

USING DATA MINING TOOLS TO PREDICTION OF GOING CONCERN ON AUDITOR OPINION - EMPIRICAL STUDY IN IRAQI COMMERCIAL

Saad Salman Awad, University of Baghdad
Ilham Mohammed Wathik, Al-IraqiaUniversity

ABSTRACT

This research aims to test a set of financial variables on a number of private commercial banks in Iraq that have proven the reality of the situation. Some of them stumbled and the continuity of others, for the purpose of determining the most appropriate ratios to confirm the auditor in his report on the company's ability to continue. The researchers used a number of tools, including Decision tree, ID3, Naive Byes, and Random Forest, to search for data in 13 banks, 7 of which are faltering and 6 continuous. This study is considered as a first attempt to test these technologies in the environment of Iraqi banks. Decision tree, ID3 has achieved a match with the actual results by 100% through two ratios DR & LTLR in the first technique and two ratios ROA in the second technique, while a technique LTLR ratio achieved a match with reality. The case is 83.33% through two percentages.

Keywords: Data Mining, Classification, ID3, Naive Byes, Decision Tree, Going Concern, Auditors Opinion, ISA 570.

INTRODUCTION

The need for data mining techniques in the field of accounting and auditing has grown rapidly, because communication systems and information technology have facilitated accounting treatments in many aspects, but have made auditing more difficult due to lose the audit trail on the one hand, and the complexity of processing operations on the other hand. Auditors used CAAT software to make audits closer to accuracy and reliability. In fact is that the method of auditing differed after it was done periodically, and it was performed in a continuous audit method through the flow of data through electronic circuits (Wang & Yang, 2014: 1). Therefore, they say that the auditing profession is now facing a new challenge in light of information technology systems. To solve this problem, auditing companies began to use the method of continuous auditing more and more and it is expected that the continuous audit will replace the annual audit in the future, in addition to that the needs of stakeholders for information in a timely manner pushed towards changing the traditional auditing method to the method of continuous auditing and from here became the need to use data mining techniques more widely.

The difficult task facing contemporary business organizations is survival and continuation, as some large organizations have collapsed without prior warning. Therefore, auditing standard 570 came to place before the auditor the responsibility to express an opinion on whether the facility is able to continue or is unable, and for the auditor to express an opinion, he must adopt his opinion To quantitative and non-quantitative information, and in the presence of the vast amount of data that the company deals with, the importance of using data mining techniques in accounting and auditing, this research deals with some of these techniques.

LITERATURE REVIEW

Rostami et al., (2011) presentation of basic concepts about data mining, its stages and the use of techniques, by financial institutions, insurance companies, tax authorities and other agencies, and the study concluded that data mining techniques can be used in predicting the behavior of customers, shareholders and stakeholders in general. several researches have been conducted in the applied field.

Mojtaba & Fahime (2012) have test a Novel approach for rule extraction from Support Vector Machine & Decision Tree for test the impact factors in predicate for auditors opinions .The researchers used the current ratio and fast ratio to identified liquidity position and liquidity ratio to identified financial health .

Fezeh & Mahdi (2013) tested this going concern on 146 Iranian companies listed in the capital market, including 73 companies exposed to bankrupt, with companies still continuing to list for the years 2001-2011 and the researchers used 42 variables, the data were entered in the Stepwise Discriminate Analysis test to arrive at the variables that distinguish continuous companies from non-continuous companies There are four ratios that distinguish between them, which are the ratio of total liabilities / total assets, operating income / sales, net income / total assets, retained earnings / total asset.

Wang & Yang (2014) identified three bases methods to data mining specifically in perpetual and periodical auditing as well as forensic audit as a new function of auditors, the methods are mathematical based methods, distance-based methods, and logic-based methods this paper identifies many types of data mining software that auditors can used in their job. Chen & Puchiu (2014) have used two properties of business intelligence and analysis which include the high dimensional the results which indicate the potential usefulness of business intelligence and analysis in going concern prediction etc.

Angle et al. (2018) provide an exclusive going concern prediction model for the hotels industrial by using computational method of variable selection. The researchers have identified the variables that could predict survival in this industrial used a sample of 252 hotels in Spain for the period from 2000-2014. The sample was divided into 126 hotels. The auditors gave an opinion of their ability to persistence and the other 126 not gave opinion of their reports. For testing purpose the researches use 12 rats financial and Non-financial for the purpose of identifying a set of explanation variables that can predict conditions of uncertainty in the hotel industry, in order to obtain a higher prediction level and fewer variables the empirical study showed that the method Ada Boost used from Boosting algorithms to nearly 100% accuracy.

Predications show that the comprehensibility of new algorithm is better than that of Support Vector Machine & high quality rules can be generated. several researches have been conducted in the applied field.

The Concepts of Data Mining

Data mining refers simply to extracting or mining knowledge from large numbers of data. Data mining can be viewed as a result of the natural evolution of information technology (Han et al., 2012:2-6). Data mining is the use of automated data analysis techniques to detect previously undetected relationships between data elements. Data mining often involves analyzing the data stored in it data warehouse (Sahu et al., 2012: 114). Thus, we can conclude that data mining is carried out on large, useful and less useful data, old and recent, qualities and quantitative, to extract data that are most appropriate for the goal that the researchers and other users intends to achieve.

The Need of Data Mining in Accounting and Auditing

The need for data mining techniques can be in several areas, but we will focus on aspects related to accounting and auditing, especially with the expansion of the data that are processed accounting and the multiplicity of financial reporting objectives and the diversity of its users.

Compliance Assessment

In the past, auditors indicated in their reports within the scope of expressing an opinion that the financial statements of the establishment were prepared in accordance with the accounting principles and rules, laws and legislations in force, but this opinion is no longer appropriate after the multiplicity of accounting alternatives and the adoption of domestic and international accounting and auditing standards, as well as the complexity and expansion of financial operations and their conduct through Information technology, and after the financial scandals that hit the financial markets of the big companies Enron, Xerox and others, which put the profession of accounting and auditing in great embarrassment, especially since some of the major auditing companies such as Author Anderson audited the financial statements of these companies. In the wake of these disasters, the Sarbance & Oxley Act was issued, which aimed to implement the governance requirements of transparency, accountability and integrity through the development of internal control systems and financial reporting. Based on this law, the Public Company Accounting Oversight Board (PACOB) was formed for the Public Shareholding Companies (Board Oversight). This law elaborated on explaining the responsibilities of management and the role of the auditor in verifying the extent to which they carry out these responsibilities related to the financial statements, and the extent to which this affects Auditor's report Based on this law, the auditor in the United States is required to evaluate the efficiency and effectiveness of the internal control system and report that to the shareholders in his report on the financial statements, as the auditor's report consists of five paragraphs that include the introduction clause, the scope clause and the determinants of the internal control system And the opinion paragraph (Khaddash et al., 2011: 599), and thus the application of this law requires high costs estimated at one million dollars for companies that have a market value ranging from 75 million dollars to 699 million dollars (Wang & Yang, 2014: 11).

In order to reduce these costs and fulfill these requirements in a timely manner and with minimal effort, auditors must use appropriate tools and techniques for analyzing financial data and data mining techniques.

Auditors Opinion in Financial Statements

The importance of the auditor's report stems from being the means by which the auditor can express his opinion on the fairness of the financial statements. The auditor's report is also one of the main references that are relied upon to determine the auditor's responsibilities, whether civil or criminal (Fernández et al., 2018).

Expectation Gap in Audit

The trends in the expectations gap go in two basic directions. The first is the difference between the public's expectation of the auditing profession and the auditor's performance. As for the second trend, it falls in the context of the difference between the perception held by the public about the implications of the auditors' responsibilities in reality (Koh & Woo, 1998).

In the context of the auditor's responsibilities, the public believes that auditors verify every transaction, ensure the accuracy of the data, the solvency of the company, detect all fraud incidents, and certify the efficiency and adequacy of the company's internal control system, all when an unqualified report is issued to the auditor (McEnroe & Martens, 2001). Accordingly, the public does not distinguish between the tasks of the auditor in auditing financial statements (traditional tasks) related to expressing an opinion on the fairness of those data, which are based on auditing standards, whether local or international, and between the auditing tasks in the special assignment such as auditing an element of financial statements such as sales or Procurement, wages, or other things in which the auditor expresses a conclusion about their validity and which are based on auditing standards and not auditing. Accordingly, the basic tasks of the auditing profession are focused on improving the requirements for preparing the auditor's report. Specifically in the International Auditing Standard ISA700 and later ISA701 in 2015, and in Iraq, the Private Companies Law No. 21 of 1997 has specified the requirements for the Sense Monitor report. Determine in paragraph 36: the extent to which the company's accounts are sound for the accounting principles followed, the extent to which the financial statements express the reality of the financial position and the result of the company's activity, the extent to which these data conform to the provisions of the Companies Law and the reference to the violations of the provisions of this law or the company contract in a manner that affects its activities or Its financial position.

Going Concern of Enterprise

The auditor must inform about the company's ability to continue as the continuation of its operations in accordance with the International Auditing Standard (ISA700) and later (ISA701), so the International Auditing and Assurance Standards Board (IAASB) issued Standard (ISA 570), which deals with the auditor's responsibility in auditing financial statements and determining the extent of management's use to assume continuity in preparing and presenting the financial statements, as well as applying this assumption in the auditor's report. From an accounting point of view, and according to the conceptual framework of the International Accounting Standards Board (IASB) issued in 2018, the imposition of continuity means that the entity prepares its financial statements according to the assumption that it continues in its activity for the foreseeable future, and therefore there is no intention of the unit to fully liquidate its activity and there is no interruption in its commercial dealings (IASB Conceptual Framework, 2018:3.9). Moreover, the assets and liabilities of that entity are recorded on the basis that the unit will be able to realize the assets, fulfill its obligations and obtain financing (if necessary) in the normal course of its work (KPMJ, 2018: 14). So, the auditor must obtain sufficient and appropriate evidence to ensure that the conditions of uncertainty associated with the unit do not affect its continuity, but the important question is how is this accomplished by the auditor? And what is the reflection of this opinion on the auditor's report?

There are a number of studies that deal and examined the factors affecting the evaluation of the continuity of the entity. A lot of research focused on analyzing the elements of the financial statements and deriving financial ratios from them, while other studies investigated non-financial indicators, including market indicators, strategic initiatives of the entity, sector indicators, and ownership structure. Bava &Trana (2018) conducted a survey of research that addressed these indicators and concluded that the most important of them are:

1. Liquidity indicators (included current ratio, fast ratio , cash from operations/total liabilities etc.)
2. Leverage indicators included debt ratio, long term liabilities/ total assets, total liabilities/total assets etc.)
3. Profitability ratios (included ROA, Retained earnings/ total assets, EBIT/total assets etc.)

Auditors should reinforce these indicators with additional measures, including management plans for future actions, to know to what extent the results of these plans are likely to improve the company's position (ISA570: paragraph 16). However, these indicators, whether financial or non-financial, face them many obstacles, including:

1. The degree of conservatism with which the financial statements are prepared. Accounting standards are characterized by a high degree of flexibility that allows the exercise of different levels of conservatism in accounting measurement. Therefore, the indicators resulting from those financial statements are also different.
2. The practice of management as preparing the financial statements for earning management techniques by exploiting the flexibility and gaps in accounting standards, and then formulating overly optimistic expectations in order to avoid the risks of being replaced by the principles.
3. Administrations resort to short or long-term borrowing methods to avoid the liquidity problem and improve its ratio, but this exposes the facility to multiple risks (internal and external)
4. The difference in the degree of influence of the factors affecting the continuity of the entity from one sector to another. This opinion confirmed that the banking sector was affected more than other economic sectors during the global financial crisis on 2008.

To avoid all these obstacles, auditors had to investigate beyond these financial and non-financial indicators through a huge amount of data to reach an appropriate opinion on the continuity of the facility, and perhaps one of the appropriate investigation tools is data mining. Several researches have been conducted in the applied field. Fezeh & Mahdi (2013) tested this going concern on 146 Iranian companies listed in the capital market, including 73 companies exposed to bankrupt, with companies still continuing to list for the years 2001-2011 and the researchers used 42 variables, the data were entered in the Stepwise Discriminate Analysis test to arrive at the variables that distinguish continuous companies from non-continuous companies. There are four ratios that distinguish between them, which are the ratio of total liabilities / total assets, operating income / sales, net income / total assets, retained earnings / total assets. Angel et al. (2018) used a sample of 252 hotels in Spain for the period from 2000-2014. The sample was divided into 126 hotels. The auditors gave an opinion of their ability to persistence and the other 126 not gave opinion of their reports. For testing purpose the researches use 12 ratios financial and Non-financial for the purpose of identifying a set of explanation variables that can predict conditions of uncertainty in the hotel industry, in order to obtain a higher prediction level and fewer variables the empirical study showed that the method Ada Boost used from Boosting algorithms to nearly 100% accuracy (Han et al., 2011).

Detection of Fraud

Wang & Yang (2014: 5-6) believes that there are two main areas for using data mining techniques to discover and identify fraud: Outlier Analysis & Benford's Law analysis. They identified the areas that were discovered through these technologies, including the discovery that the seller was issuing fraudulent invoices on a regular basis and on a serial basis that the seller had one customer, the discovery of illegal payments to members of the families of government officials, and the discovery of a senior executive issuing bills to a fake company at his home address. The private research identified different types of data mining software that auditors can use. The field of fraud detection is expanding into forensic accounting as a new professional area for accountants and auditors.

Tools of Data Mining

Classification

Classification is a common task in a human activities involving decision-making or prediction in an unknown or future situation, using the information currently available. As another point of view, classification is the process of creating a model (or function) that describes and distinguishes different data categories, for the purpose of being able to use the model to predict the category of unknown objects later. The derivative model is based on a set of training data (which is previously classified).

Decision Tree

The Decision Tree is an exploratory model that appears in the shape of a tree as its name is expressed, and precisely each of its branches represents a taxonomic question and its papers represent parts of the database that belong to the classifications that have been built. Decision tree algorithms are frequently used in Artificial Intelligence due to its many advantages in this area. What distinguishes the decision tree algorithm from other algorithms is that they are quick to learn and are often accurate in solving a wide range of problems and do not require any special preparation of your data. It is easy to build, inexpensive, gives more accuracy compared to other algorithms, the speed in the classification of records and easy to read. One of its disadvantages is that, it is biased to the qualities of the highest levels when data with class variables are used at different levels and sometimes the calculations are very complex, especially when many values are uncertain or when many outputs are linked.

Algorithm steps:

$$Info(D) = - \sum_{i=1}^m P_i \log_2(p_i) \dots\dots\dots (7)$$

$$P_i = \frac{|C_{i,D}|}{|D|} \dots\dots\dots (8)$$

Info(D): is just the average amount of information needed to identify the class label of a tuple in *D*.

$$Info_A(D) = - \sum_{j=1}^m \frac{|D_j|}{|D|} \times Info(D_j) \dots\dots\dots (9)$$

$\frac{|D_j|}{|D|}$: acts as the weight of the *j*th partition.

$Info_A(D)$: is the expected information required to classify a tuple from *D* based on the partitioning by *A*.

Information Gain: is defined as the difference between the original information requirement (i.e., based on just the proportion of classes) and the new requirement (i.e., obtained after partitioning on *A*). That is,

$$Gain(A) = Info(D) - Info_A(D) \dots\dots\dots (10)$$

In other words, Gain tells us how much information we can gain by branching out on *A*.

ID3 Algorithm

It is one of the classification algorithms where the entry of the algorithm set of data and the result is a classifier rule through which you can classify new data that has not been used before as a kind of prediction for this new data, this work is in the form of a tree structure so it is called decision tree and can be converted into a set of rules so it is also called decision rules. It creates understandable predictive rules of training data, which are fast-building and need only to test the attributes directly related to the problem in order to classify all data. One of its disadvantages is that it does not work efficiently with small samples and only one attribute is tested at the time of decision.

Algorithm Steps

The Gini index measures the impurity of D , a data partition or set of training tuples, as

$$Gini(D) = 1 - \sum_{i=1}^m p_i^2 \dots\dots\dots(14)$$

When considering a binary split, we compute a weighted sum of the impurity of each resulting partition. For example, if a binary split on A partitions D into D_1 and D_2 , the gini index of D given that partitioning is

$$Gini_A(D) = \frac{|D_1|}{|D|} Gini(D_1) + \frac{|D_2|}{|D|} Gini(D_2) \dots\dots\dots(15)$$

The reduction in impurity that would be incurred by a binary split on a discrete- or continuous-valued attribute A is

$$\Delta Gini(A) = Gini(D) - Gini_A(D) \dots\dots\dots(16)$$

Naive Bayes

It is a powerful algorithm for supervised learning. Naïve Bayes is an extension of Bayes theory. In this work, we calculate the probability that each factor contributed. It is easy and fast and needs less data to train, when the properties are not linked (Independent) it performs better than other algorithms. It works better with categorical data than with numerical data. One of its drawbacks is the hypothesis of independent properties, often there is a correlation between properties affecting each other and it doesn't work well with digital data.

Algorithm Steps

Let X is a tuple and suppose that there are m classes, C_1, C_2, \dots, C_m . Given a tuple, X , the classifier will predict that X belongs to the class having the highest posterior probability, conditioned on X . That is, the naïve Bayesian classifier predicts that tuple X belongs to the class C_i if and only if:

$$P(C_i|X) > P(C_j|X) \text{ for } 1 \leq j \leq m; j \neq i$$

$P(C_i|X)$: is the posterior probability to class i .

$P(C_j|X)$: is the posterior probability to class j .

C_i : the classes

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \dots\dots\dots(11)$$

$$P(X|C_i) = \prod_{k=1}^n P(X_k|C_i) \dots\dots\dots(12)$$

$$= P(X_1|C_i) \times P(X_2|C_i) \times \dots \times P(X_n|C_i) \dots\dots(13)$$

In order to predict the class label of X, $P(X|C_i) \cdot P(C_i)$ is evaluated for each class C_i . The classifier predicts that the class label of tuple X is the class C_i if and only if $P(X|C_i)P(C_i) > P(X|C_j)P(C_j)$ for $1 \leq j \leq m, j \neq i$

Selected Data and Analysis

The classification techniques were trained and evaluated by training a samples. Samples which are selected consist of 13 observations that represent banks with 6 going concern banks and 7 not going concern banks. We adopted 27 indicator and ratio used to classify and measure going concern of that banks. The data was gathered through the information database system of the Iraqi Stock Exchange and Central Bank of Iraq, a list of ratio sand indicators shown in Table 1.

Table 1 A SAMPLE OF RATIOS AND INDICATORS		
1	CRR	Current Rate Ratio
2	LQR	Liquidity Quick Ratio
3	COR	Current Operation Ratio
4	DR	Debt Ratio
5	LTLR	Liabilities Total Long Term
6	TLR	Total Long Term
7	ROA	Return on Assets
8	ROE	Return on Equity
9	ORR	Operation Revenue Ratio
11	WCR	Working Capital Ratio
12	ER	Equity Ratio
13	NPR	Net Profit Ratio
14	EDR	Equity Debt Ratio
15	CIR	Cash Investment Ratio
16	TRR	Total Revenue Ratio
17	NCF	Net Cash Flow
19	IR	Interest Revenue
20	CFR	Cash From Finance
21	RFC	
22	IRR	Interest Revenue Ratio
23	NPRR	Net Profit Ratio
24	EM	Equity Multiplier
27	IIR	Interest Income Ratio

Based on the sample data, the objective is to construct a prediction model for the going concern status of banks based on the twenty seven financial ratios and indicators .To derive the going concern prediction model, the sample data is partitioned into the following two data sets: a training sample include approximately 70 percent of the original samples and a validation or testing sample include the remaining 30 percent of the original samples.

Decision Tree Algorithm

This study used Decision Tree Algorithm to build the prediction model of going-concern status, it is an appropriate method to use in this situation. The decision tree algorithm results are summarized in Figure 2. As can be seen, it is a good model that indicate a good fit. The precision of positive cases is 100% and the recall of positive cases is 100% , when the value of DR ratio is greater than 0.129 the banks is ongoing and there is no problem and when value of DR and LTLR ratios is less than or equal to 0.129 and greater than 0.081 respectively , the bank is ongoing ,and when value of DR and LTLR ratios is less than or equal to 0.129 and less than or equal to 0.081 respectively , the bank is out going So the DR and LTLR ratios are very important in classification and prediction of going concern as depicted in figures 1 and 2.

accuracy: 100.00%			
	true yes	true no	class precision
pred yes	6	0	100.00%
pred no	0	7	100.00%
class recall	100.00%	100.00%	

FIGURE 1
THE CONFUSION MATRIX OF DECISION TREE ALGORITHM

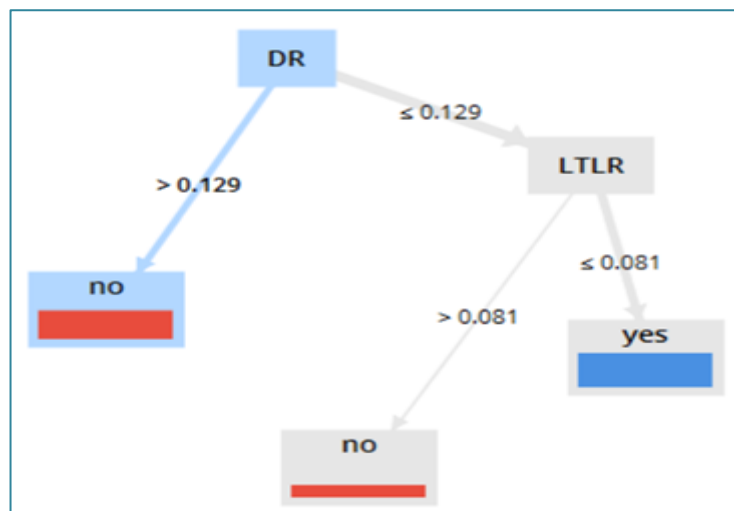


FIGURE 2
THE DECISION TREE DIAGRAM

ID3 Algorithm

The ID3 algorithm results are summarized in Figure 1. As can be seen, it is a good model that indicate a good fit. The precision of positive cases is 100% and the recall of positive cases is 100%, and classification of going concern of each bank is depicted in figures 3 and 4.

accuracy: 100.00%			
	true yes	true no	class precision
pred. yes	6	0	100.00%
pred. no	0	7	100.00%
class recall	100.00%	100.00%	

FIGURE 3
THE CONFUSION MATRIX OF ID3 ALGORITHM

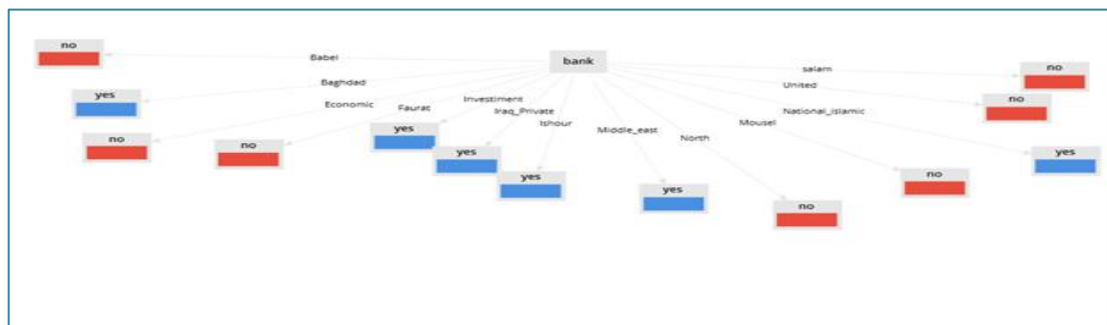


FIGURE 4
THE ID3 DIAGRAM

Naïve Bayes Algorithm

The Naïve Bayes algorithm results are summarized in Figure 1. As can be seen, the accuracy of model is 92.31%. The precision of positive cases is 100% and the recall of positive cases is 83.33%, and classification of going concern of each bank is depicted in figures 5 and 6.

accuracy: 92.31%			
	true yes	true no	class precision
pred. yes	5	0	100.00%
pred. no	1	7	87.50%
class recall	83.33%	100.00%	

FIGURE 5
THE CONFUSION MATRIX OF ID3 ALGORITHM

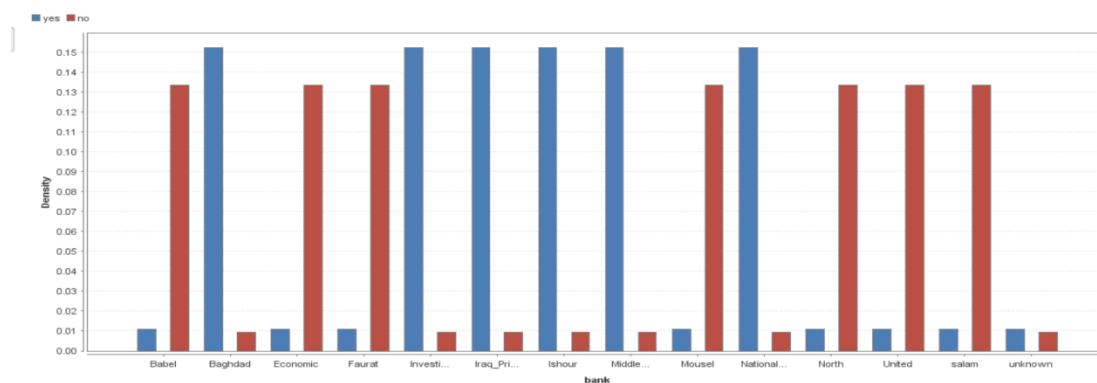


FIGURE 6
THE CONFUSION MATRIX OF ID3 ALGORITHM

Random Forestalgorithm

The random forest algorithm results are summarized in Figure 1. As can be seen, it is a good model that indicate a good fit. The accuracy of classifier is 92.31%, The precision of positive cases is 100% and the recall of positive cases is 83.33%, when the value of ROA ratio is less than or equal to 0.007 the banks is not ongoing and ROA ratios is greater than 0.039 the bank is not ongoing, and when value of ROA and LTLR ratios is less than or equal to 0.1039 and greater than to 0.081 respectively, the bank is not ongoing, So the DR, LTLR and EDR ratios are very important in classification and prediction of going concern as depicted in figures 7-9.

accuracy: 92.31%			
	true yes	true no	class precision
pred. yes	5	0	100.00%
pred. no	1	7	87.50%
class recall	83.33%	100.00%	

FIGURE 7
THE CONFUSION MATRIX OF ID3 ALGORITHM

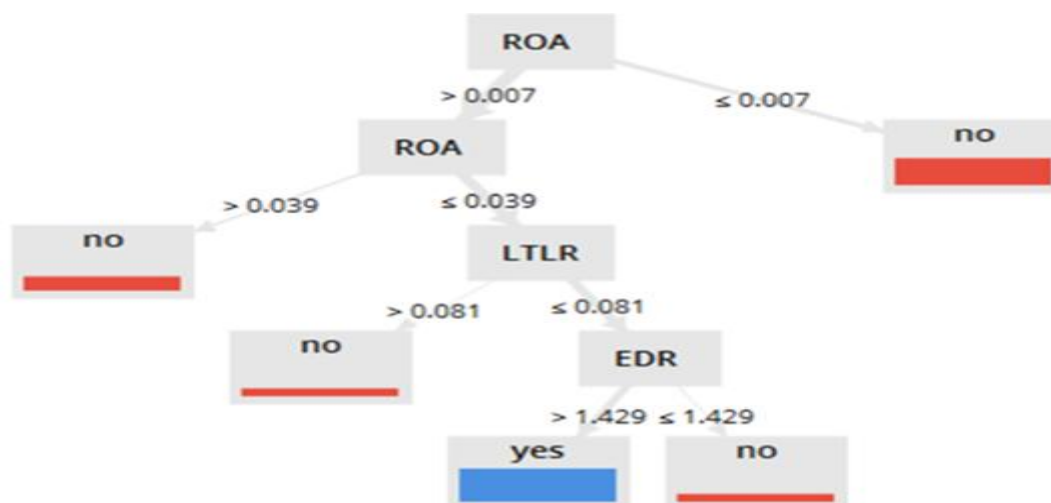


FIGURE 8
THE CONFUSION MATRIX OF ID3 ALGORITHM

```

ROA > 0.007
| ROA > 0.039: no {yes=0, no=2}
| ROA ≤ 0.039
| | LTLR > 0.081: no {yes=0, no=1}
| | LTLR ≤ 0.081
| | | EDR > 1.429: yes {yes=5, no=0}
| | | EDR ≤ 1.429: no {yes=0, no=1}
ROA ≤ 0.007: no {yes=0, no=4}

```

FIGURE 9
CLASSIFICATION AND PREDICTION

CONCLUSIONS

Recently data mining has gained widespread attention and increasing popularity in the business world. Many algorithms used for classification purposes and there was grown in usage and effectiveness. The main purpose of the paper is to predict company's going concern by using different algorithms and know the important ratios and indicators depending on real data. The results of the analysis show that the ratios DR, LTLR and ROA and EDR predicts bank's situation more accurately.

REFERENCES

- Bava, F., & Gromis di Trana, M. (2019). [ISA 570: Italian auditors' and academics' perceptions of the going concern opinion](#). *Australian Accounting Review*, 29(1), 112-123.
- Chen, Y.M., & Chiu, Y.P. (2014). [Business Intelligence and Analytics to Prediction of Going Concern using NeuroFuzzy Approach](#). In *The Asian Conference on Technology, Information & Society Official Conference Proceedings*.
- Enroe, J.E., & Martens, S.C. (2001). [Auditors & Investors' Perceptions of the Expectation Gap](#). *Accounting Horizons*, 15(4), 345-358.
- Fernández, M.Á., Serrano, J.R.S., Aguilera, D.A., & Casado, G. (2018). [Predicting going concern opinion for hotel industry using classifiers combination](#). *International Journal of Scientific Management and Tourism*, 4(1), 91-106.
- Han, J., Pei, J., & Kamber, M. (2011). [Data mining: concepts and techniques](#). Elsevier.
- IASB Conceptual Framework (2018). For Financial Reporting 2018, incp.org.co/sit/publication.

- Koh, H.C., & Woo, E. (1998). [The Expectation Gap in Auditing](#). *Managerial Auditing Journal*, 13(3), 147-154.
- KPMJ, (2018). Accounting & Auditing Update, Issue No.22, an Indian Registered Partnership.
- Mahdi, S., & Fezeh, F. (2013). [Data mining Approach to Predication of going Concern Using Classification & Regression Tree](#). *Global Journal of Managements & Business Research Accounting & Auditing*, 13(3).
- Mojtaba, S.S., & Fahime, I. (2012). [Finding Rules for Audit Opinion Prediction Through Data Mining Methods](#). *European Online Journal of Natural & Social Science*, 1(2), 28-36.
- Rostami, K. H., Omrani, H., Margavi, A. K., Asadzadeh, H., & Nazari, H. (2011). [Data Mining and Application in Accounting and Auditing](#). *Journal of Education and Vocational Research*, 2(6), 211-215.
- Sahu, H., Shrma, S., & Gandha, L. (2012). [A brief Overview on Data Mining Survey](#). *IJCTEE*, 1(3), 114-121.
- Wang, J., & Yang, J. (2014). [Data mining techniques for auditing attest function & fraud detection](#). *Journal of Forensic & Investigative Accounting*, 1(1), 1-9.