 1 1 1 1 1 −1 1 1 1 1 −1 −1 −1 1 −1 1 1	36
37	1 1 −1 1 1 1 1 1 −1 1 1 −1 −1 −1 −1 1 −1 −1 1 1 −1 1 1 1 −1 1 1 −1 1 1 1 1 1 1 −1 1 1 −1 1 −1 −1 −1 1 −1 1 1	37
38	−1 −1 −1 1 1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 1 −1 −1 −1 −1 1 −1 −1 1 1 1 1 1 −1 1 1 −1 −1 −1 −1 −1 1 −1 −1 1	38
39	1 1 −1 1 1 1 −1 1 −1 1 −1 −1 1 −1 −1 1 −1 −1 1 1 1 1 1 −1 −1 1 1 1 1 1 1 1 1 1 1 −1 1 −1 1 −1 −1 −1 1 −1 1 1	39
40	1 −1 1 1 1 1 1 1 −1 1 −1 −1 −1 −1 −1 1 −1 −1 1 1 −1 1 1 −1 −1 1 1 −1 1 1 1 1 1 1 1 1 1 1 1 −1 −1 −1 1 −1 −1 1	40
41	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	41
42	−1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	42
43	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	43
44	−1 −1 −1 −1 −1 −1 1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	44
45	−1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1 −1	45

TABLE 14. SVM utility model test results. Few cases have one class, hence were eliminated, such as case #9.

	Subjects																																															
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	% correct	β
1	45	34	36	38	32	34	35	34	33	36	35	33	35	33	31	35	32	33	36	35	33	33	36	34	33	39	33	35	41	28	28	31	31	34	37	34	31	36	35	33	34	33	32	30	36	40	0.76	0.7
2	34	45	27	35	31	37	34	35	34	35	34	36	36	34	32	36	39	34	31	36	40	32	35	35	34	38	28	36	34	37	27	26	36	35	30	33	40	31	40	34	35	34	25	33	31	33	0.76	0.7
3	36	27	45	35	27	29	32	35	36	31	36	34	34	36	34	32	31	36	33	30	30	28	29	31	36	32	34	32	34	25	25	26	28	35	38	33	26	35	32	36	37	36	29	33	35	35	0.72	0.58
4	38	35	35	45	33	35	36	39	34	41	36	32	36	34	32	32	33	34	37	36	32	34	35	33	34	40	36	34	40	35	27	30	36	41	36	33	36	35	38	34	35	34	33	31	35	39	0.78	0.7
5	32	31	27	33	45	33	24	31	22	35	24	26	24	22	22	28	27	22	33	32	28	34	31	31	22	32	26	24	32	33	33	36	38	31	30	33	32	27	30	22	23	22	33	19	29	31	0.64	0.74
6	34	37	29	35	33	45	28	33	28	33	30	30	30	28	28	32	31	28	31	30	34	32	31	31	28	36	30	32	34	37	29	28	38	33	32	31	36	33	34	28	29	28	27	29	31	33	0.7	0.74
7	35	34	32	36	24	28	45	34	37	34	37	35	37	37	35	33	36	37	30	33	33	31	36	30	37	37	29	35	35	30	22	27	27	36	31	30	31	36	37	37	36	37	30	34	32	34	0.74	0.74
8	34	35	35	39	31	33	34	45	34	39	34	32	34	34	32	34	33	34	35	38	32	34	33	33	34	38	34	34	36	33	27	28	36	39	34	31	34	33	38	34	35	34	31	31	33	35	0.76	0.58
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
10	36	35	31	41	35	33	34	39	32	45	32	30	34	32	30	32	33	32	37	38	32	36	37	33	32	38	34	34	38	33	27	32	36	39	36	31	36	31	36	32	33	32	35	29	33	39	0.76	0.58
11	35	34	36	36	24	30	37	34	43	32	45	39	41	43	41	37	38	43	30	31	37	25	32	34	43	37	33	39	33	28	18	23	27	34	37	30	29	38	37	43	42	43	26	40	32	34	0.77	0.74
12	33	36	34	32	26	30	35	32	41	30	39	45	41	41	39	37	40	41	26	31	41	25	30	34	41	35	27	37	29	30	20	23	27	34	35	32	31	38	37	41	40	41	22	38	32	34	0.76	0.74
13	35	36	34	36	24	30	37	34	43	34	41	41	45	43	41	37	40	43	28	33	41	27	34	32	43	37	31	41	33	30	18	23	27	36	37	28	31	38	39	43	42	43	26	40	34	36	0.79	0.7
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
15	31	32	34	32	22	28	35	32	43	30	41	39	41	43	45	37	38	43	26	29	37	25	30	30	43	33	29	39	29	28	14	21	25	32	35	26	27	38	35	43	42	43	24	40	30	32	0.74	0.59
16	35	36	32	32	28	32	33	34	37	32	37	37	37	37	37	45	38	37	32	33	39	27	32	34	37	37	29	37	31	30	22	25	29	34	37	32	33	34	39	37	36	37	24	34	30	36	0.75	0.7
17	32	39	31	33	27	31	36	33	40	33	38	40	40	40	38	38	45	40	29	34	42	28	33	37	40	38	26	38	32	31	21	24	30	35	34	33	34	37	40	40	39	40	25	37	29	33	0.77	0.68
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
19	36	31	33	37	33	31	30	35	28	37	30	26	28	28	26	32	29	28	45	32	28	30	33	31	28	34	38	30	36	29	31	30	34	33	36	35	32	29	32	28	29	28	33	25	29	37	0.7	0.7
20	35	36	30	36	32	30	33	38	31	38	31	31	33	31	29	33	34	31	32	45	33	37	34	32	31	37	29	33	35	32	30	31	35	36	29	34	35	30	37	31	32	31	34	28	30	34	0.73	0.58
21	33	40	30	32	28	34	33	32	39	32	37	41	41	39	37	39	42	39	28	33	45	27	32	34	39	37	27	39	31	32	22	23	31	34	35	32	35	36	39	39	38	39	24	36	30	34	0.76	0.7
22	33	32	28	34	34	32	31	34	25	36	25	25	27	25	25	27	28	25	30	37	27	45	34	28	25	33	29	27	35	32	32	39	37	36	27	30	33	30	31	25	26	25	38	24	32	34	0.68	0.74
23	36	35	29	35	31	31	36	33	32	37	32	30	34	32	30	32	33	32	33	34	32	34	45	27	32	36	32	34	38	31	27	34	30	33	34	31	34	31	34	32	31	32	33	31	31	35	0.73	0.74
24	34	35	31	33	31	31	30	33	32	33	34	34	32	32	30	34	37	32	31	32	34	28	27	45	32	36	26	30	34	27	25	26	32	33	32	33	30	33	36	32	33	32	25	29	29	33	0.71	0.7
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
26	39	38	32	40	32	36	37	38	35	38	37	35	37	35	33	37	38	35	34	37	37	33	36	36	35	45	31	37	39	32	26	31	35	38	37	32	35	38	39	35	34	35	32	32	32	36	0.79	0.74
27	33	28	34	36	26	30	29	34	31	34	33	27	31	31	29	29	26	31	38	29	27	29	32	26	31	31	45	29	33	28	28	29	31	34	35	30	29	32	31	31	32	31	32	30	30	36	0.69	0.7
28	35	36	32	34	24	32	35	34	41	34	39	37	41	41	39	37	38	41	30	33	39	27	34	30	41	37	29	45	33	30	20	23	27	32	37	26	31	34	35	41	40	41	26	40	34	34	0.76	0.7
29	41	34	34	40	32	34	35	36	31	38	33	29	33	31	29	31	32	31	36	35	31	35	38	34	31	39	33	33	45	30	28	31	33	36	35	34	33	34	35	31	32	31	34	28	34	38	0.75	0.7
30	28	37	25	35	33	37	30	33	28	33	28	30	30	28	28	30	31	28	29	32	32	32	31	27	28	32	28	30	30	45	29	26	36	35	26	33	40	29	36	28	29	28	25	29	29	31	0.68	0.7
31	28	27	25	27	33	29	22	27	16	27	18	20	18	16	14	22	21	16	31	30	22	32	27	25	16	26	28	20	28	29	45	34	32	27	22	31	28	21	24	16	17	16	31	17	27	27	0.55	0.74
32	31	26	26	30	36	28	27	28	21	32	23	23	23	21	21	25	24	21	30	31	23	39	34	26	21	31	29	23	31	26	34	45	33	30	29	28	27	28	25	21	20	21	40	20	30	30	0.61	0.58
33	31	36	28	36	38	38	27	36	25	36	27	27	27	25	25	29	30	25	34	35	31	37	30	32	25	35	31	27	33	36	32	33	45	34	29	34	37	30	33	25	26	25	34	24	28	32	0.69	0.74
34	34	35	35	41	31	33	36	39	34	39	34	34	36	34	32	34	35	34	33	36	34	36	33	33	34	38	34	32	36	35	27	30	34	45	34	33	36	35	40	34	35	34	31	31	35	39	0.77	0.74
35	37	30	38	36	30	32	31	34	37	36	37	35	37	37	35	37	34	37	36	29	35	27	34	32	37	37	35	37	35	26	22	29	29	34	45	30	29	36	33	37	36	37	30	34	32	38	0.75	0.7
36	34	33	33	33	33	31	30	31	28	31	30	32	28	28	26	32	33	28	35	34	32	30	31	33	28	32	30	26	34	33	31	28	34	33	30	45	36	31	34	28	29	28	29	25	27	35	0.69	0.74
37	31	40	26	36	32	36	31	34	29	36	29	31	31	29	27	33	34	29	32	35	35	33	34	30	29	35	29	31	33	40	28	27	37	36	29	36	45	28	37	29	30	29	26	30	28	34	0.71	0.7
38	36	31	35	35	27	33	36	33	38	31	38	38	38	38	38	34	37	38	29	30	36	30	31	33	38	38	32	34	34	29	21	28	30	35	36	31	28	45	36	38	37	38	29	35	29	35	0.75	0.74
39	35	40	32	38	30	34	37	38	37	36	37	37	39	37	35	39	40	37	32	37	39	31	34	36	37	39	31	35	35	36	24	25	33	40	33	34	37	36	45	37	38	37	26	34	32	36	0.79	0.7
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
41	34	35	37	35	23	29	36	35	44	33	42	40	42	44	42	36	39	44	29	32	38	26	31	33	44	34	32	40	32	29	17	20	26	35	36	29	30	37	38	44	45	44	23	41	33	35	0.77	0.58
42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0
43	32	25	29	33	33	27	30	31	24	35	26	22	26	24	24	24	25	24	33	34	24	38	33	25	24	32	32	26	34	25	31	40	34	31	30	29	26	29	26	24	23	24	45	21	31	31	0.64	0.74
44	30	33	33	31	19	29	34	31	42	29	40	38	40	42	40	34	37	42	25	28	36	24	31	29	42	32	30	40	28	29	17	20	24	31	34	25	30	35	34	42	41	42	21	45	31	31	0.73	0.59
45	36	31	35	35	29	31	32	33	32	33	32	32	34	32	30	30	29	32	29	30	30	32	31	29	32	32	30	34	34	29	27	30	28	35	32	27	28	29	32	32	33	32	31	31	45	35	0.7	0.7
46	40	33	35	39	31	33	34	35	34	39	34	34	36	34	32	36	33	34	37	34	34	34	35	33	34	36	36	34	38	31	27	30	32	39	38	35	34	35	36	34	35	34	31	31	35	45	0.77	0.7
																																														Grand average of percent correct =	0.73	

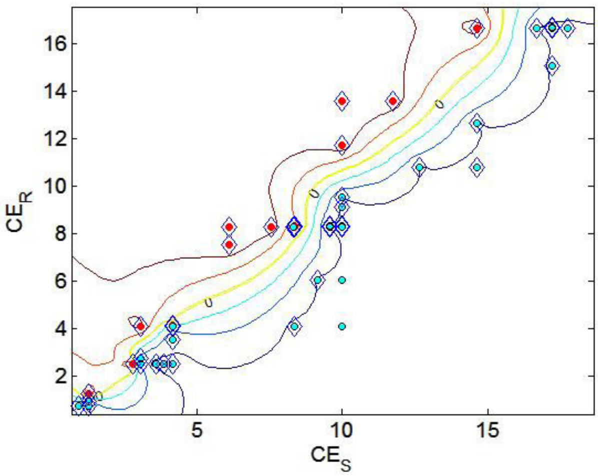

FIGURE 18. The classifier of contours of Table 12.

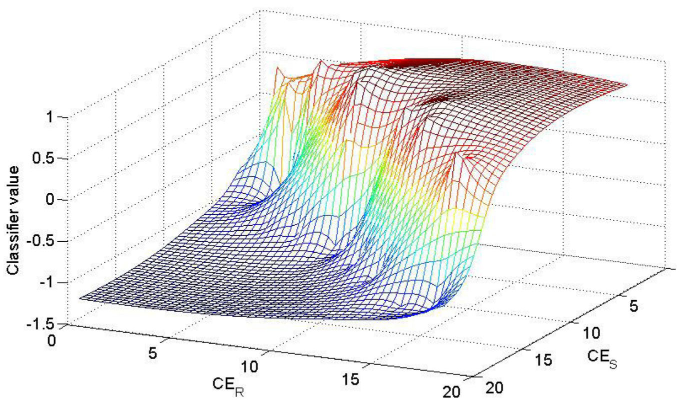

FIGURE 19. The resulting classifier of Table 12.

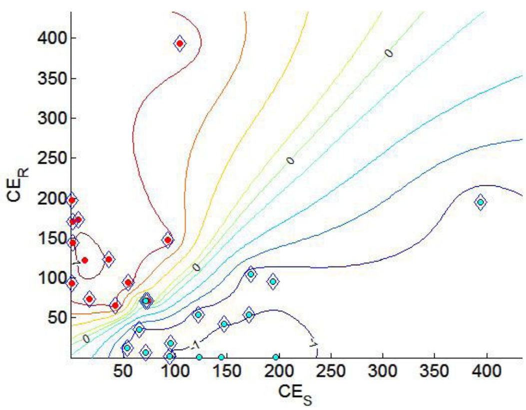

FIGURE 20. The classifier of the data in Table 15.

TABLE 15. A data set from Tversky and Kahneman [27].

	Lottery S				Lottery R				
Pair #	x_1	x_2	p_1	p_2	x_1	x_2	p_1	p_2	Class
1	50	100	0.1	0.9	0	50	0.9	0.1	-1
2	50	150	0.5	0.5	0	50	0.5	0.5	-1
3	100	200	0.5	0.5	0	50	0.1	0.9	-1
4	0	200	0.01	0.99	0	100	0.95	0.05	-1
5	50	150	0.95	0.05	0	100	0.75	0.25	-1
6	0	100	0.25	0.75	0	100	0.5	0.5	-1
7	0	50	0.1	0.9	0	100	0.25	0.75	1
8	0	200	0.99	0.01	0	100	0.05	0.95	1
9	50	150	0.05	0.95	0	200	0.99	0.01	-1
10	0	200	0.5	0.5	0	200	0.9	0.1	-1
11	50	100	0.5	0.5	0	200	0.5	0.5	1
12	0	100	0.95	0.05	0	200	0.1	0.9	1
13	0	400	0.99	0.01	0	200	0.01	0.99	1
14	100	200	0.75	0.25	0	400	0.99	0.01	-1
15	100	200	0.95	0.05	0	400	0.01	0.99	1
16	50	150	0.25	0.75	50	100	0.9	0.1	-1
17	0	50	0.5	0.5	50	100	0.5	0.5	1
18	50	100	0.9	0.1	50	100	0.1	0.9	1
19	0	200	0.1	0.9	50	150	0.95	0.05	-1
20	50	150	0.75	0.25	50	150	0.75	0.25	-1
21	100	200	0.05	0.95	50	150	0.5	0.5	-1
22	0	100	0.75	0.25	50	150	0.25	0.75	1
23	0	50	0.9	0.1	50	150	0.05	0.95	1
24	100	200	0.25	0.75	100	200	0.95	0.05	-1
25	0	100	0.5	0.5	100	200	0.75	0.25	1
26	0	100	0.05	0.95	100	200	0.5	0.5	1
27	0	200	0.9	0.1	100	200	0.25	0.75	1
28	0	400	0.01	0.99	100	200	0.05	0.95	-1

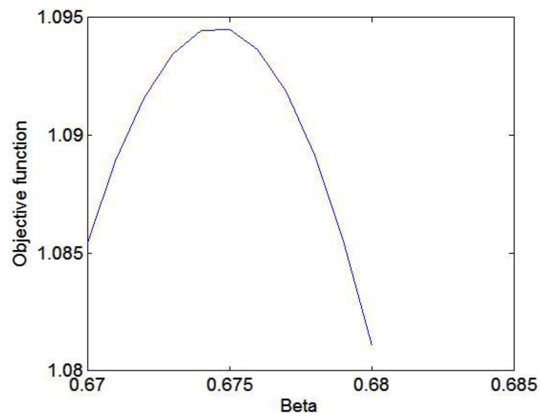

FIGURE 21. The objective function value *vs.* β.

Table 16. Test results of Table 4–9.

	Data sets (1 to 6) in Tables 4–9						
	1	2	3	4	5	6	
	30	30	30	29	26	18	β
SVM model of the Tables 4–9	25	26	29	28	25	18	0.675
Percent correct	0.83	0.87	0.97	0.97	0.96	1	
Grand average	0.93						

9. Conclusions

This paper addressed the problem of modeling human decision making under risk and uncertainty using a novel SVM approach. Supervise learning is implemented to imitate human attitude toward uncertainty in a model that helps individuals make better decisions. Unlike traditional utility estimation techniques of direct inquiry of certainty equivalents, here, the utility parameter and certainty equivalents can be computed via pair-wise comparison between lotteries. The compared lottery pairs were divided into two classes according the pattern of preference. By introducing such classes, SVM could be used to separate the data while searching for the optimal utility function parameters. Different kernels were used in the SVM model, however, a positive definite power kernel was found to be superior in its prediction and utility estimation.

Different data sets were used to train and test the model. In all data sets including the new survey and selected literature studies, the β value was ranging between 0.52–0.84, which reveals the tendency to risk aversion in the subjects. The results did demonstrate a promising model that can admit the violations to the EVT and EUT, can estimate the certainty equivalent and the utility parameter and at the same time possess remarkable prediction.

References

[1] M. Allais, Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'École Américaine. *Econometrica* **21** (1953) 503–546.

[2] P.R. Blavatskyy, Violations of Betweenness or Random Errors? *Econ. Lett.* **91** (2006) 34–38.

[3] P.R. Blavatskyy, Stochastic Expected Utility Theory. *J. Risk Uncertain* **34** (2007) 259–286.

[4] B.E. Boser, I.M. Guyon and V.N. Vapnik, A training algorithm for optimal margin classifiers. In *Proc. of the Fifth Annual ACM Workshop on Computational Learning Theory.* New York, NY (1992).

[5] D. Butlera and G. Loomesb, Imprecision as an account of violations of independence and betweenness. *J. Econ. Behav. Organiz.* **80** (2011) 511–522.

[6] M. Colyvan, No Expectations. *Mind* **116** (2006) 695–702.

[7] M. Colyvan, Relative Expectation Theory. *J. Philos.* **105** (2008) 37–44.

[8] C. Cortes and V. Vapnik, Support-vector network. *Mach. Learn.* **20** (1995) 273–297.

[9] D. Dalalah, W.A. Al-Rawabdeh and H. Alshraideh, The beta stochastic utility (β-SU). *Stoch. Anal. Appl.* **34** (2016) 456–482.

[10] K. Easwaran, Strong and Weak Expectations. *Mind* **117** (2008) 633–641.

[11] K. Easwaran, Regularity and Hyperreal Credences. *Philos. Rev.* **123** (2014) 1–41.

[12] U.W. Elke, A.-R.E. Blais and N.E. Betz, A Domain-specific Risk-attitude Scale: Measuring Risk Perceptions and Risk Behaviors. *J. Behav. Decis. Making* **15** (2002) 263–290.

[13] T. Fine, Evaluating the Pasadena, Altadena, and St Petersburg Gambles. *Mind* **117** (2008) 613–632.

[14] T. Hastie, R. Tibshirani and J. Friedman, The elements of statistical learning: Data mining inference and prediction, 2nd edition. Springer, New York (2009).

[15] C.A. Holt and S.K. Laury, Risk aversion and incentive effects. *Am. Econ. Rev.* **92** (2002) 1644–1655.

[16] D. Kahneman and A. Tversky, Judgment Under Uncertainty: Heuristics and Biases. *Sci. Mag.* **185** (1974) 1124–1131.

[17] S.K. Laury and C.A. Holt, Further reflections on prospect theory. *Exp. Econ.* **12** (2008) 405–40.

[18] Y. Lee, Y. Lin and G. Wahba, Multicategory support vector machines, theory, and application to the classification of microarray data and satellite radiance data. *J. Am. Stat. Assoc.* **99** (2004) 67–81.

[19] G. Loomes and R. Sugden, Testing Different Stochastic Specifications of Risky Choice. *Economica* **65** (1998) 581–598.

[20] J.M. Machina, Choice under Uncertainty: Problems Solved and Unsolved. *J. Econ. Perspect.* **1** (1987) 121–154.

[21] J. Marron, Distance-weighted discrimination. *J. Am. Stat. Assoc.* **102** (2007) 1267–1271.

[22] X. Qiao, H.H. Zhang, Y. Liu, M.J. Todd and J.S. Marron, Asymptotic properties of distance-weighted discrimination. *J. Am. Stat. Assoc.* **105** (2010) 401–414.

[23] X. Shen, G.C. Tseng, X. Zhang and W.H. Wong, On ψ-Learning. *J. Am. Stat. Assoc.* **98** (2003) 724–734.

[24] M. Siamak and A.K.J. Suykens, Learning solutions to partial differential equations using LS-SVM. *Neurocomputing* **159** (2015) 105–116.

[25] C. Starmer, Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk. *J. Econ. Literature* **38** (2000) 332–382.

[26] M. Thalos and O. Richardson, Capitalization in the St. Petersburg game: Why statistical distributions matter. Politics Philosophy Economics (2013) 1–22.

[27] A. Tversky and D. Kahneman, Advances in Prospect Theory: Cumulative Representation of Uncertainty. *J. Risk Uncertain* **5** (1992) 297–323.

[28] Zhigang Wang, Zengshum Zhao, Shifeng Weng and Changshui Zhang, Solving one-class problem with outlier examples by SVM. *Neurocomputing* **149** (2015) 100–105.

[29] G. Wu, J. Zhang and R. Gonzalez, Decision under Risk. In The Blackwell Handbook of Judgment and Decision Making, edited by D. Koehler and N. Harvey. Blackwell, Oxford, England (2004) 399–423.

[30] J. Zhu and T. Hastie, Kernel Logistic Regression and the Import Vector Machine. *J. Comput. Graph. Stat.* **14** (2005) 185–205.

[31] J. Zhu, H. Zou, S. Rosset, T. Hastie, Multi-class adaboost. *Stat. Interface* **2** (2009) 349–360.

[32] Li Ziqiang, Mingtian Zhou, LiJia Xu, Hao Lin, Haibo Pu, Training sparse SVM on the core sets of fitting-planes. *Neurocomputing* **130** (2014) 20–27.