



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

An out-of-sample framework for TOPSIS-based classifiers with application in bankruptcy prediction

Citation for published version:

Ouenniche, J, Pérez-Gladish, B & Bouslah, K 2018, 'An out-of-sample framework for TOPSIS-based classifiers with application in bankruptcy prediction', *Technological Forecasting and Social Change*, vol. 131, pp. 111-116. <https://doi.org/10.1016/j.techfore.2017.05.034>

Digital Object Identifier (DOI):

[10.1016/j.techfore.2017.05.034](https://doi.org/10.1016/j.techfore.2017.05.034)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Technological Forecasting and Social Change

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



An Out-of-Sample Framework for TOPSIS-based Classifiers
with Application in Bankruptcy Prediction

Jamal Ouenniche¹ (Jamal.Ouenniche@ed.ac.uk)
University of Edinburgh, Business School
29 Buccleuch Place, Edinburgh EH8 9JS, United Kingdom

Blanca Perez-Gladish (bperez@uniovi.es)
University of Oviedo, Faculty of Economics and Business,
Oviedo, Asturias 33006, Spain

Kais Bouslah (kbhb@st-andrews.ac.uk)
University of St Andrews, School of Management,
Gateway Building, North Haugh, St Andrews KY16 9RJ, United Kingdom

Abstract: Since the publication of the seminal paper by Hwang and Yoon (1981) proposing Technique for Order Performance by the Similarity to Ideal Solution (TOPSIS), a substantial number of papers used this technique in a variety of applications requiring a ranking of alternatives. Very few papers use TOPSIS as a classifier (e.g. Wu and Olson, 2006; Abd-El Fattah, 2013) and report a good performance as in-sample classifiers. However, in practice, its use in predicting discrete variables such as risk class belonging is limited by the lack of an out-of-sample evaluation framework. In this paper, we fill this gap by proposing an integrated in-sample and out-of-sample framework for TOPSIS classifiers and test its performance on a UK dataset of bankrupt and non-bankrupt firms listed on the London Stock Exchange (LSE) during 2010-2014. Empirical results show an outstanding predictive performance both in-sample and out-of-sample and thus opens a new avenue for research and applications in risk modelling and analysis using TOPSIS as a non-parametric classifier and makes it a real contender in industry applications in banking and investment. In addition, the proposed framework is robust to a variety of implementation decisions.

Keywords: Out-of-Sample Prediction, TOPSIS Classifier, k-Nearest Neighbour Classifier, Bankruptcy, Risk Class Prediction

¹ Corresponding author: Jamal Ouenniche
Email address: Jamal.Ouenniche@ed.ac.uk

1. Introduction

Multi-criteria decision analysis (MCDA) methodologies are widely used for addressing a variety of problems; namely, selection problems, ranking problems, sorting problems, classification problems, clustering problems, and description problems, where selection problems are concerned with identifying the best alternative or a subset of best alternatives; ranking problems are concerned with constructing a rank ordering of alternatives from best to worst; sorting problems are concerned with classifying alternatives into pre-defined and ordered homogenous groups or classes; classification problems are concerned with classifying alternatives into pre-defined and unordered homogenous classes; clustering problems are concerned with classifying alternatives into not pre-defined and not ordered homogenous classes; and description problems are concerned with identifying major distinguishing features of alternatives and perform their description based on these features. In this paper, we are focusing on the solution of classification problems, or equivalently predicting class belonging. To be more specific, we are concerned with the implementation of classifiers and their performance evaluation both in-sample and out-of-sample.

One popular MCDA methodology is Technique for Order Performance by the Similarity to Ideal Solution (TOPSIS) proposed by Hwang and Yoon (1981) and used in many application areas – see Behzadian et al (2012) for a review including a sample of application areas. This methodology was originally designed for solving ranking problems. In fact, TOPSIS provides a ranking of alternatives based on similarity scores, where the similarity score of each alternative is a function of the distances between the alternative and a couple of benchmarks commonly referred to as the positive and the negative ideal solutions. Later on, TOPSIS has been adapted for solving classification problems. However, to the best of our knowledge, TOPSIS classifiers and their performance evaluation has so far been restricted to in-sample analyses only (e.g., Tansel IÇ and Yurdakul, 2010). In sum, an out-of-sample framework for TOPSIS as a classifier is lacking. The aim of this paper is to fill this gap by proposing a new integrated framework for implementing a full classification analysis; namely, in-sample classification and out-of-sample classification. The proposed framework is intended to make TOPSIS classifiers real contenders in practice and to increase confidence in their use in a variety of critical application areas such as the prediction of risk class belonging (e.g., bankruptcy prediction, distress prediction, fraud detection, credit scoring).

The remainder of this paper unfolds as follows. In section 2, we provide a detailed description of the proposed integrated in-sample and out-of-sample framework for TOPSIS classifiers and discuss implementation decisions. In section 3, we empirically test the performance of the proposed framework in bankruptcy prediction of companies listed on the London Stock Exchange (LSE) and report on our findings. Finally, section 4 concludes the paper.

2. An Integrated In-Sample – Out-of-Sample Framework for TOPSIS Classifiers

In the forecasting literature, nowadays prediction models – whether designed for predicting a continuous variable (e.g., the level or volatility of the price of a strategic commodity such as crude oil) or a discrete one (e.g., risk class belonging of companies listed on a stock exchange) – have to be implemented both in-sample and out-of-sample to assess their ability to reproduce or forecast the response variable in the training sample and to forecast the response variable in the test sample, respectively. The rationale behind the necessary implementation and performance evaluation of prediction models both in-sample and out-of-sample lies in the fact that if you feed a properly designed prediction model with some information, it should be able to reproduce/predict that information; therefore, in real life settings, in-sample performance is not enough to qualify a prediction model as a good one. Because the future is unknown, out-of-sample implementations and evaluations are used to simulate the future. Out-of-sample implementation and evaluation frameworks are available for parametric prediction models (e.g. statistical models); however, this is not the case for all non-parametric ones (e.g., TOPSIS classifiers).

Hereafter, we shall present our integrated implementation and evaluation framework for TOPSIS classifiers – see Figure 1 for a graphical depiction of the process. For illustration purposes, we shall customize the presentation of the proposed framework to a bankruptcy application where we reproduce a classical bankruptcy prediction model; namely, the multivariate discriminant analysis (MDA) model of Taffler (1984), within a TOPSIS classifier framework. Recall that Taffler’s MDA model focuses on liquidity and makes use of four drivers; namely, Current Assets to Total Liabilities; Current Liabilities to Total Assets; Number of Credit Intervals; and Profit Before Tax to Current Liabilities. Note that lower values are better than higher ones for Current Liabilities to Total Assets and Number of Credit Intervals, whereas higher values of Current Assets to Total Liabilities and Profit Before Tax to Current Liabilities are better than lower ones.

Input: A training sample $X_E = \{x_{i,j}^E; i = 1, \dots, \#X_E, j = 1, \dots, m\}$ of cardinality $\#X_E$ and a test sample $X_T = \{x_{i,j}^T; i = 1, \dots, \#X_T, j = 1, \dots, m\}$ of cardinality $\#X_T$, where each observation i in X_E or X_T is an alternative (e.g., LSE listed firm-year observation) along with a set of relevant features (e.g., bankruptcy drivers) for the analysis under consideration (e.g., Current Assets to Total Liabilities; Current Liabilities to Total Assets; Number of Credit Intervals; Profit Before Tax to Current Liabilities) of cardinality m , and the observed risk or bankruptcy status Y ;

Phase 1: In-Sample Analysis

Step 1: Choose a *normalisation method* (see Table 1) along with a *weighting scheme* w (see Table 2) and use them to transform both training sample data $(x_{i,j}^E; i = 1, \dots, \#X_E, j = 1, \dots, m)$ and test sample data $(x_{i,j}^T; i = 1, \dots, \#X_T, j = 1, \dots, m)$ into their normalised counterparts $(r_{i,j}^E; i = 1, \dots, \#X_E, j = 1, \dots, m)$ and $(r_{i,j}^T; i = 1, \dots, \#X_T, j = 1, \dots, m)$, respectively, where $x_{i,j}^E$ (respectively $x_{i,j}^T$) denote the value of feature or driver j of alternative i in the training (respectively, test) sample and $r_{i,j}^E$ (respectively $r_{i,j}^T$) denote the standardized value of feature j of alternative i in the training (respectively, test) sample.

Step 2: Compute two virtual benchmarks r^+ and r^- – commonly referred to as the ideal positive solution and the ideal negative solution, respectively – as follows, where F^- (respectively, F^+) denote the set of features for which lower (respectively, higher) values are better:

$$r_j^+ = \min_{i=1, \dots, \#X_E} r_{i,j}^E, \text{ if } j \in F^- \text{ or } \max_{i=1, \dots, \#X_E} r_{i,j}^E, \text{ if } j \in F^+; j = 1, \dots, m,$$

and

$$r_j^- = \max_{i=1, \dots, \#X_E} r_{i,j}^E, \text{ if } j \in F^- \text{ or } \min_{i=1, \dots, \#X_E} r_{i,j}^E, \text{ if } j \in F^+; j = 1, \dots, m.$$

Step 3: Choose the *metric* d_{TOPSIS} to use for computing distances between alternatives – see Table 3 – and compute the distances $d_{TOPSIS}(r_i^E, r^+)$ and $d_{TOPSIS}(r_i^E, r^-)$ between each alternative i in the training sample X_E ($i = 1, \dots, \#X_E$) and the virtual benchmarks (i.e., ideal positive and negative solutions) r^+ and r^- , respectively, using the pre-specified metric.

Step 4: Choose the *type of similarity score* – which is appropriate for the application at hand – amongst the following:

$$s_i^+ = d_{TOPSIS}(r_i^E, r^+) / (d_{TOPSIS}(r_i^E, r^-) + d_{TOPSIS}(r_i^E, r^+));$$

$$s_i^- = d_{TOPSIS}(r_i^E, r^-) / (d_{TOPSIS}(r_i^E, r^-) + d_{TOPSIS}(r_i^E, r^+));$$

and, for each alternative i in the training sample X_E ($i = 1, \dots, \#X_E$), compute such score.

Step 5: Use the appropriate scores computed in the previous step to classify alternatives in the training sample X_E according to a *user-specified classification rule* into, for example, risk or bankruptcy classes, say \hat{Y}_E . Then, compare the TOPSIS based classification of alternatives in X_E into risk classes; that is, the predicted risk classes \hat{Y}_E , with the observed risk classes Y_E of alternatives in the training sample, and compute the relevant in-sample performance statistics. The choice of a decision rule for classification depends on the nature of the classification problem; that is, a two-class problem or a multi-class problem. In bankruptcy prediction we are concerned with a two-class problem; therefore, we shall provide a solution that is suitable for these problems. In fact, we propose a TOPSIS score-based cut-off point procedure to classify alternatives in X_E . The proposed procedure involves solving an optimization problem whereby the TOPSIS score-based cut-off point, say ρ , is determined so as to optimize a given classification performance measure, say π (e.g., Type I error, Type II error, Sensitivity, Specificity), over an interval with a lower bound, say ρ_{LB} , equal to the smallest TOPSIS score of alternatives in X_E (i.e., $\min_i r_i^+$ or $\min_i r_i^-$ depending on the decision made at step 4) and an upper bound, say ρ_{UB} , equal to the largest TOPSIS score of alternatives in X_E (i.e., $\max_i r_i^+$ or $\max_i r_i^-$ depending on the decision made at step 4). In sum, the proposed procedure is based on a performance measure-dependent approach. A generic procedure is summarised hereafter into three steps. Note that, in most applications, the performance measure π is a non-linear function. The choice of a specific optimization algorithm for the implementation of the generic procedure outlined above depends on whether the performance measure π is differentiable or not and if it is non-differentiable, whether it is quasi-convex or not. To be more specific, if π is differentiable, then one could choose Bisection Search; if π is twice differentiable, then one could choose Newton's Method; if π is non-differentiable but quasi-convex, then one could choose Golden Section Search, Fibonacci Search, Dichotomous Search, or a brute force search such as Uniform Search. For details on these standard non-linear programming algorithms, the reader is referred to Bazaraa et al. (2006).

Step 5a: Compute ρ_{LB} and ρ_{UB} ;

Step 5b: Find the optimal value of ρ with respect to π , say ρ^* , within the interval $[\rho_{LB}, \rho_{UB}]$ using the relevant non-linear programming search algorithm amongst the ones mentioned above;

Step 5c: Classify observations in X_E into two classes; namely bankrupt and non-bankrupt observations or firms; that is, determine the predicted risk classes \hat{Y}_E so that firms with TOPSIS similarity scores less (respectively, greater) than ρ^* are assigned to a bankruptcy class and those with TOPSIS scores greater (respectively, less) than or equal to ρ^* are assigned to a non-bankruptcy class if an ideal positive (respectively, an ideal negative) benchmark was chosen to compute TOPSIS scores;

Notice that the last step of this generic procedure classifies alternatives in the training sample into two classes; namely bankrupt and non-bankrupt firms or alternatives, and thus the output is the optimal TOPSIS score-based cut-off point ρ^* along with the predicted risk classes \hat{Y}_E . Compare the predicted risk classes \hat{Y}_T with the observed ones Y_T and compute the relevant in-sample performance statistics.

Phase 2: Out-of-Sample Analysis

Step 6: Use an appropriate *algorithm* to classify alternatives in X_T into, for example, risk or bankruptcy classes, say \hat{Y}_T . Then, compare the predicted risk classes \hat{Y}_T with the observed ones Y_T and compute the relevant out-of-sample performance statistics. Note that alternatives i in the test sample X_T could be classified using a decision rule similar to the one used for classifying alternatives in the training sample – see Step 5c, where ρ^* is the optimal cut-off score determined in Step 5b which is based on the training sample. This naïve classification rule might fail to predict the right class belonging for an alternative $i \in X_T$, because alternative i might score better (respectively, worse) than the ideal positive (respectively, negative) benchmark on one or several criteria; instead, we propose an instance of case-based reasoning; namely, the k-nearest neighbour (k-NN) algorithm which could be described as follows:

Initialization Step

Choose the Case Base as X_E and the Query Set as X_T ;

Choose a *distance metric* d_{k-NN} to use for computing distances between alternatives. In our implementation, we tested several choices amongst the following: Euclidean, Cityblock, and Mahalanobis;

Choose a *classification criterion*. In our implementation, we opted for the most commonly used one; that is, the majority vote;

Iterative Step

// Compute distances between queries and cases

FOR $i_1 = 1$ to $|X_T|$ {

FOR $i_2 = 1$ to $|X_E|$ {

 Compute $d_{k-NN}(Alternative_{i_1}, Alternative_{i_2})$; }

// Sort cases in ascending order of their distances to queries and classify queries

FOR $i_1 = 1$ to $|X_T|$ {

 Sort the list $L_{i_1} = \{(i_2, d_{k-NN}(Alternative_{i_1}, Alternative_{i_2})) ; i_2 = 1, \dots |X_E|\}$ in ascending order of distances and use the first k entries in the list $L_{i_1} (1:k, .)$ to classify $Alternative_{i_1}$ according to the chosen criterion; that is, the majority vote; }

Output: In-sample and out-of-sample classifications or risk class belongings of alternatives along with the corresponding performance statistics.

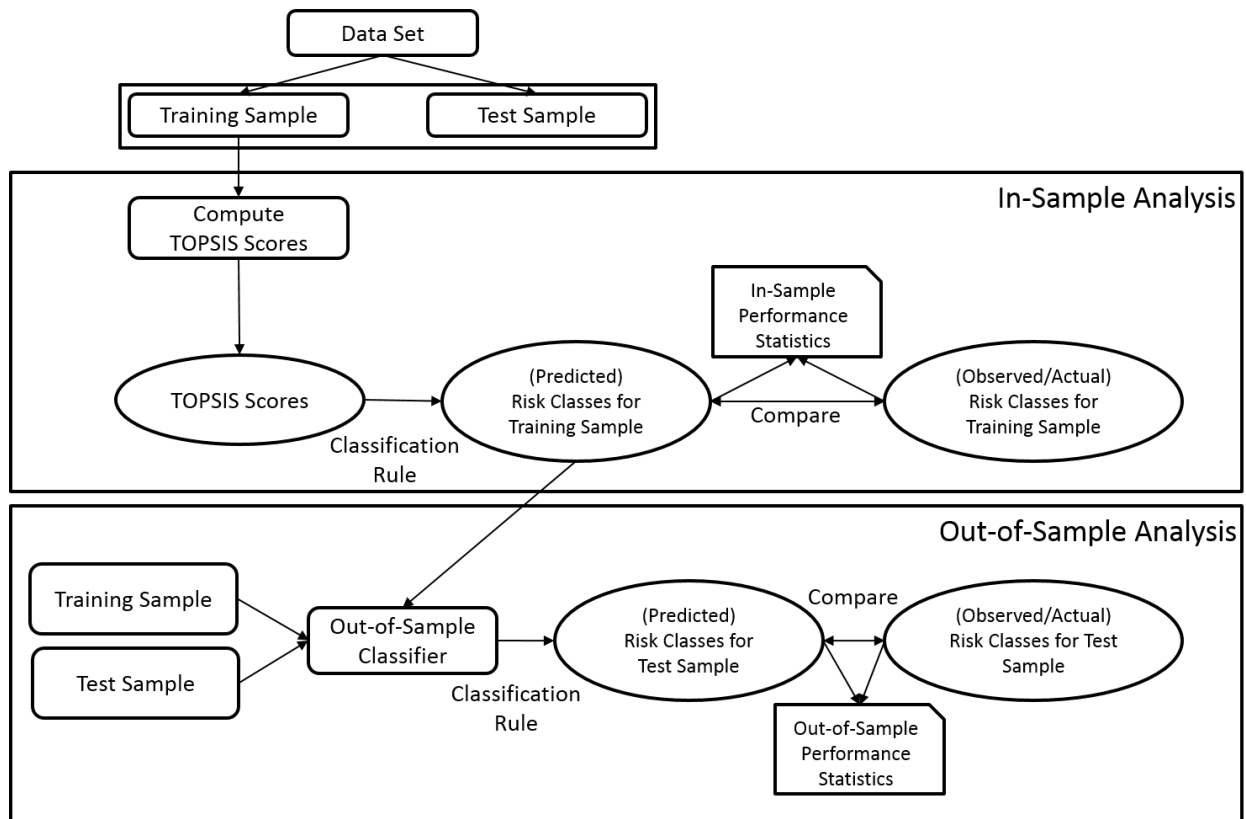


Figure 1: Generic Design of In-Sample and Out-of-Sample Analyses of TOPSIS Classifiers

Normalization Method	Formulae	Sample of References
Vector Normalization	$r_{i,j} = \frac{x_{i,j}}{\ x_j\ } = \frac{x_{i,j}}{\sqrt{\sum_{k=1}^{\#X} x_{k,j}^2}} \text{ If } j \in F^+$ $r_{i,j} = \frac{1/x_{i,j}}{\ 1/x_j\ } = \frac{1/x_{i,j}}{\sqrt{\sum_{k=1}^{\#X} (1/x_{k,j})^2}} \text{ If } j \in F^-$	Amiri et al. (2009), Isiklar and Büyüközkan (2007), Jahanshahloo et al. (2006), Li (2010), Peng et al. (2011), Secme et al. (2009), Shyur (2006), Wu et al. (2009), Wu et al. (2010), Yu et al. (2011), Zandi and Tavana (2011), Zhang et al. (2010)
Linear Scale Transformation (Max-Min)	$r_{i,j} = \frac{x_{i,j} - \min_{k=1, \dots, \#X} x_{k,j}}{\max_{k=1, \dots, \#X} x_{k,j} - \min_{k=1, \dots, \#X} x_{k,j}} \text{ If } j \in F^+$ $r_{i,j} = \frac{\max_{k=1, \dots, \#X} x_{k,j} - x_{i,j}}{\max_{k=1, \dots, \#X} x_{k,j} - \min_{k=1, \dots, \#X} x_{k,j}} \text{ If } j \in F^-$	Lin et al. (2010), Bai et al. (2014)
Linear Scale Transformation (Max)	$r_{i,j} = \frac{x_{i,j}}{\max_{k=1, \dots, \#X} x_{k,j}} \text{ If } j \in F^+$ $r_{i,j} = 1 - \frac{x_{i,j}}{\max_{k=1, \dots, \#X} x_{k,j}} \text{ If } j \in F^-$	Ertugrul and Karakasglu (2009), KarimiAzari et al. (2011), Sun (2010), Sun and Lin (2009), Vahdani et al. (2012)
Linear Scale Transformation (Sum)	$r_{i,j} = \frac{x_{i,j}}{\sum_{k=1}^{\#X} x_{k,j}} \text{ If } j \in F^+$ $r_{i,j} = \frac{1/x_{i,j}}{\sum_{k=1}^{\#X} (1/x_{k,j})} \text{ If } j \in F^-$	Huang and Peng (2012)

Table 1: Sample of Commonly Used Normalization Methods

Type of Weighting Process	Description
Subjective Assignment of Weights	Preferential Weights (Amiri et al., 2009; Benitez et al., 2007; Ertugrul and Karakasglu, 2009; Jahanshahloo et al., 2006; KarimiAzari et al., 2011; Li, 2010; Peng et al., 2011; Wang and Lee, 2007; Ye, 2010; Zandi and Tavana, 2011; Zhang et al., 2010); Analytical Hierarchy Process (AHP) based Methods (Aydogan, 2011; Isiklar and Büyüközkan, 2007; Khademi-Zare et al., 2010; Lin et al., 2010; Secme et al., 2009; Shyur, 2006; Sun, 2010; Sun and Lin, 2009; Vahdani et al., 2012; Wu et al., 2009; Wu et al., 2010; Yu et al., 2011)
Objective / Data-driven Assignment of Weights	Equal weights (Chang et al., 2010); Centroid weights; Entropy weight method (Chang et al., 2010); Coefficient of Variation weight method (Chang et al., 2010); Weights obtained from Regression Techniques (Olson, 2004; Wu and Olson, 2006); Data Envelopment Analysis (Chen et al., 2009)

Table 2: Sample of Commonly Used Weighting Schema

Distance	Formulae	Sample References
Euclidean distance (Minkowski with p=2)	$d_{i,k} = \sqrt{\sum_{j=1}^m (x_{i,j} - x_{k,j})^2}$	Amiri et al. (2009), Aydogan (2011), Benitez et al. (2007), Chang et al. (2010), Ertugrul and Karakasglu (2009), Huang and Peng (2012), Isiklar and Büyüközkan (2007), Jahanshahloo et al. (2006), KarimiAzari et al. (2011), Li (2010), Lin et al. (2010), Peng et al. (2011), Secme et al. (2009), Vahdani et al. (2012), Wu et al. (2009), Wu et al. (2010), Ye (2010), Yu et al. (2011), Zandi and Tavana (2011), Zhang et al. (2010), Vega et al. (2014)
Manhattan/Cityblock distance (Minkowski with p=1)	$d_{i,k} = \sum_{j=1}^m x_{i,j} - x_{k,j} $	Chang et al. (2010), Khademi-Zare et al. (2010), Tan (2011), Vega et al. (2014)
Mahalanobis distance	$d_{i,k} = \sqrt{(x_i - x_k)^t \Sigma^{-1} (x_i - x_k)}$ $\Sigma = \text{variance-covariance matrix}$	Chang et al. (2010), Vega et al. (2014)

Table 3: Sample of Commonly used Distance Metrics

In the next section, we shall report on our empirical evaluation of the proposed framework.

3. Empirical Results

In order to assess the performance of the proposed framework, we considered a sample of 6605 firm-year observations consisting of non-bankrupt and bankrupt UK firms listed on the London Stock Exchange (LSE) during 2010-2014 excluding financial firms and utilities as well as those firms with less than 5 months lag between the reporting date and the fiscal year. The source of our sample is DataStream. The list of bankrupt firms is however compiled from London Share Price Database (LSPD) – codes 16 (Receivership), 20 (in Administration) and 21 (Cancelled and Assumed valueless). Information on our dataset composition is summarised in Table 4. As to the selection of the training sample and the test sample, we have chosen the size of the training sample to be twice the size of the test sample. The selection of observations was done with random sampling without replacement so as to ensure that both the training sample and the test sample have the same proportions of bankrupt and non-bankrupt firms. A total of thirty pairs of training sample-test sample were generated.

Observations (2010-2014)	Nb.	%
Bankrupt Firm-Year Observations	407	6.16%
Non-Bankrupt Firm-Year Observations	6198	94.38%
Total Firm-Year Observations	6605	100%

Table 4: Dataset Composition

There are many parametric bankruptcy prediction models (e.g., Altman, 1968; Ohlson, 1980; Zmijewski, 1984; Taffler, 1984). In our experiment, we reworked a standard and well known parametric model in the TOPSIS framework; namely, the multivariate discriminant analysis (MDA) model of Taffler (1984), to provide some empirical evidence on the merit of the proposed framework. Recall that Taffler’s model makes use of four explanatory variables: current liabilities to total assets, number of credit intervals, profit before tax to current liabilities, and current assets to total liabilities. We report on the performance of the proposed framework using four commonly used metrics; namely, Type I error (T1), Type II error (T2), Sensitivity (Sen) and Specificity (Spe), where T1 is the proportion of bankrupt firms predicted as non-bankrupt, T2 is the proportion of non-bankrupt firms predicted as bankrupt, Sen is the proportion of non-bankrupt firms predicted as non-bankrupt, and Spe is the proportion of bankrupt firms predicted as bankrupt.

Since both the TOPSIS classifier and the k-NN classifier, trained on the classification done with TOPSIS, require a number of decisions to be made for their implementation, we considered several combinations of decisions to find out about the extent to which the performance of the proposed framework is sensitive or robust to these decisions. Recall that, for the TOPSIS classifier, the analyst has to choose (1) the normalization method, (2) the weighting scheme, (3) the metric to use for computing distances between each alternative and the virtual benchmarks, d_{TOPSIS} , (4) the type of similarity score to use, and (5) the classification rule. On the other hand, for the k-NN classifier, the analyst has to choose (1) the metric to use for computing distances between alternatives, d_{k-NN} , (2) the classification criterion, and (3) the size of the neighbourhood k . Our choices for these decisions are summarised in Table 5.

TOPSIS	
Decision	Options Considered and Justification, if relevant
Normalization Method	Vector Normalization; Linear Scale Transformation (Sum); Linear Scale Transformation (Max); and Linear Scale Transformation (Max-Min).
Weighting Scheme	Equal weights. In practice, any of the weighting schema outlined in Table 2 could be used; however, for testing the performance of our proposed framework, we have chosen to avoid any subjectivity due to user (authors) preferences. On other hand, we avoided using objective or data-driven weighting schema because they all rely on assumptions of the underlying methodologies which might be justified or hold in one application context but not in a different context.
Metric d_{TOPSIS}	Euclidean distance; Manhattan/Cityblock distance; and Mahalanobis distance. Note that, for our dataset, the performance of the proposed framework turned out not to be sensitive to the choice of the metric to measure the distance between observations. Therefore, we only report results for the commonly used metric; namely, Euclidean distance.
Type of Similarity Score	s^+ and s^- . Note that, for our dataset, the performance of the proposed framework turned out not to be sensitive to s^+ and s^- . Therefore, we only report results for s^+ .
Classification Rule	TOPSIS score-based cut-off point procedure, where the choice of the cut-off point optimises a specific performance measure (i.e., T1, T2, Sen, Spe)
k-NN	
Decision	Options Considered and Justification, if relevant
Metric d_{k-NN}	Euclidean, Cityblock, Mahalanobis.
Classification Criterion	Majority vote. Several criteria could have been used such as a Weighted Vote, but once again our choice is made so as to avoid any personal (subjective) preferences.
Size of the neighbourhood k	$k = 3; 5; 7$. The results reported are for $k = 3$ since higher values delivered very close performances but required more computations.

Table 5: Implementation Decisions for TOPSIS and k-NN

Hereafter, we shall provide a summary of our empirical results and findings. Table 6 provides a summary of in-sample and out-of-sample statistics on the performance of the MDA model of Taffler (1984) reworked within our proposed framework, which is an integrated in-sample – out-of-sample framework for TOPSIS classifiers. In sum, the proposed framework is meant to equip TOPSIS with a mechanism to perform out-of-sample prediction where an instance of case-based reasoning; namely, k-NN, is trained on the outcome or in-sample classification of TOPSIS.

With respect to in-sample performance of the proposed TOPSIS classifier, our results demonstrate that TOPSIS provides an outstanding classifier regardless of the choices of its implementation decisions – see Table 6. In fact, in-sample, our TOPSIS classifier does not wrongly classify any non-bankrupt firm as demonstrated by Type II error of 0% and Sensitivity of 100%.

In-sample Performance						
Normalization Method	Distance Metric	Statistics	Performance Measure			
			T1	T2	Sen.	Spe.
Vector normalization	Euclidean	Min	0%	0%	100%	100%
		Max	0%	0%	100%	100%
		Average	0%	0%	100%	100%
		Std. Dev.	0%	0%	0%	0%
Linear Scale Transformation Max-Min	Euclidean	Min	0.3690%	0%	100%	99.6310%
		Max	0.3690%	0%	100%	99.6310%
		Average	0.3690%	0%	100%	99.6310%
		Std. Dev.	0%	0%	0%	0%
Linear Scale Transformation Sum	Euclidean	Min	0%	0%	100%	100%
		Max	0%	0%	100%	100%
		Average	0%	0%	100%	100%
		Std. Dev.	0%	0%	0%	0%
Linear Scale Transformation Max	Euclidean	Min	0.3690%	0%	100%	99.6310%
		Max	0.3690%	0%	100%	99.6310%
		Average	0.3690%	0%	100%	99.6310%
		Std. Dev.	0%	0%	0%	0%
Out-of-Sample Performance						
Normalization Method	Distance Metric	Statistics	Performance Measure			
			T1	T2	Sen.	Spe.
Vector normalization	Euclidean	Min	0%	0%	100%	100%
		Max	0%	0%	100%	100%
		Average	0%	0%	100%	100%
		Std. Dev.	0%	0%	0%	0%
Linear Scale Transformation Max-Min	Euclidean	Min	0.7353%	0%	100%	99.2647%
		Max	0.7353%	0%	100%	99.2647%
		Average	0.7353%	0%	100%	99.2647%
		Std. Dev.	0%	0%	0%	0%
Linear Scale Transformation Sum	Euclidean	Min	0.7353%	0%	100%	99.2647%
		Max	0.7353%	0%	100%	99.2647%
		Average	0.7353%	0%	100%	99.2647%
		Std. Dev.	0%	0%	0%	0%
Linear Scale Transformation Max	Euclidean	Min	0%	0%	100%	100%
		Max	0%	0%	100%	100%
		Average	0%	0%	100%	100%
		Std. Dev.	0%	0%	0%	0%

Table 6: Summary Statistics of The Performance of The Proposed Framework

On the other hand, most bankrupt firms are properly classified as demonstrated by a very small range (0% to 0.3690%) of Type I error, and a very high Specificity ranging from 99.6310% to

100%. Notice that the in-sample performance is slightly effected by the choice of the normalisation method. Note that compared to Multivariate Discriminant Analysis (MDA), the performance of TOPSIS is by far superior – see Table 7.

Next, we provide empirical evidence to demonstrate that the proposed out-of-sample framework achieved a very high performance in classifying firms listed on LSE into the right bankruptcy category – see Table 6. In fact, regardless of which TOPSIS and k-NN implementation decisions are made, the out-of-sample performance of the proposed framework is outstanding. In fact, ideal results are obtained with Vector normalization and Linear Scale Transformation (Max); that is, T1 and T2 being 0% and sensitivity and specificity being 100%. These performances are slightly lower when data is normalized using the Linear Scale Transformation (Max-Min) or Linear Scale Transformation (Sum) with T1 being 0.7353% and specificity being 99.2647%. Notice that the classification of non-bankrupt firms has not been affected by the change in the normalization method. Out-of-sample, the proposed framework also proves to be superior to MDA – see Table 7.

In-sample Performance				
Statistics	Performance Measure			
	T1	T2	Sen.	Spe.
Min	97.0500%	0.1900%	99.3700%	0%
Max	100%	0.6300%	99.8100%	2.9500%
Average	98.8200%	0.2600%	99.7400%	1.1800%
Std. Dev.	0.6700%	0.0900%	0.0900%	0.6700%
Out-of-Sample Performance				
Statistics	Performance Measure			
	T1	T2	Sen.	Spe.
Min	0%	0%	0.1500%	0%
Max	100%	99.8500%	100%	100%
Average	82.2000%	17.0100%	82.9900%	17.8000%
Std. Dev.	37.4300%	37.6600%	37.6600%	37.4300%

Table 7: Summary Statistics of The Performance of MDA

4. Conclusions

The validation of prediction models requires both in-sample and out-of-sample evaluation of their performance. TOPSIS classifiers however lack a proper framework for performing their out-of-sample evaluation. In this paper, we filled this gap by proposing an instance of the case-based reasoning methodology; namely, k-nearest neighbour, trained on the outcome of a TOPSIS

classifier. We assessed the performance of the proposed framework using a UK dataset on bankrupt and non-bankrupt firms. Our results demonstrate its outstanding prediction performance. In addition, the outcome of the proposed framework is robust to a variety of implementation decisions. Last, but not least, the proposed out-of-sample framework makes TOPSIS classifiers real contenders for practitioners.

References

- Abd-El Fattah, I.M., Khedr, W.I., & Sallam, K.M. (2013). A TOPSIS based Method for Gene Selection for Cancer Classification. *International Journal of Computer Applications*, 67(17), 39–44.
- Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance*, 23, 589-609.
- Amiri, M., Zandieh, M., Soltani, R., & Vahdani, B. (2009). A hybrid multi-criteria decision-making model for firms competence evaluation. *Expert Systems with Applications*, 36, 12314–12322.
- Aydogan, E.K. (2011). Performance measurement model for Turkish aviation firms using the rough-AHP and TOPSIS methods under fuzzy environment. *Expert Systems with Applications*, 38, 3992–3998.
- Bai, C., Dhavale, D., & Sarkis, J. (2014). Integrating Fuzzy C-Means and TOPSIS for performance evaluation: An application and comparative analysis. *Expert Systems with Applications*, 41, 4186–4196.
- Bazaraa, M.S., Sherali, H.D., & Shetty, C.M. (2006). *Nonlinear Programming: Theory and Algorithms*. (3rd edition), New Jersey: John Wiley & Sons Inc.
- Behzadian, M., Otaghsara, S.K., Yazdani, M., & Ignatius, J. (2012). A state-of-the-art survey of TOPSIS applications. *Expert Systems and Applications*, 39, 13051–13069.
- Benitez, J.M., Martin, J.C., & Roman, C. (2007). Using fuzzy number for measuring quality of service in the hotel industry. *Tourism Management*, 28, 544–555.
- Chang, C.H., Lin, J.J., Lin, J.H., & Chiang, M.C. (2010). Domestic open-end equity mutual fund performance evaluation using extended TOPSIS method with different distance approaches. *Expert Systems with Applications*, 37, 4642–4649.
- Chen, Y., Li, K.W., Xu, H., & Liu, S. (2009). A DEA-TOPSIS method for multiple criteria decision analysis in emergency management. *Journal of Systems Science and Systems Engineering*, 18(4), 489–507.
- Ertugrul, I., & Karakasglu, N. (2008). Comparison of fuzzy AHP and fuzzy TOPSIS methods for facility location selection. *International Journal of Advanced Manufacturing Technology*, 39, 783–795.

- Hwang, C.L., & Yoon, K. (1981). *Multiple Attribute Decision Making*, Springer-Verlag, Berlin.
- Huang, J.H., & Peng, K.H. (2011). Fuzzy Rasch model in TOPSIS: A new approach for generating fuzzy numbers to assess the competitiveness of the tourism industries in Asian countries. *Tourism Management*, 33(2), 456–465.
- Isiklar, G., & Büyüközkan, G. (2007). Using a multi-criteria decision making approach to evaluate mobile phone alternatives. *Computer Standards & Interfaces*, 29, 265–274.
- Jahanshahloo, G.R., Lotfi, F.H., & Izadikhah, M. (2006). An algorithmic method to extend TOPSIS for decision-making problems with interval data. *Applied Mathematics and Computation*, 175, 1375–1384.
- KarimiAzari, A.R., Mousavi, N., Mousavi, S.F., & Hosseini, S.B. (2011). Risk assessment model selection in construction industry. *Expert Systems with Applications*, 38, 9105–9111.
- Khademi-Zare, H., Zarei, M., Sadeghieh, A., & Saleh Owlia, M. (2010). Ranking the strategic actions of Iran mobile cellular telecommunication using two models of fuzzy QFD. *Telecommunications Policy*, 34, 747–759.
- Li, D.F. (2010). TOPSIS-based nonlinear-programming methodology for multiattributes decision making with interval-valued intuitionistic fuzzy sets. *IEEE Transactions on Fuzzy Systems*, 18(2), 299–311.
- Lin, C.L., Hsieh, M.S., & Tzeng, G.H. (2010). Evaluating vehicle telematics system by using a novel MCDM technique with dependence and feedback. *Expert Systems with Applications*, 37, 6723–6736.
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18, 109-131.
- Olson, D.L. (2004). Comparison of Weights in TOPSIS Models. *Mathematical and Computer Modelling*, 40, 721–727.
- Peng, Y., Wang, G., Kou, G., & Shi, Y. (2011). An empirical study of classification algorithm evaluation for financial risk prediction. *Applied Soft Computing*, 11, 2906–2915.
- Secme, N.Y., Bayrakdaroglu, A., & Kahraman, C. (2009). Fuzzy performance evaluation in Turkish banking sector using analytic hierarchy process and TOPSIS. *Expert Systems with Applications*, 36, 11699–11709.
- Shyur, H.J. (2006). COTS evaluation using modified TOPSIS and ANP. *Applied Mathematics and Computation*, 177, 251–259.
- Sun, C.C., & Lin, G.T.R. (2009). Using fuzzy TOPSIS method for evaluating the competitive advantages of shopping websites. *Expert Systems with Applications*, 36, 11764–11771.
- Sun, C.C. (2010). A performance evaluation model by integrating fuzzy AHP and fuzzy TOPSIS methods. *Expert Systems with Applications*, 37, 7745–7754.

- Taffler, R.J. (1984). Empirical models for the monitoring of UK corporations. *Journal of Banking and Finance*, 8(2), 199–227.
- Tan, C. (2011). A multi-criteria interval-valued intuitionistic fuzzy group decision making with Choquet integral-based TOPSIS. *Expert Systems with Applications*, 38, 3023–3033.
- Tansel IÇ Y., & Yurdakul M. (2010). Development of a quick credibility scoring decision support system using fuzzy TOPSIS. *Expert Systems with Applications*, 37, 567–574.
- Vahdani, B., Hadipour, H., & Tavakkoli-Moghaddam, R. (2012). Soft computing based on interval valued fuzzy ANP – a novel methodology. *Journal of Intelligent Manufacturing*, 23(5), 1529–1544.
- Vega, A., Aguarón, J., García-Alcaraz, J., & Moreno-Jiménez, J.M. (2014). Notes on dependent attributes in TOPSIS. *Procedia Computer Science*, 31, 308–317.
- Wang, Y.J., & Lee, H.S. (2007). Generalizing TOPSIS for fuzzy multiple-criteria group decision-making. *Computers and Mathematics with Applications*, 53, 1762–1772.
- Wu, D., & Olson, D.L. (2006). A TOPSIS Data Mining Demonstration and Application to Credit Scoring. *International Journal of Data Warehousing and Mining*, 2(3), 16–26.
- Wu, C.R., Lin, C.T., & Lin, Y.F. (2009). Selecting the preferable bank assurance strategic alliance by using expert group decision technique. *Expert Systems with Applications*, 36, 3623–3629.
- Wu, C.S., Lin, C.T., & Lee, C. (2010). Optimal marketing strategy: A decision making with ANP and TOPSIS. *International Journal of Production Economics*, 127, 190–196.
- Ye, F. (2010). An extended TOPSIS method with interval-valued intuitionistic fuzzy numbers for virtual enterprise partner selection. *Expert Systems with Applications*, 37, 7050–7055.
- Yu, X., Guo, S., Guo, J., & Huang, X. (2011). Rank B2C e-commerce websites in e-alliance based on AHP and fuzzy TOPSIS. *Expert Systems with Applications*, 38, 3550–3557.
- Zandi, F., & Tavana, M. (2011). A fuzzy group quality function deployment model for e-CRM framework assessment in agile manufacturing. *Computers & Industrial Engineering*, 61, 1–19.
- Zhang, L., Gao, L., Shao, X., Wen, L., & Zhi, J. (2010). A PSO-Fuzzy group decision making support system in vehicle performance evaluation. *Mathematical and Computer Modeling*, 52, 1921–1931.
- Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59-82.