An emerging online business decision making architecture in a dynamic economic environment

Sin-Jin Lin\textsuperscript{a}, Te-Min Chang\textsuperscript{b} and Ming-Fu Hsu\textsuperscript{c,*}

\textsuperscript{a}Department of Accounting, Chinese Culture University, Yang-Ming-Shan, Taipei, Taiwan, R.O.C
\textsuperscript{b}Department of Information Management, National Sun Yat-Sen University, Kaohsiung, Taiwan, R.O.C
\textsuperscript{c}English Program of Global Business, Chinese Culture University, Yang-Ming-Shan, Taipei, Taiwan, R.O.C

Abstract. It is widely acknowledged that a corporate’s profitability that decreases dramatically not only threatens both potential and current investors, but also can freeze stock market transactions as well as deteriorate the flow of economic resources. However, far too little attention has been paid to this issue, which also has been deemed as a main trigger for a financial crisis. To confront this problem, a decision support system can be built up to evaluate a corporate’s operating performance. Thus, this study introduces a novel architecture for forecasting the online operating performance of a firm when entering data at different time intervals. The introduced architecture is grounded on multiple data envelopment analysis specifications, dynamic fuzzy c-means (DFCM), and extreme support vector machine (ESVM). Because obtaining a comprehensible architecture is essential for achieving high accuracy in today’s knowledge-based economy, this study advances the opaque nature of the introduced architecture and extracts the inherent knowledge to represent it in a transparent, human-readable format. The decision logics can be judged or examined by users, which will increase the acceptance rate of the architecture as well as enhance its practical application. Our built-up architecture, tested by real cases, is a promising alternative for financial performance forecasting.

Keywords: Decision making, fuzzy c-means, knowledge management, artificial intelligence, data envelopment analysis

1. Introduction

Modelling corporate survival and comprehending the main triggers for determining a corporate’s existence have drawn considerable attention from both academics and practitioners over the last two decades, because financial distress precedes financial troubles, such as liquidity risk, credit risk, financial default, or insolvency risk \cite{15, 36}. Tinco and Wilson \cite{43} and Chang et al. \cite{5} also stated that financial distress can impede the development of socioeconomics, obstruct the circulation of economic resources, raise the unemployment rate, and cause fatal damage to both potential and current investors. If a pre-warning model of financial distress is trustful and reliable, then top-level managers can initiate remedial treatment to prevent the situation from getting worse, and stakeholders can subsequently adjust their investment portfolio so as to maximize their personnel wealth under anticipated risk exposure. Thus, how to establish an accurate forecasting model for managers to form decisions is an urgent and important aspect in finance, especially as today’s global economic growth has begun to slow down \cite{17, 30}.

Decisions on the financial risk of a corporate borrower have traditionally relied exclusively on sub-
jective judgments made by human experts adopting past experiences and some guiding principles [8, 42]. However, two critical challenges with the aforementioned method are the difficulty to achieve consensus estimates and the fact that decisions tend to be reactive rather than predictive. Financial ratio analysis has been widely viewed as a popular managerial tool to determine the economic activity of corporate and has also gained acceptance for describing a corporate’s operating status and for providing insights and clarifications when making a variety of business judgments. In this vein, such analysis can be executed to reduce uncertainty in decision making via providing a trustworthy measurement of the planning, investing, operating effectiveness, and health of business operations [40].

The literature has widely employed a mixture of ratio analysis and statistical or artificial intelligence (AI) techniques to construct financial distress prediction models. Examples include Altman [1] who applied discriminant analysis (DA) to establish a prediction model that can determine which corporates belong to the category of bankruptcy and which corporates belong to the category of non-bankruptcy; Ohlson [33] utilized logistic regression (LR) to forecast a corporate’s financial distress; Wilson and Sharda [47] set up a neural network (NN) to forecast the discontinuity of a corporate; Jo et al. [18] executed case-based reasoning (CBR) to evaluate the health of a corporate; and Kim and Sonh [22] used support vector machine (SVM) for default prediction. Unfortunately, compared to the large strand of studies on financial distress prediction and credit risk prediction, research works on corporate operating performance forecasting are quite rare, even as it is deemed to be the main trigger for a financial crisis. Kamei [21] further indicated that up to 99% of financial distress arises due to poor management. How to robustly evaluate corporate operating performance thus turns out to be an essential issue.

Return on assets (ROA), return on investment (ROI), and return on equity (ROE) are widely designated as performance measures and belong to the categories of one input and one output variable. However, merely utilizing one input and one output variable to depict the whole structure of a corporate’s operating status is not reliable and trustful in today’s highly turmoil economic environment. In other words, these measures have not yet been able to fully explain the relationships between previous and future financial performances [5, 14]. Recent efforts have shown that this fact may be caused by omitting essential information that can be substantially extracted from financial reports and other text documents [12]. To overcome such an obstacle this study takes up a mathematical liner programming algorithm as the performance measure - namely, data envelopment analysis (DEA), which not only can provide an overall measure of performance without pre-setting an objective function (i.e., cost minimization and profit maximization), but also can simultaneously be open to multiple input and output variables.

DEA provides a score that relies heavily on pre-determined input and output variables. The score is affected by the inclusion or exclusion of an input or an output [34]. For making the efficiency judgment more robust and compact, we prefer to go beyond a score (i.e., a single DEA specification) and advance an alternative DEA specification - i.e., combining inputs and outputs in several different ways - that can achieve two benefits at the same time. On one hand, this method is effective for examining the robustness and effectiveness of the result. On the other hand, a set of numerous efficiency scores yields a wide range of information for clustering and classifying the observations [41]. This study takes two input and three output variables into consideration that can generate 21 different DEA specifications. However, too much information could confuse decision makers and mislead the decision outcome. In order to make the analyzed results accessible to non-specialists, we perform a clustering technique.

K-means and fuzzy c means (FCM) are two well-known methods of data clustering. Compared to K-means, FCM is more comprehensive and accurate [45]. When the new data are heterogeneous or different from the existing data, then it is likely that all cluster sets including the data of each cluster and even the number of clusters may be changed [37]. In the above case, FCM lacks handling capability, because it cannot update the clusters. To solve this problem, a dynamic clustering technique (i.e., dynamic fuzzy c means: DFCM) should be carried out. By doing so, managers can analyze the changes/movements within clusters (i.e., from inefficiency cluster to efficiency cluster) so as to know the effectiveness of operating strategies and quickly respond to market changes.

Analized results derived from multiple DEA specifications and DFCM can be fed into an emerging artificial intelligence (AI) technique - namely, extreme support vector machine (ESVM), which is
developed from extreme learning machine (ELM) and support vector machine (SVM) - to construct the forecasting model [16, 28]. It not only preserves the benefit of ELM, such as extremely fast learning speed, but also has better generalization capability than conventional ELM due to its output bias term and regularization scheme [13, 49]. One of the critical challenges of ESVM is a lack of comprehensibility. In other words, the inherent decision logic of ESVM is obscure (i.e., black-box). Cano et al. [4] also stated that the comprehensibility of the model is a new challenge as essential as is its accuracy. If the decision logics embedded into the model cannot be examined or judged by human experts, then they have a higher chance of seeing a declining rate of model acceptance as well as confronting an impedance to its practical applications [7, 47, 48]. To overcome this challenge, we execute a rule generation algorithm that maximizes interpretability with more compact and a fewer number of rules in order to open up the black-box of ESVM as well as to represent the decision logics in a human-readable and transparent format. Managers can take the rules as a handbook to allocate their own economic resources to the right place and maximize their wealth under anticipated risk levels. The introduced model, examined by real cases, is a promising alternative for operating performance forecasting.

The reminder of this paper is organized as follows. Section 2 briefly outlines the proposed methodologies. Section 3 explains the research design. The paper ends with conclusions in Section 4.

2. Methodologies

2.1. Data envelopment analysis: DEA

Data envelopment analysis (DEA) introduced by Charnes et al. [6] is a data-oriented algorithm utilized in assessing the relative performance of a set of homogeneous entities called decision-making units (DMUs) by means of setting multiple input and output variables. Since it does not require a priori assumptions and has other merits, it has rapidly caught many researchers’ attentions [19].

Suppose we have m DMUs. Each DMU\(_j\)\((j = 1, \ldots, m)\) generates h outputs \(y_{rj}(r = 1, \ldots, h)\) by utilizing \(n\) inputs \(x_{ij}(i = 1, \ldots, n)\). Each DMU’s efficiency score can be derived from Equation (1).

\[
\max \frac{\sum_{r=1}^{h} u_{r} y_{r0}}{\sum_{i=1}^{n} g_{i} x_{i0}} \\
\text{s.t.} \frac{\sum_{r=1}^{h} u_{r} y_{r0}}{\sum_{i=1}^{n} g_{i} x_{i0}} \leq 1, \ j = 1, \ldots, m.
\]

\[
u_{r}, g_{i} \geq \varepsilon, \ i = 1, \ldots, n; \ r = 1, \ldots, h.
\]

2.2. Dynamic fuzzy c-means: DFCM

Clustering is the procedure of grouping s series of data from an unlabeled pattern into a number of groups called clusters, so that a similar pattern is grouped into one cluster [5, 11]. In real applications, the data are not easy to be classified into crisp clusters. To handle the aforementioned task, fuzzy clustering (i.e., fuzzy c means, FCM), which allocates data into different clusters via implementing fuzzy logics to determine the effective means for discovering overlapping clusters, was introduced by Dunn [10] and improved by Bezdek [3]. The objective function of FCM is to minimize the following equation [29, 38]:

\[
F(U, Z, \lambda, X) = \sum_{i=1}^{k} \sum_{j=1}^{n} (u_{ij})^{\lambda} \| x_{j} - z_{i} \|^{2}
\]

where \(\lambda\) denotes a fuzzy factor; \(k\) expresses the number of clusters; \(Z = (z_{1}, z_{2}, \ldots, z_{k})^{T}\) represents the vector of cluster centers containing the centers of the \(k\) clusters; \(n\) depicts the number of data-points; \(X = (x_{1}, x_{2}, \ldots, x_{n})^{T}\) indicates the datapoint vector; and \(U = [u_{ij}]_{k \times n}\) denotes the membership matrix consisting of membership \(u_{ij}\), which represents the membership of \(x_{j}\) in the \(i-th\) cluster, and the Euclidean distance norm expressed as \(\| \cdot \|\).

We normalize and fuzzify the membership by \(\lambda\). Iterative approaches such as an alternative optimization algorithm can be used to solve the task of minimization of \(F(U, Z, \lambda, X)\). When \(\lambda > 1\), the solution that minimizes \(F(U, Z, \lambda, X)\) is shown in Equation (3).

\[
u_{ij} = \left( \sum_{q=1}^{k} \left( \frac{\| x_{j} - z_{q} \|}{\| x_{j} - z_{i} \|} \right)^{2/(\lambda-1)} \right)^{-1}
\]
where \(1 \leq i \leq k, 1 \leq j \leq n\), and the center of the \(i-th\) cluster is appears in Equation (4).

\[
Z_i = \frac{\sum_{j=1}^{n} (u_{ij})^2 x_j}{\sum_{j=1}^{n} (u_{ij})} \tag{4}
\]

Although FCM performs well in the clustering task, it loses its capability when it has to update the clusters. In order to add the dynamic capability into FCM, two new parameters, i.e., membership threshold \((Y_{th})\) and FCM error \((EF_{CM})\), should be decided. The former is represented as the maximum acceptable level for the degree of memberships; the latter is expressed as the maximum acceptable difference between two centers of the clusters acquired in two successive steps utilizing FCM \([37, 39]\). We employ the Xie-Beni index as a validity index to determine how to better represent the data in obtained clusters. The mathematical formulation is shown in Equation (5).

\[
I \cdot Z_{XB}(U, Z; X) = \sum_{i=1}^{k} \left( \sum_{j=1}^{n} (u_{ij})^2 \left\| x_j - z_i \right\|^2 \right) / k \cdot (\text{Min}_{i \neq j} \left\{ \left\| z_i - z_j \right\| \right\}) \tag{5}
\]

If the newly entered data’s maximum degree of membership is greater than or equal to the membership threshold, then it means that the data belong to one of the clusters. If the degree is less than threshold, then it must be tested to see whether or not there is a better decision for the number of clusters (i.e., \(k\)). This is determined by comparing with \(k - 1\) and \(k + 1\) centers of the existing clusters. The assessing criteria is based on the Xie-Beni index. A more detailed illustration of DFCM can be seen in Fathabadi \([11]\) and Rezaee et al. \([37]\).

### 2.3. Extreme support vector machine: ESVM

Extreme support vector machine (ESVM) is a sort of a single hidden layer feedforward neural network (SLFN) \([27]\). The input \(x \in \mathbb{R}^J\) can be mapped onto a future space via an activation function \(\Theta(G', x)\), in which the hidden layer parameters \(G'\) from the input layer can be decided randomly. Sequentially, the output layer deals with a regularized least square task in the future space, where the regularizations are put on both \(g\) and \(h\). The purpose of ESVM is to identify an approximate decision surface of SVM, \(g^T x - h = \{\pm 1\}\), where \(g\) and \(h\) represent the orientation and relative location of the decision surface, respectively. ESVM can be obtained by substituting the equality constraint for the inequality constraint of SVM, and the formulation is shown in Equation (6).

\[
\min_{(g,h,\zeta) \in \mathbb{R}^{J+1}} \frac{C}{2} \left\| \zeta \right\|_2^2 + \frac{1}{2} \left\| g \right\|_2^2 \tag{6}
\]

\[
s.t. \quad B(\Theta(G', D)g - he) + \zeta = 0
\]

where \(C\) denotes the regularization parameter, \(\zeta\) represents the slack variable, and \(B\) depicts the diagonal square matrix. In Equation (6), \(D\) is the data sample matrix of size \(n \times m\), where each row is one sample \(x\), and \(\Theta(G', D) = (\Theta(G'x_1), \ldots, \Theta(G'x_n))^T\). The feature mapping function in the hidden layer is denoted as \(\Theta(G', x) : \mathbb{R}^f \rightarrow \mathbb{R}^\tilde{f}\), where \(f\) and \(\tilde{f}\) represent the dimension of the input data and the number of hidden nodes, respectively. The sigmoid kernel \((\Theta(G', x) = \text{sigmoid}(G^T x))\) is taken as a mapping function of ESVM, where \(G'\) is a matrix of size \(f \times \tilde{f}\) that can be randomly decided. The model is a quadratic programming optimization problem.

Liu et al. \([28]\) indicated that the solution to the aforementioned problem is equivalent to handling the following Equation (7).

\[
\begin{bmatrix}
    g \\
    h
\end{bmatrix} = \left( \frac{I}{C} + K_\Theta^T K_\Theta \right)^{-1} K_\Theta^T B \tag{7}
\]

where \(K_\Theta = \left[ \Theta(G', D), -e \right] \in \mathbb{R}^{n \times (J+1)}\). Zhu et al. \([49]\) indicated that ESVM’s generalization capability outperforms ELM in almost all classification tasks and reaches almost comparable accuracies to SVM. With the simple calculation procedure of ESVM, it can also reach an extremely fast learning speed without deteriorating its forecasting quality. Thus, this study uses ESVM.

### 3. Research design

#### 3.1. The data

Taiwan has been admired as one of the economic miracles of the world, especially its electronics industry, which emerged in the later part of the 20\textsuperscript{th} century and is now an integral part in the global supply chain
of electronic devices (such as personal computers and integrated circuits) as well as an essential capital market for global investors. Over the years, the Taiwan government has announced numerous financial policies or incentives, such as tax preferences/credits, land, human resources, capital, training courses, etc., for this specific industry, which has turned into the backbone of the local stock market. In particular, stock transactions covering the electronics industry are on average about 60% of the whole stock market turnover. Due to numerous special characteristics, we take this specific industry as our research target and collect all data from public websites, such as Taiwan Economic Journal Data (TEJ), Taipei Exchange (TE), and Taiwan Stock Exchange (TSE), for the period 2015-2017.

3.2. The dependent variable

How to appropriately assess a corporate’s operating performance is an urgent task in this current highly turbulent economic situation. Apart from those studies that merely employed one input and one output variable to determine operating performance, this study executes DEA as it can handle multiple input and output variables simultaneously and provide a performance score. Furthermore, in order to yield a more synthetic and overarching evaluated outcome, this study prefers to go beyond a single performance score and extends into multiple DEA specifications so as to depict a corporate’s intrinsic operating status.

Before the performance score calculation, the input and output variables should be decided first. In accordance with prior research, we choose total assets (TA) and total equity (TE) as input variables and designate three profitability ratios, return on assets (ROA), profit margin (PM), and return on equity (ROE), as output variables. To examine the representativeness of the chosen variables, we conduct the Pearson correlation.

Table 1 lists the results. We can see that all the chosen variables have a significantly positive correlation - that is, no chosen variable should be deleted.

<table>
<thead>
<tr>
<th>Variable</th>
<th>TA</th>
<th>TE</th>
<th>ROA</th>
<th>PM</th>
<th>ROE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TA</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>TE</td>
<td>0.673**</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ROA</td>
<td>0.352**</td>
<td>0.364**</td>
<td>1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>PM</td>
<td>0.416**</td>
<td>0.440**</td>
<td>0.452**</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>ROE</td>
<td>0.581**</td>
<td>0.469**</td>
<td>0.652**</td>
<td>0.952**</td>
<td>1</td>
</tr>
</tbody>
</table>

TA denotes I1; TE denotes I2; ROA denotes O1; PM denotes O2; ROE denotes O3.; *** denotes p < 0.01; ** denotes p < 0.05; * denotes p < 0.1.

Table 2

The multiple DEA specifications

<table>
<thead>
<tr>
<th>DEA1: I1-O1</th>
<th>DEA2: I1-O2</th>
<th>DEA3: I1-O3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEA4: I1-O1, O2</td>
<td>DEA5: I1-O1, O3</td>
<td>DEA6: I1-O2, O3</td>
</tr>
<tr>
<td>DEA7: I1-O1, O2, O3</td>
<td>DEA8: I1-O1, O2</td>
<td>DEA9: I1-O2, O2</td>
</tr>
<tr>
<td>DEA10: I1-O3</td>
<td>DEA11: I1-O1, O2</td>
<td>DEA12: I1-O2, O3</td>
</tr>
<tr>
<td>DEA13: I1-O2, O3</td>
<td>DEA14: I1-O2, O2, O3</td>
<td>DEA15: I1-O1, O3</td>
</tr>
<tr>
<td>DEA16: I1-O2, O2</td>
<td>DEA17: I1-O2, O3</td>
<td>DEA18: I1-O2, O1, O2</td>
</tr>
<tr>
<td>DEA19: I1-O3</td>
<td>DEA20: I1-O2, O2</td>
<td>DEA21: I1-O2, O1, O2</td>
</tr>
</tbody>
</table>

TA denotes I1; TE denotes I2; ROA denotes O1; PM denotes O2; ROE denotes O3.

prior operating strategies. Non-specialists can use the information of category change to modify their investment portfolios.

3.3. The independent variables

Kamei [20] indicated that the main reason for financial distress is bad operating performance. Based on this concept, because we realize that bad operating performance highly correlates to financial distress, the independent variables conducted in financial distress prediction can be taken as predictors in this study. Table 3 presents the selected predictors.

3.4. Assessment criteria

The most widely used criterion for measuring a model’s forecasting quality is accuracy (ACC) or error rate. However, merely using one criterion to reach a final conclusion is not reliable and robust enough. Thus, in addition to ACC, we take another four measuring criteria into consideration: precision (PRE), recall (REC), false positive rate (FP) and F-measure (FM). Table 4 shows the confusion matrix.
### Table 3

<table>
<thead>
<tr>
<th>Types</th>
<th>Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profitability</td>
<td>V1: NI/S</td>
<td>Net income to sales</td>
</tr>
<tr>
<td></td>
<td>V2: EBIT/TA</td>
<td>Earnings before interest and tax to total assets</td>
</tr>
<tr>
<td></td>
<td>V3: NI/FA</td>
<td>Net income to fixed assets</td>
</tr>
<tr>
<td>Solvency</td>
<td>V4: TL/TA</td>
<td>Total liabilities to total assets</td>
</tr>
<tr>
<td></td>
<td>V5: CA/CL</td>
<td>Current assets to current liabilities</td>
</tr>
<tr>
<td></td>
<td>V6: TL/TE</td>
<td>Total liabilities to total equity</td>
</tr>
<tr>
<td>Operational capabilities</td>
<td>V7: S/CA</td>
<td>Sales to current assets</td>
</tr>
<tr>
<td></td>
<td>V8: S/TA</td>
<td>Sales to total assets</td>
</tr>
<tr>
<td>Structural soundness</td>
<td>V9: FA/TA</td>
<td>Fixed assets to total assets</td>
</tr>
<tr>
<td></td>
<td>V10: TE/FA</td>
<td>Total equity to fixed assets</td>
</tr>
<tr>
<td>Capital expansion capability</td>
<td>V11: EPS</td>
<td>Earnings per share</td>
</tr>
<tr>
<td></td>
<td>V12: NA/OS</td>
<td>Net asset to outstanding shares</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>Predicted/Actual</th>
<th>Superior</th>
<th>Inferior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superior</td>
<td>True positive (TP)</td>
<td>False positive (FP)</td>
</tr>
<tr>
<td>Inferior</td>
<td>False negative (FN)</td>
<td>True negative (TN)</td>
</tr>
<tr>
<td>PRE: TP/(TP+FP); REC: TP/(TP+FN); FM:2*(PRE*REC)/(PRE+REC); FP: FP/(FP+TN); ACC: (TP+TN)/(TP+TN+FP+FN)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3.5. Empirical results

Prediction is an effective managerial tool in the process of planning that can give users a predicted outcome about future events, by relying on utilizing past and existing data and computational techniques. Due to the merits of prediction techniques, users can initiate some treatments to avoid financial troubles as well as reach sustainable development. We describe the proposed architecture as follows.

**Step 1. Construction of multiple DEA specifications**

Rather than using a single DEA specification (i.e., a single score), this study goes beyond a score and advances a more overarching and comprehensive synthesized performance measure to depict a corporate’s intrinsic operating status. Thus, we utilize two input and three output variables and offer 21 different DEA specifications.

**Step 2. Online data clustering**

Decision makers are unable to reach a consensus judgment when too much information (i.e., each corporate herein has 21 different DEA specifications) is involved. To overcome this problem, one can employ a clustering technique. FCM is a very well-known clustering technique, but because the numbers of clusters need to be updated in different time series, a conventional clustering technique like FCM loses its effectiveness. Thus, we use DFCM and enter the data for the period 2015-2017 into it. The updating process continues until the number of clusters no changes. Finally, the data are grouped into seven categories.

**Step 3. Essential feature identification**

In many financial forecasting tasks, such as credit risk assessment and financial crisis prediction, samples are represented as feature vectors in which many features are contaminated by some degree of error (i.e., redundant, vague, or irrelevant). Data with high dimensionality will lead to some disadvantages in the forecasting task, such as high computing cost/burden, over-fitting, bad performance, and large storage requirement [25, 26]. Thus, feature selection (FS) is an inevitable and required prior stage before forecasting model construction.

The basic idea of FS is determining a reduced subset from the original dataset without impeding the model’s forecasting quality as well as enhancing the computing efficiency. Peng et al. [35] introduced an emerging hybrid FS mechanism, naming it minimum Redundancy Maximum Relevance (mRMR). It can be divided into a two-stage mechanism, which first reduces the number of features by ranking and sequentially chooses the top ranked features by a wrapper-based algorithm [20]. It has widely demonstrated its usefulness and effectiveness. Thus, mRMR is conducted, and Fig. 1 shows its selected features.

**Step 4. Forecasting model construction and evaluation**

The analyzed outcomes derived from multiple DEA specifications, DFCM, and mRMR are then entered into ESVM to construct the forecasting model. We divide the data into two subsets: training subset and testing subset. The former is used to construct the model, and the latter is performed to examine the usefulness of the constructed model.
Five-fold cross-validation is executed to prevent the over-fitting problem. To test the usefulness of FS, we set up two scenarios: with FS and without FS. Table 5 shows the result. We can see that the model with FS achieves superior forecasting quality and a less biased outcome. This finding is in accordance with prior work done by Uysal and Gunal [44] who stated that FS can boost forecasting accuracy as well as overcome dimensionality. A statistical examination (i.e., t-test) is next performed to confirm the result does not just happen by chance.

Multiple DEA specifications\textsubscript{DFCM\_mRMR\_ESVM} in Scenario 1; Multiple DEA specifications\textsubscript{DFCM\_ESVM} in Scenario 2

To reach a more reliable outcome, we take our introduced architecture as a benchmark and compare it with three AI techniques: support vector machine (SVM), neural network (NN), and extreme learning machine (ELM). One of the non-parametric statistics, called the Friedman test, poses the advantage of easy-to-use, depicts the overall performance of models in the form of rank instead of dubious averages, and is taken as an assessment criterion under several of our comparisons [9]. Table 6 shows the results. We see that the introduced model outperforms the other three AI models under all assessment criteria. This finding is in line with Liu et al. [28] who indicated that ESVM not only has the same merit as ELM, such as extremely fast learning speed, but also possesses comparable forecasting quality and superior generalization ability like SVM. The model, examined by real cases, is a promising alternative model for performance forecasting.

\textbf{Step 5. Extraction of decision logics}

Although ESVM obtains accurate forecasting quality, the embedded decision logic is regarded as a black-box. In other words, the decision logics are opaque and unintuitive to the users. Black-box forecasting models prevent users from tracing the logic behind a forecasted outcome and acquiring the useful knowledge previously unknown from the model [4]. These models also do not permit human judgment and inspection, they are not directly understandable by decision makers, and they are unable to identify and discover which relevant features forecast the class labels. This weakness (i.e., opacity) impedes the model’s real-life knowledge discovery applications where both preciseness and interpretability are required, such as credit risk assessment [31] and decision support system [2], since the forecasting model must explain the reasons for the forecasted outcome. To open up the opaque nature of ESVM, this study, grounded on a pedagogical structure, extracts the inherent knowledge from ESVM by an emerging rule learner called DataSqueezer [23]. This learner not only provides compact, efficient, human-readable rules, but is also robust to large quantities of missing data. Table 7 presents the extracted knowledge. We can see that a corporate with superior operating performance usually exhibits higher profitability, suitable financial leverage, efficient asset utilization, and large risk absorbing capability. Take Rule 2 and Rule 3 as example, the managers can consider adopting some promoting strategies to stimulate the sales volumes and modifying the capital structure to avoid encountering heavy financial burden so as to reach a sustainable development.

\textbf{3.6. Robustness examination}

Most prior studies have reached a final conclusion that merely relied on one pre-decided dataset that is not reliable and trustful in today’s highly volatile atmosphere. To make our research findings more robust, we consider two scenarios: (1) change the performance measure, and (2) change the clustering technique.

Table 8 represents the results in scenario 1. We see that the performance measure determined by multiple DEA specifications has better discriminant capability than the other three performance measures (i.e., ROA, ROE and single DEA) - that is, the data that went through multiple DEA specifications and then fed into DFCM can generate more clusters. This means that DFCM uses the entered data with sufficient information (i.e., data that are information-contained) to appropriately discriminate the difference between each cluster. If the data (i.e., ROA, ROE single DEA) entered into DFCM can only generate fewer clusters, then it means that the entered data does not have sufficient information (i.e., data that are not information-contained) for DFCM to discriminate
the difference between each cluster. This finding is in comparison to Sagarra et al. [41] who indicated that the utilization of alternative DEA specifications - i.e. combining inputs and outputs in several dissimilar ways - can provide a wider and robust set of information for classifying and clustering the observations. The proposed architecture still outperforms the other three AI techniques under the three different performance measures.

Table 9 shows the result in scenario 2. We see that DFCM performs better than the other two clustering techniques (i.e., FCM, K-means) [32]. The proposed architecture achieves outstanding forecasting quality with less biased outcomes in all situations.

4. Conclusions and future works

Corporate financial performance forecasting helps decision makers form appropriate judgments and avoid dramatic financial crises, especially during a downturn in the economic landscape. However, as most previous studies only relied on one input and one output variable to determine a corporate’s operating status, these simple measures cannot describe the intrinsic and real operating status of corporates being examined. To reach a more robust analyzed outcome, this study goes beyond a single score and further implements an overarching and comprehensive synthesized measure, called multiple DEA specification, to evaluate a corporate’s operating performance. However, too much information will mislead the decision findings. Thus, DFCM is conducted to overcome this obstacle as well as to make the analyzed outcome more accessible to non-specialists.

We also note that DFCM generates dynamic clusters that can be used to represent a corporate’s operating performance change over time. Through this, top-level managers can consider the potential implications and further realize the effectiveness of prior implemented business strategies. The outcome derived from multiple DEA specifications and DFCM are then entered into ESVM to construct the forecasting architecture, and the introduced model, examined by real cases, reaches a satisfactory forecasting quality and less biased outcome.

The fundamental idea in this research is inspired by a hybrid mechanism that can complement the error made by a single mechanism as well as increase the preciseness of the forecasted outcome. Even a fraction of improvement in forecasting accuracy can translate into a large amount of future monetary savings. In today’s knowledge-based economy, obtaining a comprehensible forecasting model may be as essential as reaching high preciseness. To open up the opaque nature of ESVM, a rule learner is
executed that expresses the knowledge in a transparent and interpretable format. If the decision logics can be judged or tested by users, then the model’s acceptance rate will increase considerably, thus enhancing its real-life applications.

Future works can consider two potential research directions. First, this study only focused on the electronics industry in Taiwan, which suggests that the ability to generalize the outcomes may be limited. Future works can look into other industries in different countries. Second, future research can extend the introduced architecture to a much more advanced format, i.e., classifier ensemble, in order to increase forecasting accuracy.
Disclosure statement

No conflict of interest exists in the submission of this manuscript.

Acknowledgements

The authors would like to thank the Ministry of Science and Technology, Taiwan, R.O.C. for financially supporting this work under contracts MOST No.106-2410-H-034-046-MY2, No.106-2410-H-034-011-MY3 and No.107-2410-H-110-025-MY2.

References


