Prediction Model for Financial Distress Using Proposed Data Mining Approach By

Raghan Mohammed Hadi¹ Shatha H. Jafer Al-khalisy² Najlaa Abd Hamza³

¹Mustansiriyah University – College of Administration and Economics
²University of Technology- computer science Department
³University of Baghdad/ College of Nursing

Email: najlaa_al_mayyaehe@yahoo.com

Received: 23/1/2019 Revised: 10/3/2019 Accepted: 11/3/2019

Available online: 17/4/2019

Abstract:
The problem of financial distress researches are the lack of awareness of banks about the risks of financial failure and its impact on the continuity of its activity in the future, as the traditional methods used to predict financial failure through financial analysis based on financial ratios in a single result gives misleading results cannot be relied upon to judge the continuity of the activity of banks. With an increase in the number of failed banks and their inability to continue. Which requires the discovery of modern techniques that serve as an early warning of the possibility of failure and lack of continuity. The research aims to apply data mining technology to predict the financial failure of banks, and how it can provide information that helps to judge the extent to which banks continue to operate. This effort suggested founded back propagation artificial neural network to build predict system. The proposed module evaluated with banks from Free Iraq Stock Exchange dataset the investigational outcomes displays capable method to identify failure banks with great discovery rate and small wrong terror rate.

Keywords: Data mining, back propagation, Financial Distress, banks.
1- Introduction

The assumption of continuity for accounting thought is that the accounting unit since its inception is a continuous and continuous unit of activity, without regard to the natural age of the owners. That is, the general purpose of imposing continuity (intuitive continuity) is to separate the age of the economic unit and the age of the owners as individuals independent of each other and have separate objectives and plans to pursue. And that the natural view of economic unity is the long-term continuation of its activity without separation from previous periods and later, until the liquidation of the actual. Since the imposition of continuity was considered one of the fundamental assumptions in accounting theory because it is one of the cornerstones of the foundations and rules of accounting measurement, which left its impact on professional practices in terms of avoiding unity and some potential losses, the imposition of continuity [3].

The economic unit is continuous for a period that is only money at the end of an indefinite period. This period is sufficient to achieve its objectives. This assumption is considered as a justification for ignoring the current liquidation values when presenting the assets and liabilities in the statement of financial position [2].

Cases of imposing continuity

The current reality was often saw the disappearance of economic units, whether companies or organizations and solutions economic units in place, and this shows that the imposition of continuity is the imposition of the present and not the future, i.e., economic units will operate for an indefinite period of time until the achievement of its objectives and economic plans and three cases can distinguish of imposing continuity, which are as follows:

1. Normal (natural) condition:

   This case indicates that it is expected that the economic unit will continue to operate for a long period of time in order to be able to pay its obligations and benefit from its resources, i.e the establishment is able to continue the activity without providing a specific and specific information at present, and that the imposition of continuity relates to past and present activity And is not inconsistent with activity for future periods. Thus, this assumption is based on a logical interpretation that at any point or time in a given time the enterprise can continue its work [5].

2. The state of the mind or the limited duration

   Is the case where the age of the economic unit is predetermined, such as the grant of a government concession to invest a particular facility for a period of time For a period of time and the follow-up of the investment of this facility after the end of the period to the government or to the donor of the right of freedom. This case shows that the economic unit lasts for a known period at present and is called the case of the concession or duration of the concession. In such cases, continuity is explained by the useful life of the asset or the duration of the concession, In light of this, accounting measurement and financial statements are prepared as the economic life of the unit approaches. In this case, it is necessary to exclude the imposition of continuity and the adoption of a method that takes into account the remaining period and the nature of the liquidation process to determine whether the liquidation is gradual or forced. In the case of liquidation, the accounting measure is adopted according to the expected net value [1].

3. The existence of indicators at the present time allowing the presumption of discontinuity

   Is an abnormal condition that clarifies and identifies indicators of discontinuity in terms of probability of liquidation or actual liquidation status. It is a predictive state that represents a constraint on the normal state of continuity, i.e., it is similar to the forecasting process, may be correct or absent. In this case, Report on the validity and fairness of the financial statements [1].

Related works

The approach presented in [7] the notice in the forecast of shared insolvency is cumulative owing to the insinuations related through this wonder (e.g. financial, and communal) for depositors, creditors,
The discovery and examination process on the upcoming depositors of Iraqi business was an attempt to determine whether such businesses are insolvency or dependable. The sampling was chosen from the Iraqi market which is still suffering from the ensuing war in the field, and the statistical data were collected for the year 2018. Several economical models were previously done and tested, and the growth of the selection and examination was completed, which was considered as the training phase of the model. The main objective of this paper is to design an efficient data mining system that can detect financial failure in banks to identify the risk associated with loans and borrowers, while reducing false positive alerts with affordable computational cost. The proposed data mining system able to detect Continuity banks from Banks that will be spared. The detection process in multi-level system made through two consequence detection module where cluster module used to detect normal continuity banks from abnormal and classifier module used to detect the type of failure. Multi-level data mining system: the proposed system consist of four steps which is First step: data standardization, Second step: feature selection, Third step: training phase, and Fourth step: testing phase, as shown in figure (1).

The Proposal System

The main objective of this paper is design an efficient data mining system that can detect financial failure in Banks to identify the risk associated with loans and borrowers, while reducing false positive alerts with affordable computational cost. The proposed data mining system able to detect Continuity banks from Banks that will be spared. The detection process in multi-level system made through two consequence detection module where cluster module used to detect normal continuity banks from abnormal and classifier module used to detect the type of failure. Multi-level data mining system: the proposed system consist of four steps which is First step: data standardization, Second step: feature selection, Third step: training phase, and Fourth step: testing phase, as shown in figure (1).
### Data Standardization

1. Data collection (obtaining a database of banks from the Iraqi Stock Exchange as a dataset used).
2. Preprocessing is carried out through the extraction of several early warning control standards used to measure the soundness of finance presentation. These ideals are occupied as needles to measure the presentation of sets and before to classify them and detect economic imbalances popular their presentation before they are not unprotected to rainy monetary difficulties that lead to their failure. And the most important of these criteria is the system of assessment of banks according to the indicators resulting from the field examination process called CAMELS, and was assumed in the selection of ratios on several criteria, including: (A) to be meaningful. (B) Facilitate its interpretation. (C) Access to its components is easy.
3. Alter the price of emblematic piece (CAMELS criteria) to consecutive numeral price. Then BP and FCM process take arithmetical cost.

### Piece selection

1. Compute information of the class (Failure banks or continue banks) using the equation depended on results of CAMELS criteria:

\[
\text{information class(m)} = \sum_{i=1}^{m} \frac{\text{distinct class } i}{\text{number of data samples}} \log_2 \left( \frac{\text{distinct class } i}{\text{number of data samples}} \right)
\]

2. Compute the entropy for each feature in the dataset using equation

\[
\text{entropy(Feature)} = \sum_{j=1}^{v} \frac{\text{number of samples}_{ij} + \cdots + \text{number of samples}_{mj}}{\text{number of data samples}} \times \text{information}(C_{ij} + \cdots + C_{mj})
\]

3. Compute the information gain for each feature using equation

\[
\text{Gain(Feature)} = \text{Information class}(C_1, \ldots, C_m) \quad \text{– entropy(Feature)}
\]

### Training phase

- The first level: Train with FCM algorithm to build cluster module to detect the failure bank from continuity banks.
- The second level: train with BP algorithm to build classifier module to detect the type of bank in financial cluster.

#### First level of training phase in the proposed system as details

1. Initialize the cluster centre randomly.
2. Compute the membership matrix for each sample using
3. Update the cluster centre using
4. Allocate group category aimed at each bank through maximum association price.

\[
\text{membership}_{i \text{ of sample } i \text{ regard to cluster } j} = \frac{1}{\sum_{j=1}^{k} (\text{sample } i - \text{cluster center } j)/\text{sample } i - \text{cluster center } k)^2/\text{fuzzienss} - 1}
\]

---

**Figure 1: Block diagram of proposed system**
Second level of proposed system

1. Isolate the failure bank in cluster failure to explicit failure clusters according to the type of failure according to CAMELS criteria.
2. Prepare the weightiness vectors since effort elements to concealed elements and from concealed elements to production elements arbitrarily.
3. Calculate wanted production for each failure cluster from labelled failure bank s.
4. Every piece in failure cluster bank embodies contribution element \( X_i \) before transmissions to every neuron in concealed coating.
5. Calculate the production of every element in the concealed coatings \( z_j \) by

\[
Hidden_j = \text{bias} + \sum_{i=1}^{\text{number of input unit}} \text{input unit}_i \times \text{weight}_{ij}
\]

6. Now the production of concealed coating is managed by

\[
hidden(j) = \frac{1}{1 + \exp^{-hidden_j}}
\]

7. Calculate the production of all component hip the production coating \( Y_k \) by equality

\[
output of k = \text{bias} + \sum_{j=1}^{\text{number of hidden units}} \text{hidden}_j \cdot \text{weight}_{jk}
\]

8. Now the production treated by equality

\[
output(k) = \frac{1}{1 + \exp^{-output_k}}
\]

9. Calculate the error value between output units and hidden units using equation

\[
\text{error of output}_k = \text{output}_k(1 - \text{output}_k)(\text{target}_k - \text{output}_k)
\]

10. Calculate the error value between hidden units and input units using equation below:

\[
\text{error hidden}_j = hidden_j (1 - \text{hidden}_j) \sum_{k=1}^{\text{number of output unit}} \text{error output}_k \times \text{weight}_{jk}
\]

11. Inform the masses \( w_{ij} \) and \( w_{jk} \) among coatings rendering to equalities.

\[
\text{weight}_{jk}(\text{new}) = \text{learning rate} \times \text{error output}_k \times \text{hidden}_j + \text{momentum} \times \text{weight}_{jk}(\text{old})
\]

\[
\text{weight}_{ij}(\text{new}) = \text{learning rate} \times \text{error output}_k \times \text{input}_i + \text{momentum} \times \text{weight}_{ij}(\text{old})
\]

12. Calculate the mean square error value \( E_m \) using equation

\[
\text{error} = \frac{1}{2} \sum_{\text{number of sample}} \sum_k \left( \text{target}_k - \text{output}_k \right)^2
\]

Results and Experiments

Banks from Free Iraq Stock Exchange dataset were evaluated by the proposed algorithm. The projected procedure educated with 23 banks carefully chosen since dataset contains period from 2015 to 2017 as a samples besides the other private financial statements of free Iraq stock exchange to agree standard banks since failure banks and too notice the category of banks. Three estimation standards castoff to measure the projected element. To checker the competence of the projected module two trials Lead, in the first trials the procedure examination through (26) banks contain financial statements. The effects acquire display the great discovery rate of the model equal to (0.92) and little wrong terror rate (0.04). The additional trials behavior with (30) registers separately also include failure banks. The outcomes of the trials raining in identical variety and exposed in table (2), and table (1) describe the evaluation measures used in the proposed system. The topology of network shown in table (3).
Table (1): The evaluation measures in the proposed system

<table>
<thead>
<tr>
<th>Name of measure</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Negative (TN)</td>
<td>Specify the correct prediction of continue banks.</td>
</tr>
<tr>
<td>Correct positive (TP)</td>
<td>Specify the correct prediction of failure banks.</td>
</tr>
<tr>
<td>incorrect positive (FP)</td>
<td>Specify the wrong prediction of continue behavior as failure.</td>
</tr>
<tr>
<td>incorrect negative (FN)</td>
<td>Specify mistake prediction of failure behavior as continue banks.</td>
</tr>
<tr>
<td>Accurateness (ACC)</td>
<td>$\frac{TP + TN}{TP + TN + FP + FN}$</td>
</tr>
<tr>
<td>Discovery rate (DR)</td>
<td>$DR = \frac{TP}{TP + FN}$</td>
</tr>
<tr>
<td>Incorrect alarm rate (FAR):</td>
<td>$FAR = \frac{FP}{TN + FP}$</td>
</tr>
</tbody>
</table>

Table (2): Trial result of the proposed algorithm

<table>
<thead>
<tr>
<th>Number of trial</th>
<th>Number of samples (banks from 2015-2017)</th>
<th>DR</th>
<th>Accuracy</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>26</td>
<td>0.92</td>
<td>0.97</td>
<td>0.03</td>
</tr>
<tr>
<td>Trial 2</td>
<td>30</td>
<td>0.90</td>
<td>0.98</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table (3): Constraints of Back Propagation Algorithm

<table>
<thead>
<tr>
<th>Constraints name</th>
<th>Value parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of layer</td>
<td>3</td>
</tr>
<tr>
<td>basis</td>
<td>1</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.5</td>
</tr>
<tr>
<td>Number of component in contribution layer</td>
<td>30</td>
</tr>
<tr>
<td>Number of component in concealed layer</td>
<td>25</td>
</tr>
<tr>
<td>Number of component in production layer</td>
<td>5</td>
</tr>
<tr>
<td>Malicious square error</td>
<td>0.0001</td>
</tr>
<tr>
<td>Maximum amount of repetition</td>
<td>25</td>
</tr>
</tbody>
</table>
References


8. سيف عباد رزاق محمد رمط، سيف عدد محمود يحيى فتحي، “Data mining and decision-making (An Applied Study on KhashmAlgirba Dam), مجلة النيل الأبيض للدراسات والبحوث العدد الثالث، مارس 2014م.

9. سيف عبد الرازق جبر الوطار، ”استخدام أساليب التحليل المالي في التنبؤ بإغلاق الشركات’، المساهمة الصناعية: دراسة على عينة من الشركات الصناعية العراقية المدرجة في سوق العراق "المساهمة كنسبة الإدارة والأعمال"، جامعة "الحدود المالية الموصل - النزاع"، المجلة العدد 32، السنة 2010."
نموذج التنبؤ بالضائمة المالية باستخدام طريقة تنقيب البيانات المقترحة

النجم محمد هادي
نجلاء عبد حمزة

كلية الإدارة والاقتصاد - الجامعة المستنصرية
قسم علوم الحاسوب - الجامعة التكنولوجية
كلية التمريض - جامعة بغداد

الخلاص:
مشكلة البحث في موضوع الضائمة المالية هي قلة الوعي حول بنوك الأسواق المالية التي تعطي نتائج مضللة لا يمكن الاعتماد عليها في استمرارية أنشطة البنوك. الأمر الذي يتطلب اكتشاف التكتينات الحديثة التي تساعدها إدرازًا مبكرًا لاتمام فشل البنوك وعدم الاستمرارياتها. يهدف البحث إلى تطبيق تكنولوجيا استخراج البيانات للتنبؤ بالفشل المالي للبنوك. في هذا البحث استخدمت خوارزمية الشبكات العصبية الاصطناعية (BP) لبناء نظام كشف فشل أو استمرارية البنوك. النظام المقترح اختبر باستخدام بيانات المصارف الموجودة في سوق العراق الحر للأوراق المالية وافهرت النتائج ان النظام المقترح يشكل طريقة واعدة لتنبؤ فشل البنوك بطريقة عالية مع نسبة انذار كاذبة منخفضة.