PREDICTING GOING CONCERN OPINION FOR HOTEL INDUSTRY USING CLASSIFIERS COMBINATION

Manuel Ángel Fernández[1]

José Ramón Sánchez

David Alaminos

Gisela Casado

Abstract:

The prediction of qualifications for going concern has been the focus of attention of the accounting and financial research with the purpose of creating models that help auditors to assess the normal course of business. Previous studies about going concern opinion prediction have been developed exclusively for manufacture and financial companies. However, there are no previous experiences of companies from the hotel industry. In the last decades, hotel industry has become one of the largest industries with the greatest expansion in the world, and this industry has its own features that we should pay special attention to. This paper provides an exclusive going concern prediction model for the hotel industry using computational methods of variable selection and classifiers combination. According to the results, companies that hold a high proportion of current assets, low return on asset margin, high leverage ratio, low current ratio, possess establishments that have a non-vacational style and don't belong to a chain, are more likely to get a going concern opinion. The document offers a view of the challenges faced by auditors in the hotel industry and how the implementation of a proper model to foresee opinions of going concern can help auditors to cope with these challenges.

Keywords: Classifiers combination, Going concern opinion, prediction, Computational methods of variable selection, Hotel industry

JEL Classification Codes: C63, M42, L83

[1] Department of Finance and Accounting, University of Málaga, Málaga, Spain

1. INTRODUCTION

The going concern accounting principle is one of the most important when preparing financial statements and a great part of the financial information is based on the hypothesis that the company will continue with the normal development of its activity in the future. The current accounting rules demand that the auditor should assess the ability of a company to continue as a going concern. These assessments are useful to foresee an eventual bankruptcy and provide explanation about it (Chen and Church, 1992). Owing to this, the prediction of qualifications for going concern has been the focus of attention of the accounting and financial research with the purpose of creating models that help auditors to assess the normal course of business.

Previous studies about businesses' continuity predictions have been developed exclusively for manufacturing and financial companies. Therefore, there are no previous experiences of companies from the hospitality industry. However, in the last decades, tourism has been continually growing and it has become one of the largest sectors with the greatest expansion in the world and this is best represented by hotels. In 2015, tourism’s contribution to world GDP rose for the sixth year in a row reaching a total of 9.8 % (WTTC, 2015).

Hotel auditing faces a unique activity in many ways and all the knowledge that we can obtain about other industries’ work would be very useful for us to understand how the hotel industry works. The hotel sector has its own features that we should pay special attention to (Gémar et al, 2016). A good example of this is the inability to increase the amount of accommodation in a short time frame despite a growing demand added to the high fixed costs that make it difficult to adapt the costs structure to occupancy fluctuation (Mattimoe, 2000). In addition, the literature points out that the hospitality industry is very sensitive to business cycles (Chen and Yeh, 2012) and that hotels go bankrupt for financial reasons different from those of other sectors (Youn and Gu, 2010; Fernández-Gámez, Cisneros-Ruíz and Callejón-Gil, 2016). Taking this into account, it is essential that the auditor has work tools specially designed for the hotel industry that includes management and operations aspects of the company being audited. The use of these tools to predict going concern audit opinions would complete and improve the auditing process and it would let the auditors be more confident when issuing their reports. As previously stated, tourist industry has been increasingly contributing to world GDP, and therefore it is of ever growing importance that professionals are provided with better and more sophisticated tools.

The present study intends to shed light on the research regarding going concern prediction by developing an exclusive model for the hotel industry. With this objective, a sample of 252 hotels has been used from which we obtained both financial and non-financial information for the period 2000-2014. In order to obtain a high level of prediction accuracy, computational methods for the selection of variables and for classifier combination have been applied to the data. Such methods have not been used before for going concern predictions but the conclusions of the previous research point out that classifier combination of artificial intelligence considerably overcomes the performance of individual classifiers when the act/operate in a scenery where multiple and independent data turn up (Kittler et al., 1998). Artificial intelligence techniques can be an alternative to solve classification problems, then they are more liable to make predictions as other statistical and conventional methods used to predict the firms’ continuity (Martens et al., 2008; Koh and Low, 2004). Nevertheless, although artificial intelligence classifiers have reached a high prediction level it would be necessary to carry out further studies to define the eventual best model to predict continuity opinions (Yeh, Chi and Li, 2014). Considering this fact, the present study develops a model to predict continuity opinions for tourist companies by using classifier combination methods. This method has been indeed successfully used both in the economy and other scientific fields such as value predictions of stock markets (Kin, Min and Han, 2006; Tsai and Hsiao, 2010), financial distress (Sun and Li, 2009; Alfaro, Gámez and García, 2008), temporary series prediction (Inoue and Kilian, 2005) and for phoneme sets and satellite images (Kuncheva, Bezdek and Duin, 2001). The results obtained let us determine which factors are better going concern predictors for hotel business as well as a prediction model with an accuracy rate higher than 99%.

The rest of this study has been organized as follows: section 2, we carry out a review of the literature about predictions of audit opinion qualified as going concern; section 3, we present methodological aspects used in the research; section 4, we detail data and variables that have been used; section 5, we analyse the empirical results obtained. Finally, we show the conclusions of this study and their implications.

2. LITERATURE REVIEW

The International Auditing Standards, according to the precepts proposed by IAASB (issued by the IFAC), include the NIA 570 relating to the auditor's responsibility in the auditing of financial statements related to going concern. This rule proposes that the audit opinions must evaluate if the going concern hypothesis is suitable for the elaboration of financial statements. In addition, these opinions should identify any uncertainty that could cause any doubt to the firm's continuity.

The importance of evaluating the continuity of the companies has been a concern not only for legislators but also for the scientific investigation. Since McKee's seminal work (1976) after the publication of the Statement of Auditing Standards n.2 as the first auditing rule that detailed specific considerations in terms of audit opinions about the going concern principle, several works have been published ranging from international legislative harmonization (Martin, 2000; Cordos and Fülöp, 2015), the importance of the audit opinions quality (Myers et al., 2014; Mo et al., 2015; Matsumura, Subramanyam and Tucker, 1997); the going concern opinion's effects (Citron et al., 2008; O'Reilly, 2010); to the development of certain models to predict audit opinions about going concern (Bellovary, Giacomino and Akers, 2007; Koh and Low, 2004).

In the beginning, the works made to predict going concern opinions reached an accuracy slightly higher than 80% by using multiple and discriminatory analysis (McKee, 1976; Mutchler, 1985). Later on, with the objective of improving the results of the initial models they used Probit analysis (Dopuch et al., 1987; Koh and Brown, 1991) and logistic regression (Menon and Schwartz, 1987; Cornier, Magnan and Morard; 1995; Mutchler, Hopwood and McKeown, 1997; Gaeremynck and Willekens, 2003). Together with statistic techniques previously mentioned, other studies were made in which they applied neural networks computational techniques (Klersey and Dugan, 1995; Koh and Tan, 1999; Anandarajan and Anandarajan, 1999). Since 2000 the use of computational techniques has been extended with the works of Lenard, Alam and Madey (2001), who propose a hybrid system of discriminating analysis and neural networks from which they got an accuracy of 92%. Koh and Low (2004) compare the usefulness of such neural networks, decision trees and logistic regression in predicting a company's going concern status. Their classification results indicate the potential of data mining techniques in a going concern prediction context. Meanwhile, Martew et al. (2008) used data mining for ant colonies optimization (AntMiner+) and other techniques such as support vectors machines or decision trees. Rough sets are one of the latest techniques being used (Yeh et al., 2014).

Based on the previous literature, which has focused on the audit opinion predictions about going concern, we can deduce that certain variables have turned out to be the best predictors. On one hand, there are financial variables that refer to liquidity, indebtedness, profitability and activity (Yeh et al., 2014; Hung and Shih, 2009; Martens et al., 2008). On the other hand, non-financial variables such as the effect of the corporate governance as a matter of financial information risk, the intellectual capital (Yeh et al., 2014), the company's size (Anandarajan and Anandarajan, 1999) and the kind of auditor (Yeh et al., 2014; Martens et al., 2008; Ireland, 2003).

3. RESEARCH METHODOLOGY

Going concern prediction models try to solve a classification problem by using an explicative set of variables to estimate dependent and binary variables because it reflects two possible situations: companies that have received going concern audit opinions and those who haven't. Consequently, to solve this classification problem, at least two questions must be answered, these are: to find out which explicative variables have a higher quality of explanation and which classification method has the highest accuracy for the dataset.

In order to solve the first issue raised, and considering that data mining techniques have limitations to assess the effect of independent variables on the dependent ones (Koh and Low, 2004), this study makes use of six methods of selection of attributes available in the set of automatic learning algorithms for data mining works of Waikato University (WEKA). According to Hall and Holmes´ proposal (2003), this method selection included three algorithms that evaluate subsets of attributes, specifically two which are classified as Filters: Consistency Subset Eval and Classifier Subset Eval, that select and evaluate attributes in a different way from the learning algorithm; and a third one classified as Wrapper, Wrapper Subset Eval, that make use of some classifier to determine the desirability of a subset. In addition, another three algorithms classified as evaluators of individual attributes are applied together with the Ranker method, allowing the sorting of attributes according to their qualities: ChiSquaredAttributeEval, GainRatioAttributeEvaly and InfoGainAttributeEval. Finally, with the aim of verifying the effectiveness of the selection of variables, the decision tree C4.5 has been used in order to compare the classification accuracy before and after such variables selection (Hall, 2002).

As for the second issue raised regarding the election of the classification method, the present study has chosen methods of classifier combination. To this end, four of the most important individual classifiers within the supervised classification (Mitchell, 2006) have been selected: Naïve Bayes Classifier (NBC), Algoritmo C4.5 (C4.5), Multilayers Perceptron (MLP) and Support Vector Machine (SVM). Subsequently we could reach conclusions about their combination by using the Adaboost, Random Forest and Mayority Voting methods. The classifier combination can be justified upon its statistical, computational and representational advantages. Instead of using a single classifier, several classifiers can be considered and combined in such a way that the different advantages of each of them can be taken (Kuncheva, 2004).

Amongst the classifier combination methods, the stand out is the so called Bootstrap, which uses samples without replacement that are the same size as the original training dataset and are suitable for classifiers that are very sensitive to small changes in training data such as decision trees or artificial neural nets (Alfaro, Gámez and García, 2004). One of the most commonly used Bootstrap techniques is Boosting, which employs a basic classifier as a subroutine to produce a new classifier with a high accuracy level. It involves applying the classification method successively and focusing each time on the learning of different cases of the data training. Once the process has concluded, it results in a final classifier that combines basic classifiers and that offers high accuracy both with the training data and the testing data. One of the Boosting algorithm version most used is the so called AdaBoost (Freund and Schapire, 1996), which starts from the training set:

$$T_r = \{(X_1,Y_1), (X_2,Y_2), ..,(X_n,Y_n)\} \qquad [1]$$

Each observation X_i is assigned a weight $w_b(i)$ that equals l/n, in order to update such weights in each iteration. As a result, and based on the training data, we get a basic classifier $C_b(X_i)$ whose error gets calculated as shown in [2].

$$\sum_{i=1}^{N} w_b(i) f_b(i) \quad \text{meaning} \quad f_b(i) = \begin{cases} 0 \; si \; C_b(x_i) = y_i \\ 1 \; si \; C_b(x_i) \neq y_i \end{cases}$$

[2]

The constant variable α_b will be calculated based on the classifier's errors. This variable will be used to update the weights by increasing the weighting of the wrongly classified observations and by diminishing the weighting of the correctly classified. The process is carried out in all the iterations (from $b=1, 2, 3,..., B$) resulting in a final classifier as a lineal combination of the basic classifiers that are weighted by the constant variable α_b. In addition, as the number of iterations grows, the errors of the Adaboost combined classifier tend to zero (Freund and Schapire, 1996).

As a complement to Adaboost and to get robust conclusions, this study uses two methods with the objective of evaluating the accuracy on the classifier combination: Random Forest and Mayority Votin. If the main classifier used is a decision tree, the Random Forest involves all these techniques (Breiman, 2001). We will call Random Forest any classifier set where the basic classifier is a decision tree built from a vector of random numbers Θ_t, where Θ_t are independent and equally distributed variables and in addition, the result of the final classifier is obtained through non-weighted voting. As it happens in every classifier set, the aim of this kind of method is that the algorithm gets the proper randomness that is needed to maximize the independence of the trees while keeping at the same time a reasonable precision. In the Random Forest case, these qualities are measured all together and are called strength and correlation. The strength of the classifier set is defined as in [3].

$$S= E_{X,Y}\, mr(X,Y) \qquad [3]$$

where *mr(X,Y)* is the margin function of a Random Forest that in case of having two classes, it is defined as in [4].

$$mr(X, Y) = E_{\Theta}[c(X, \Theta) = Y] - E_{\Theta}[c(X, \Theta) \neq Y] = 2.E_{\Theta}[c(X; \Theta) = Y] - 1 \quad [4]$$

where $E_{\Theta}[c(X, \Theta) = Y]$ *is* the limit of the trees proportion c_t that, given a pattern x, classifies it correctly when T increases:

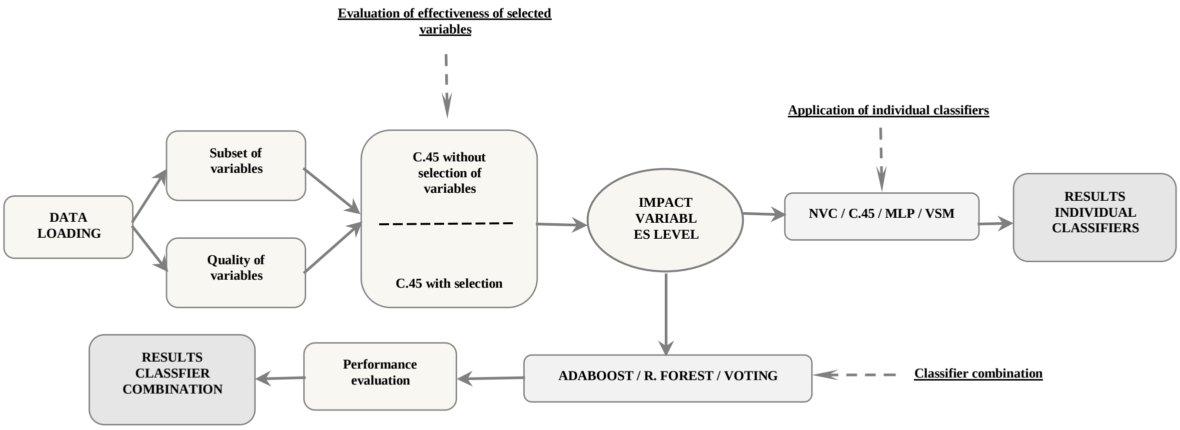
$$\frac{1}{T}\sum_{t=1}^{T} I(c_t(x, \Theta_t)) = y \rightarrow E_{\Theta}[c(X, \Theta) = Y] \quad [5]$$

Finally, the so-called Majority Voting (MV) is a technique related to a simple 'consensus' schema widely used in the literature. Its approach comes from the electoral context and has been used in different research works. Amongst them we should highlight Hazen et al. (2000), Lam and Suen (1997) and Battiti and Colla (1994). Supposing the starting labels of the individual methods to appear as binary c-dimensional vectors as the following … $[d_{i,1}, d_{i,2}, \ldots, d_{i,c}]^T \epsilon \{0,1\}^c$ *for* $i = 1, \ldots, L$, where $d_{i,j} = 1$ if the CL_i method labels the dataset as belonging to w_j or 0 if it doesn't. This way, the so-called plural vote becomes an overall decision for the class w_k if the condition given in [6] is fulfilled.

$$\sum_{i=1}^{L} d_{i,k} = {}_{j=1}^{c}max \sum_{i=1}^{L} d_{i,j} \quad [6]$$

Figure 1 presents a summary of the methodological process applied in this work. The process begins with the selection of attributes or explicative variables, considering the subsets of variables that have better and more reliable results. Then we apply the individual classifiers with the selected variables and as result, we obtain its classification precision that can be compared with the one obtained through classifier combinations.

Figure 1
Methodological process applied



4. DATA AND VARIABLES

The present investigation uses a sample of 252 Spanish hotels that have presented financial information for the period 2000-2014 and that appear in the SABI database of Bureau Van Dijk that provides fundamental data on around one million Spanish companies including profit and loss accounts, balance sheets and some additional data such as the date of foundation of the company, its status and its activity presented to the Commercial Registry. It is therefore a balanced sample that associates companies that have going concern opinion with those having a non-qualified audit opinion taking into account the year and the size (according to total assets), and all of them having been selected from those having code 70 in the Standard Industrial Classification (SIC). This matching criterion is consistent with those used in previous investigation about the selection of sampling methods for the prediction of going concern opinions (Martens et al., 2008; Levitan and Knoblett, 1985). The sample included 126 companies with going concern opinions and 126 companies without qualified opinions were also included at random.

The independent variable in our investigation is GCO that is equal to 1 if the auditor gave a going concern opinion and 0 if not. On the 21st December 2010, the Instituto de Contabilidad y Auditoría de Cuentas (Spanish accounting regulator) published a resolution that modified the previous Norma Técnica de Informes de Auditoría (Accounting Opinions Technique Rules). Since then, opinions given have had a new structure and content in terms of qualification classification. Before 2010 these situations of continuity problems were treated as uncertainties, which suggested a qualified opinion in case it was of a significant nature. The auditor, when encountering this situation, was faced with producing a report with qualifications, or even, taking into account the significance of the qualification or its accumulation with other qualifications, denying his opinion if he considered this to be the case for technical reasons. From 2010, these uncertainties about the possibility of keeping the firm in business as well as any kind of uncertainty haven't been included in an audit opinion as a qualification but as an emphasis paragraph and therefore without an effect on the audit opinion. With the objective of reaching a homogeneity when treating these incidents in the audit reports and not to break such homogeneity in the treatment of the information about the years under study, each audit report of the firms selected for the fiscal years 2010 onwards has been thoroughly analysed. In these years, the regulatory change takes place and those companies whose audit reports include an emphasis paragraph will be considered an exception due to doubts about keeping the firm in business and therefore without incidence on the audit opinion.

As independent variables, according to Koh and Low (2004) proposals, we selected a great number of variables used in previous researches, allowing the different models to select the most reliable variables through its internal algorithms. In addition, other typical variables of the hotel industry (Fernández and Becerra, 2015) were also selected. Table 1 shows the definition of the variables used. It represents a set of 16 variables that includes 8 financial ratios, 6 non-financial indicators and 2 control variables. It particularly includes liquidity variables, for they were considered to be determinant on the decision of reporting going concern audit opinions. (Anandarajan and Anandarajan, 1999; Yeh et al., 2014; Martens, 2008; Hung and Shih, 2009).

Profitability variables (Martens et al., 2008; Anandarajan and Anandarajan, 1999: Hung and Shih, 2009) and leverage variables (Martens et al., 2008; Anandarajan and Anandarajan, 1999: Hung and Shih, 2009) are also included.

To complete the selection of independent variables we have included other variables to capture the corporate governance effect as a risk factor of financial information (Hung and Shih, 2009; Beasley, 1996: Abbott, 2000; Wang and Deng, 2006). Besides, there are other 3 variables included that refer to intellectual capital, then it is accepted that it has an important role in terms of ensuring a higher financial benefit (Yeh et al., 2014). Likewise, two variables related to characteristics of tourism companies, one of them about the kind of establishment and the other one about the chain membership (Fernández and Becerra, 2015).

Finally, as control variables, we include the company size then it is proved that it is significantly associated with going concern decisions (Ireland, 2003; Anandarajan and Anandarajan, 1999) and the kind of auditor, then big auditing firms (Big 4) tend to issue more conservative reports (Yeh et al., 2014; Ireland, 2003; Martens et al., 2008).

Table 1
Independents variables

Categories	**Variable**	**Definition**
LIQUIDITY	CRR	Current Assets/Current Liabilities
	LQR	Cash + Cash Equivalents/Total Assets
	WCR	Working Capital/Total Assets
PROFITABILITY	ATR	Net Sales/Total Assets
	ROA	Earnings before Interest and Taxes//Total Assets
	RTR	Retained Earnings/Total Assets
LEVERAGE	DCR	Cash Flows/Total Debts
	LVR	Total Debt/Total Assets
CORPORATE GOVERNANCE	NMB	Number of Members on Board
INTELLECTUAL CAPITAL	IC1	Net Sales/Workforce
	IC2	EBIT/ Workforce
	IC3	R&D Expenses/Total Assets
HOTEL CHARACTERISTICS	TYH	1=Vacation, 2=Cultural, 3=Mixed
	CHM	1 if Hotel is Chain Membership=1, and 0 if not
CONTROL VARIABLES	SIZ	Log Total Assets
	AUD	1 if Auditor is Big 4 , and 0 if not

5. RESULTS

The main descriptive statistics of the selected variables for the present investigation are shown in table 2. Hotels with going concern opinion (GCO=1), compared with those not having it (GCO=0) are characterized by a higher liquidity average (LQR), a higher sales proportion in relation to total assets (ATR) and high leverage level (LVR). However, they present lower average values in terms of current ratio (CRR), working capital ratio (WCR), return on assets (ROA) and retained earnings (RTR). Table 3 shows the correlation between variables. According to the results obtained, it can be deduced that independent variables present a high correlation in respect to the dependent variable (GCO), excepting those referred to corporate governance and intellectual capital.

Table 2
Descriptive statistics

		CRR	LQR	WCR	ATR	ROA	RTR	DCR	LVR	NMB	IC1	IC2	IC3	SIZ
	Mean	1.059	0.407	-0.273	1.278	-0.084	-0.628	0.032	1.305	3.520	78.692	-0.405	0.018	10.032
GCO= 1	Median	0.651	0.239	-0.015	0.591	-0.076	-0.204	-0.012	0.830	3.000	74.811	-2.163	0.027	9.938
	SD	1.692	0.210	0.720	1.154	0.110	0.793	0.255	1.251	2.629	43.606	9.785	0.011	1.302
	N	126	126	126	126	126	126	126	126	126	126	126	126	126
	Mean	1.908	0.289	0.086	0.759	0.061	0.218	0.183	0.503	4.314	259.543	32.920	0.021	10.048
GCO= 0	Median	1.066	0.177	0.020	0.512	0.047	0.276	0.119	0.389	3.000	81.384	7.767	0.032	10.183
	SD	2.819	0.188	0.104	0.345	0.082	0.503	0.204	0.342	4.763	121.096	11.123	0.029	1.845
	N	126	126	126	126	126	126	126	126	126	126	126	126	126
t-test		2.540	-4.002	4.129	-4.065	3.996	4.708	2.481	-4.399	1.378	1.302	1,639	-0.284	1.894
p-value		0.008	0.000	0.000	0.000	0.000	0.000	0,022	0.000	0.129	0.143	0.076	0.735	0.118

Notes: Variable definitions: *CRR* is the ratio Current Assets/Current Liabilities; *LQR* is the ratio Cash + Cash Equivalents/Total Assets; *WCR* is the ratio Working Capital/Total Assets; *ATR* is the ratio Net Sales/Total Assets; *ROA* is the ratio Earnings before Interest and Taxes//Total Assets; *RTR* is the ratio Retained Earnings/Total Assets; *DCR* is the ratio Cash Flows/Total Debts; *LVR* is the ratio Total Debt/Total Assets; *NMB* is the Number of Members on Board; *IC1* is the ratio Net Sales/Workforce; *IC2* is the ratio EBIT/ Workforce; *IC3* is the ratio R&D Expenses/Total Assets; SIZ is the natural logarithm of the company's Total Assets.

able 3
Correlation matrix

	CRR	LQR	WCR	ATR	ROA	RTR	DCR	LVR	NMB	IC1	IC2	IC3	TYH	CHM	SIZ	AUD	GCO
CRR	1	0.254**	0.305**	-0.090	0.112	0.117	0.192**	-0.134**	0.028	-0.039	-0.005	-0.060	-0.028	-0.043	0.102	0.097	-0.163**
LQR		1	-0.137*	0.560**	-0.257**	-0.347**	-0.093	0.366**	0.013	-0.071	-0.050	0.151*	0.167*	0.148*	-0.357**	0.371**	0.235**
WCR			1	-0.339**	0.695**	0.792**	0.168**	-0.873**	0.040	-0.001	0.122	0.022	-0.187**	-0.131*	0.245**	-0.063	-0.240**
ATR				1	-0.271**	-0.280**	-0.093	0.293**	-0.019	-0.026	-0.069	0.229**	0.214**	0.087	-0.527**	0.171**	0.254**
ROA					1	0.644**	0.221**	-0.694**	0.024	0.025	0.226**	0.070	-0.158*	-0.124	0.191**	-0.187**	-0.263**
RTR						1	0.174**	-0.923**	0.037	0.028	0.114	0.063	-0.156*	0.056	0.348**	-0.290**	-0.291**
DCR							1	-0.163**	-0.034	0.007	0.079	0.013	-0.171*	0.008	0.089	-0.031	0.141*
LVR								1	-0.048	-0.015	-0.093	-0.024	0.159*	0.028	-0.264**	0.211**	-0.257**
NMB									1	-0.009	0.095	0.090	0.238**	0.177**	0.219**	0.229**	-0.090
IC1										1	0.581**	-0.013	0.035	0.021	0.026	-0.089	-0.090
IC2											1	0.008	-0.001	0.034	0.134*	-0.055	-0.108
IC3												1	0.062	-0.032	-0.111	-0.076	0.015
TYH													1	0.207**	-0.108	0.149*	0.234**
CHM														1	0.210**	0.160*	-0.138*
SIZ															1	-0.021	-0.167**
AUD																1	0.272**
GCO																	1

Notes: Variable definitions: *CRR* is the ratio Current Assets/Current Liabilities; *LQR* is the ratio Cash + Cash Equivalents/Total Assets; *WCR* is the ratio Working Capital/Total Assets; *ATR* is the ratio Net Sales/Total Assets; *ROA* is the ratio Earnings before Interest and Taxes//Total Assets; *RTR* is the ratio Retained Earnings/Total Assets; *DCR* is the ratio Cash Flows/Total Debts; *LVR* is the ratio Total Debt/Total Assets; *NMB* is the Number of Members on Board; *IC1* is the ratio Net Sales/Workforce; *IC2* is the ratio EBIT/ Workforce; *IC3* is the ratio R&D Expenses/Total Assets; *TYH* is a nominal variable equals to 1 if type of hotel is vacation, 2 if type of hotel is cultural and 3 if type of hotel is mixed; *CHM* is a dichotomous variable equals to 1 if a hotel belongs to a chain and 0 otherwise; *SIZ* is the natural logarithm of the hotel´s Total Assets: *AUD* is a dichotomous variable equals to 1 if auditor belongs to BIG 4 and 0 if not ; *GCO* is a dichotomous variable equals to 1 for going concern opinion and 0 if not.

According to the methodological process proposed, in the first place we proceeded to select the variables. As it was previously exposed, for this selection we have employed filter methods and methods based on models by using the C4.5 decision tree. Each evaluating method chose a subset of between 6 and 9 variables, as it is shown in table 4. For the selection of variables there have been also used three individual evaluators by using Ranker method, which provides a list of the ordered variables depending on their quality. In table 5 the results of each method are shown. Some variables coincided in the selection made by most of methods. In this sense, we can observe that the three evaluating methods of individual variables (ChiSquared, GainRatio e InfoGain) provided equal selections with small exceptions.

Table 4
Subsets of variable selection

Consistency	Classifier	Wrapper
CRR	ROA	WCR
LQR	DCR	DCR
ROA	LVR	LVR
DCR	TYH	IC2
LVR	CHM	SIZ
TYH	AUD	AUD
CHM		
SIZ		
AUD		

Table 5
Evaluation of the quality of the variables

ChiSquared		GainRatio		InfoGain	
1	DCR	1	DCR	1	DCR
2	CRR	2	ROA	2	CRR
3	ROA	3	CRR	3	ROA
4	TYH	4	TYH	4	LQR
5	LQR	5	LQR	5	LVR
6	LVR	6	LVR	6	TYH
7	CHM	7	CHM	7	CHM
8	AUD	8	WCR	8	AUD
9	SIZ	9	SIZ	9	IC2
10	IC2	10	IC2	10	SIZ
11	WCR	11	ATR	11	WCR
12	ATR	12	AUD	12	ATR
13	IC1	13	IC1	13	IC1

14	RTR	14	IC3	14	IC3
15	IC3	15	RTR	15	NMB
16	NMB	16	NMB	16	RTR

In order to contrast the selection of variables made by the different methods, we submit the dataset together with the selection of variables made, to the classifier C4.5. It is considered that a selection method significantly improves or gets worse in terms of statistics, when the difference of the change ratio in respect to the original classificatory reaches +/- 1%, as Hall and Holmes (2003) propose. The percentage of errors in the classification due to the application of C4.5 allowed us to prove that before the selection of variables, the precision percentage of the data classification was 91.62%. In table 6 can be seen that the precision percentage increased in most cases if we apply C4.5 after the selection of variables and got fewer errors with the Wrapper (94.72%) and Classifier (93.98) methods.

Table 6
Accuracy after the selection of variables (%)

Consistency	**Classifier**	**Wrapper**	**ChiSquared**	**GainRatio**	**InfoGain**
92.87	93.98	94.72	92.41	93.08	92.57

Finally, figure 2 shows the variables that were eventually selected as more significant. This selection was made based on the results obtained by the different methods used. For this reason, it was considered the frequency of appearance of the different variables in every subset. Those subsets selected by the methods having lower percentage of errors will be considered more relevant. Then, they were ordered according to the frequency of selection and in the case of the evaluating method Ranker we considered those variables that had the first positions.

Figure 2
Selected variables

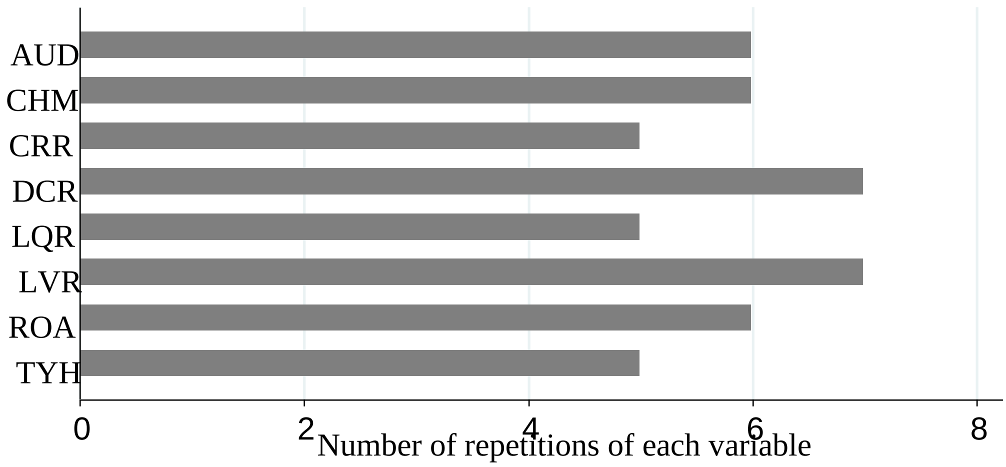


Table 7 shows the results that were obtained by applying individual classifiers once the selection of variables was made. The highest classification percentage was obtained with NBC, reaching 95.20%.

Table 7
Accuracy of individual classifiers (%)

NBC	C4.5	MLP	SVM
95.20	94.99	91.32	90.75

Finally, table 8 shows the results obtained by applying the different methods proposed of classifier combination. As expected, the classification precision improves in all the methods used.

Table 8
Accuracy of classifier combination (%)

Adaboost	Voting	Random Forest
99.22	97.81	95.30

In accordance with the results obtained we could firstly verify the evidence of a set of variables that can be considered to predict audit reports that are qualified as going concern for hotel industry. The set of variables selected shows that both financial and non-financial aspects are crucial for the problem under study. The financial aspects are those related to liquidity, profitability and leverage. Therefore, a high proportion of current assets (LQR), low return-on-assets (ROA), high leverage ratio (LVR) and low current ratio (CRR) increase the probability of going concern opinion. What is more, non-financial aspects of hotel industry having to do with the kind of establishment and the chain membership, then companies that develop a 'cultural' o 'mixed' style and don't belong to a chain have more probability of getting a going concern opinion. Finally, the kind of auditor (AUD) has also turned to be relevant for this study then it has been detected that big auditing companies are likely to issue going concern reports. These results coincide with those pointed out by Anandarajan and Anandarajan (1999), Yeh et al. (2014), Martens et al. (2008) y Hung and Shih (2009) related to financial variables. However, they differ from those obtained by Beasley (1996), Abbott (2000) y Wang and Deng (2006) for not having verified the importance of the corporate governance effect. Additionally, the variable that refers to the kind of auditor was also selected in previous studies (Anandarajan and Anandarajan, 1999; Yeh et al.2014; Martens et al., 2008).

Secondly, the results reached in this study prove that the accuracy of the prediction of the continuity of the hotels' businesses can be improved through the classifier combination regardless of the variables used. The algorithms version Boosting called AdaBoost has been the combination method most stable in terms of accuracy, improving significantly the results obtained in the previous literature about going concern prediction (Yeh et al. 2014; Bellovary, Giacomino and Akers, 2007; Kow and Low, 2004).

6. CONCLUSION

The going concern accounting principle is one of the most important concepts when preparing financial statements. Because of that, its importance in the prediction of audit opinions has been the center of attention for accounting and financial investigation over decades and has been used to design models to help auditors in the evaluation of the operation and continuity of businesses.

Previous studies about businesses' continuity predictions have been carried out exclusively for manufacturing and financial companies; on the contrary, there isn't anything similar for companies of the hospitality industry. For this reason, hotel auditing faces a unique activity in many ways and it is therefore important that the auditor has tools especially designed for the hotel industry that cover management aspects as well as the operations of the entity being audited. With the aim of covering this gap in the literature, this study tries to shed light on the research into going concern prediction by developing an exclusive method for the hotel industry.

According to the results of this study, we have identified the explicative set of variables that can predict going concern in the hotel industry. In order to get a higher prediction level and a smaller number of variables, the proposed procedure obtains a set of easy-to-understand decision rules that can help to interpret the auditing information and allow auditors to identify the most important variables for a system of control of opinions about continuity.

What is more, the empirical study carried out lets us verify that the classifier combination approach proposed is very efficient for classifying opinions about going concern in the hotel industry. The AdaBoost method derived from the Boosting algorithms has reached a classification accuracy of close to 100%.

The approach of some countries such as Spain to the tourist industry suggests a higher implication in the investigation tasks. Dealing with this gap from a scientific point of view that tries to improve the set of tests of the auditors who work for this sector leads to an investigation milestone from which future lines of investigation can be derived. For that reason, other applications can be considered to evaluate the proposed perspective in this study such as the use of samples related to different companies of the tourist sector such as restaurants and travel agencies. Also, the use of other explicative variables referring to the credit rating of the companies and to their reputations. It would be also interesting to explore the utilization of qualitative information about the different circumstances that provoke opinions about continuity.

REFERENCES

Abbott, L.J., 2000. The effects of audit committee activity and independence on corporate fraud. Managerial Finance 26, 55-67.

Alfaro, E., Gámez M., and García N., 2004. New classification ensemble methods. 57th International Atlantic Economic Conference, Lisboa.

Alfaro, E., Gámez, M., and García, N., 2008. Linear discriminant analysis versus adaboost for failure forecasting. Revista Española de Financiación y Contabilidad 137, (2008) 13-32.

Anandarajan, M., and Anandarajan, A., 1999. A comparison of machine learning techniques with a qualitative response model for auditors' going concern reporting. Expert Systems with Applications 16(4), 385-392.

Beasley, M.S., 1996. An empirical analysis of the relation between the board of director composition and financial statement fraud. The Accounting Review 71(4), 443-465.

Bellovary, J.L., Giacomino D.E., and Akers, M.D., 2007. A review of going concern prediction studies: 1976 to present. Journal of Business and Economic Research 5, 9-28.

Breiman, L., 2001. Random forests. Machine Learning 45, 5-32.

Chen, K.C.W., and Church, B.K., 1992. Default on debt obligations and the issuance of going-concern opinions. Auditing: A Journal of Practice and Theory 11(2), 30-50.

Chen, C.M., and Yeh, C.Y., 2012. The causality examination between demand uncertainty and hotel failure: a case study of international tourist hotels in Taiwan. International Journal of Hospitality Management 31(4), 1045-1049.

Citron, D., Taffler, R., and Jinn-Yang, U., 2008. Delays in reporting price-sensitive information: the case of going concern. Journal of Accounting and Public Policy 27, 19-37.

Fernández, M.A., and Becerra, R., 2015. An analysis of Spanish hotel efficiency. Cornell Hospitality Quarterly 56(3), 248-257.

Fernández-Gámez, M.A., Cisneros-Ruiz, A.J., and Callejón-Gil, A.M., 2016. Applying a probabilistic neural network to hotel bankruptcy prediction. Tourism and Management Studies 12 (1), 40-52.

Freund, Y., and Schapire, R.E., 1996. Experiments with a new boosting algorithm. Proceedings of the Thirteenth International Conference on Machine Learning, Morgan Kaufmann, 148-156.

Gémar, G., Moniche, L. and Morales, A.J., 2016. Survival analysis of the Spanish hotel industry. Tourism Management 54, 428-438.

Hall, M.A., and Homes, G., 2003. Benchmaking attribute selection techniques for discrete class data mining. IEEE Transactions on Knowledge and Data Engineering 16(6), 1437-1447.

Hung, Y.C., and Shih, Y.N., 2009. A prediction model of going-concern from the viewpoint of sustainable development. The 1st International Conference on Information Science and Engineering.

Inoue A., and Kilian, L., 2005. How useful is bagging in forecasting economic time series? A case study of U.S. CPI Inflation, NCSU and University of Michigan.

Ireland, J.C., 2003. An empirical investigation of determinants of audit reports in the UK, Journal of Business Finance and Accounting 30(7-8), 975-1015.

Kim, M.J., Min, S.H., and Han, I., 2006. An evolutionary approach to the combination of multiple classifiers to predict a stock price index. Expert Systems with Applications 31, 241-247.

Kittler, J., Hatef, M., Duin, R.P.W., and Matas, J., 1998. On Combining Classifiers. *IEEE* Transactions on Pattern Analysis and Machine Intelligence 20, 226-238.

Koh, H.C., and Low, C.K., 2004. Going concern prediction using data mining techniques. Managerial Auditing Journal 19(3), 462-476.

Kuncheva, L.I., 2004. Combining pattern classifiers. Methods and algorithms. John Wiley and Sons, New Jersey.

Kuncheva, L.I., Bezdek, J.C., and Duin, R.P.W., 2001. Decision templates for multiple classifier fusion: and experimental comparison. Pattern Recognition 34, 299-314.

Martens, D., Bruynseels, L., Baesens, B., Willekens, M., and Vanthienen, J., 2008. Predicting going concern opinion with data mining. Decision Support Systems 45, 765-777.

Matsumura, E.M., Subramanyam, K.R. and Tucker, R.R., 1997. Strategic auditor behavior and going-concern decisions. Journal of Business Finance and Accounting 24(6), 727-758.

Mattimoe, R.B., 2000. An institutional study of the factors impacting on room rate pricing strategy in the Irish hotel industry. The Sixth Inter-disciplinary perspectives on accounting conference, University of Manchester.

Sun, J., and Li, H., 2009. Financial distress prediction based on serial combination of multiple classifiers. Expert Systems with Applications 36, 8659-8666.

Tsai C.F., and Hsiao, Y.C., 2010. Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. Decision Support Systems 50, 258-269.

Wang, Z.J., and Deng, X.L., 2006. Corporate governance and financial distress. The Chinese Economy 39(5), 5-27.

WTTC., 2015. Travel and tourism economic impact 2015. World Travel and Tourism Council, London.

Yeh, C.C., Chi, D. J., and Lin, Y. R., 2014. Going-concern prediction using hybrid random forests and rough set approach. Information Sciences 254, 98-110.

Youn, H., and Gu, Z., 2010. Predicting Korean lodging firm failures: An artificial neural network model along with a logistic regression model. International Journal of Hospitality Management 29, 120–127.