

Prediction Model for Financial Distress Using Proposed Data Mining Approach By

Raghad 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_mayyahee@yahoo.com

Recived : 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, ie, 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, ie, 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,

contestants, administration, though this is a traditional problematic in the monetary works. Dual types of replicas are usually accepted for insolvency forecast: (i) secretarial relations founded replicas and (ii) shop founded replicas. In the previous, traditional arithmetical methods for example discriminant examination or logistic reversion replicas must be castoff, though in the last the Moody's KMV model was accepted. The paper founded on the examination of the development of numerous monetary pointers throughout a three year dated. A structure was industrialized, surrounding a whole of 15 replicas. These change in the data mining procedure (for example artificial neural networks or decision trees), the statistics castoff and the contribution qualities accepted. The trials were lead consuming the original Commercial Intellect Progress Pottery of the SQL Server. Respectable consequences were attained, with presentations amid 86% and 99% for each 15 models. In [8] the usage of methods data mining delivers for organizations in wholly spaces the capability to discover and emphasis on the greatest significant material in the files, similarly converging methods data mining in addition to to shape a upcoming predictions and examination performance and tendencies, which lets to income the correct choices and occupied in a opportune way. In [9] the disappointment of business is measured to be the most significant topics deliberate by numerous writers and assistants as for the inactive effects that might consequence from them on the general cheap in addition to on the complete humanity. Consequently, numerous investigators happening to discovery a technique which they can forecast catastrophe previously it receipts place. As a consequence, huge educations secondhand the economic charges as tackles for convinced numerical replicas created on together solitary and numerous division examination methods for the option of disappointment forecast. In addition to appropriate models. The extant investigation purposes at discovery a dependable method for disappointment forecast done relating the Altman Model in a amount of Iraqi paying businesses. The position of the investigation has originated from the meaning of disaster topic amongst numerous gatherings connected to the business. The investigation tricky dishonesties in that the Iraqi causal businesses, depositors and further festivities are not alert of threats that principal the businesses to letdown in the upcoming. The classical was practical on example

contained of 17 Iraqi causal business afterward procurement the essential info around them. Numerous results were reached; the greatest significant of which is Altman Model's accurateness in forecasting the letdown of Iraqi businesses. The education decided that Altman Model for letdown forecast would be accepted by way of a method of economic examination dependable to the assessment of businesses presentation. In [6]

Offerings another method for financial suffering forecast schemes. The technique is founded on a kind of neural network, which is named mixture associative recollection with conversion. Though numerous dissimilar neural network buildings have positively rummage-sale to forecast recognition danger and business disappointment, the control of associative reminiscences for monetary choice creation has not been traveled in any complexity as however. In [4] the presentation of the mixture associative recall with conversion is likened to four outdated neural networks, a support vector machine used for prediction competences. The new consequences ended actuality data sets demonstration that the associative recall planned establishes a suitable answer for insolvency and recognition danger forecast, execution meaningfully healthier than the respite of replicas below period inequity and statistics meeting circumstances in footings of the factual optimistic degree and the symmetrical nasty of correct optimistic and factual undesirable rates.

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 equator depended on results of CAMELS criteria :

infomation 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

$$\text{entropy(Feature)} = \sum_{j=1}^v \frac{\text{number of samples}_{ij} + \dots + \text{number of samples}_{mj}}{\text{number of data samples}} \times \text{infomation}(C_{ij} + \dots + C_{mj})$$

3. Compute the information gain for each feature using equation

$$\text{Gain(Feature)} = \text{Information class}(C_1, \dots, C_m) - \text{entropy(Feature)}$$

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

$$\text{cluster center } j = \frac{\sum_{i=1}^{\text{number of sample}} \text{membership}_{\text{of sample } i \text{ regard to clstuer } j}^{\text{fuzzienss}} \cdot \text{sample } i}{\sum_{i=1}^{\text{number of sample}} \text{membership}_{\text{of sample } i \text{ regard to cluster } j}^{\text{fuzzienss}}}$$

5. Allocate group category aimed at each bank through maximum association price.

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

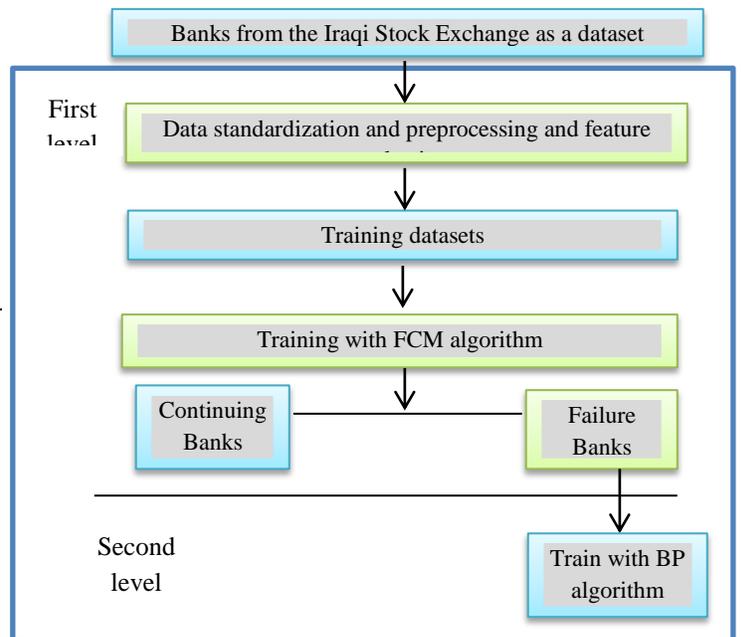


Figure 1: Block diagram of proposed system

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 banks.
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

$$j = \text{bais} + \sum_{i=1}^{\text{number of input unit}} \text{input unit}_i \text{ weight}_{ij}$$

6. Now the production of concealed coating is managed by

$$\text{hidden}(j) = \frac{1}{1 + \exp^{-\text{hidden } j}}$$

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

$$\text{output of } k = \text{bais} + \sum_{j=1}^{\text{number of hidden units}} \text{hidden } j \cdot \text{weight}_{jk}$$

8. Now the production treated by equality

$$\text{output}(k) = \frac{1}{1 + \exp^{-\text{output } k}}$$

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

$$\begin{aligned} \text{error of output } k &= \text{ouptu}_k(1 \\ &- \text{outptu}_k)(\text{target}_k \\ &- \text{output}_k) \end{aligned}$$

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

$$\begin{aligned} &= \text{error hidden } j \\ &= \text{hidden}_j (1 - \\ &\text{hidden}_j) \sum_{k=1}^{\text{output number}} \text{error output}_k \text{ weight}_{jk} \end{aligned}$$

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

$$\begin{aligned} \text{weight}_{jk}(\text{new}) &= \text{learning rate} \\ &* \text{error output}_k * \text{hidden}_j \\ &+ \text{momentum} \\ &* \text{weight}_{jk}(\text{old}) \end{aligned}$$

41

$$\begin{aligned} \text{weight}_{ij}(\text{new}) &= \text{learning rate} \\ &* \text{error ouptu}_k * \text{input}_i \\ &+ \text{momentum} \\ &* \text{weight}_{ij}(\text{old}) \end{aligned}$$

12. Calculate the mean square error value E_m using equation

$$\text{error} = \frac{1}{2} \sum_{\text{number of sample}} \sum_k (\text{target}_{pk} - \text{output}_{pk})^2$$

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

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

Table (2): Trial result of the proposed algorithm

Number of trial	Number of samples (banks from 2015-2017)	DR	Accuracy	FAR
Trial 1	26	0.92	0.97	0.03
Trial 2	30	0.90	0.98	0.04

Table (3): Constraints of Back Propagation Algorithm

Constraints name	Value parameters
No. of layer	3
basis	1
Learing rate	0.5
Number of component in contribution layer	30
Number of component in concealed layer	25
Number of component in production layer	5
Malicious square error	0.0001
Maximum amount of repetition	25

References

1. Amano, R. and S. van Norden, "A Forecasting Equation for the Canada-U.S. Dollar Exchange Rate." In The Exchange Rate and the Economy, Proceedings of a conference held at the Bank of Canada, June 2010. Ottawa: Bank of Canada, 2010.
2. Elhassan H. I. M. and Mohammed S. M., "Using data mining techniques to forecast the volume of exports and imports", Sudan Customs- IT Department, 2017.
3. Hamid A. J. and Ahmed T. M., "Developing Prediction Model Of Loan Risk In Banks Using Data Mining", Machine Learning and Applications: An International Journal (MLAIJ) Vol.3, No.1, March 2016.
4. Itkar S., Kulkarni U., "Efficient frequent pattern mining using auto-associative memory neural network" , Brit. J. Appl. Sci. Technol. 4 (2014) 3160–3178.
5. Dr. Khareem Ali A. A., Prof. Dr. Omer F. K., "The Use Nets of Artificial Neural Networks to Predict of Macroeconomic Multidimensional model in Iraq 1996 to 2007", 2013.
6. Lessmann S., Baesens B., Seow H.V, Thomas L.C, Benchmarking state-of the art classification algorithms for credit scoring: An update of research, Eur. J. Oper. Res. 247 (2015) 124–136.
7. Santos M.F., Cortez P., Pereira J. and Quintela H., "Corporate bankruptcy prediction using data mining techniques", Data Mining VII: Data, Text and Web Mining and their Business Applications, WIT Transactions on Information and Communication Technologies, Vol 37, 2006 WIT Press www.witpress.com, ISSN 1743-3517 (on-line) doi:10.2495/DATA060351.
8. الشفيق جعفر محمود , سيف الدين عثمان فتوح. "Data mining and decision-making (An Applied Study on KhashmAlgirba Dam), مجلة النيل الأبيض للدراسات والبحوث العدد الثالث مارس ٢٠١٤م
9. سيف عبد الرزاق محمد الوتار , وحيد محمود رمو , استخدام أساليب التحليل المالي في التنبؤ بفشل الشركات " المساهمة الصناعية: دراسة على عينة من الشركات الصناعية العراقية المدرجة في سوق العراق *المساهمة كلية الإدارة والاقتصاد-جامعة "للأوراق المالية الموصل تنمية الرافدين العدد ١٠٠ مجلد ٣٢ لسنة ٢٠١٠.

نموذج التنبؤ بالضائقة المالية باستخدام طريقة تنقيب البيانات المقترحة

نجلاء عبد حمزة^٣

شذى حبيب جعفر^٢

رغد محمد هادي^١

كلية الادارة والاقتصاد- الجامعة المستنصرية^١

قسم علوم الحاسوب- الجامعة التكنولوجية^٢

كلية التمريض- جامعة بغداد^٣

المتخلص:

مشكلة البحث في موضوع الضائقة المالية هي قلة الوعي حول بنوك الأسواق المالية التي تعطي نتائج مضللة لا يمكن الاعتماد عليها في استمرارية أنشطة البنوك. الأمر الذي يتطلب اكتشاف التقنيات الحديثة التي تعطينا إنذارًا مبكرًا لاحتمال فشل البنوك وعدم الاستمراريته. 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 للاوراق المالية واطهرت النتائج ان النظام المقترح يشكل طريقة واعدة تتميز بدقة عالية مع نسبة انذار كاذبة منخفضة.