sciendo
Review of Economic Perspectives – Národohospodářský obzor
Vol. 22, Issue 3, 2022, pp. 219–239, DOI: 10.2478/revecp-2022-0010

Best proxy to determine firm performance using financial ratios: A CHAID approach

Muhammad Yousaf[1] and Sandeep Kumar Dey[2]

Abstract: The main purpose of this study is to investigate the best predictor of firm performance among different proxies. A sample of 287 Czech firms was taken from automobile, construction, and manufacturing sectors. Panel data of the firms was acquired from the Albertina database for the time period from 2016 to 2020. Three different proxies of firm performance, return of assets (RoA), return of equity (RoE), and return of capital employed (RoCE) were used as dependent variables. Including three proxies of firm's performance, 16 financial ratios were measured based on the previous literature. A machine learning-based decision tree algorithm, Chi-squared Automatic Interaction Detector (CHAID), was deployed to gauge each proxy's efficacy and examine the best proxy of the firm performance. A partitioning rule of 70:30 was maintained, which implied that 70% of the dataset was used for training and the remaining 30% for testing. The results revealed that return on assets (RoA) was detected to be a robust proxy to predict financial performance among the targeted indicators. The results and the methodology will be useful for policy-makers, stakeholders, academics and managers to take strategic business decisions and forecast financial performance.

Keywords: Czech firms, Decision tree, financial ratios, firm performance, return on assets

JEL Classification: G00, L25

Received: 24 March 2022 / Accepted: 26 July 2022 / Sent for Publication: 9 September 2022

Introduction

One of the best methods to determine a firm's financial condition is by doing an analysis of financial ratios. Tracy (2012) argued that a financial ratio is executed by comparing two or more variables from the firm's financial statement for a given period of time. The analysis of financial ratios is used to determine the stability and accountability of a firm. It gives an idea to the tax bodies, government, stakeholders, and shareholders of the firm's financial performance. The analysis is important to determine a firm's financial performance and position. Hence financial ratios define the relationship between

[1] Faculty of Management and Economics, Tomas Bata University in Zlin, Mostni 5139, Zlin 76001, Czech Republic, e-mail: usaf880@yahoo.com
[2] Faculty of Management and Economics, Tomas Bata University in Zlin, Mostni 5139, Zlin 76001, Czech Republic, and Czech Mathematical Society, Prague, Czech Republic, e-mail: dey@utb.cz

accounting values, which is generally expressed mathematically (Larasati & Purwanto, 2022).

Financial ratios play an important role in evaluating and analysing the performance of a business, as these ratios forecast the firm's financial situation and operation performance (Zaini & Mahmuddin, 2019). Financial ratios could be useful for investors to observe whether a firm is worth enough to invest in and how is the trend in the future. Therefore, the analysis of financial ratios can help investors in making investment decisions and predicting a firm's future performance. The analysis might warn early if the firm's financial conditions slow down. The analysis is important for the firm's managers for the purpose of rewards. The analysis of the financial ratios is useful for economists, academics, and researchers to forecast potential future financial conditions. The analysis of the ratios is also beneficial for the stakeholders in making financing decisions.

Delen et al. (2013) used four decision tree algorithms to investigate the relationship between financial ratios and firm performance. However, it has been observed that the research by Delen et al. (2013) has many flaws. For example, the authors have utilized an inverted coincidence matrix to conclude their findings of critical parameters like True Positive and True Negative. These flaws encourage us to research with accurate calculations in the context of Czech firms. Abdel-Basset et al. (2020) argued that the findings and results should be precise; that's why the authors employed four different models by using financial ratios to reach accurate outcomes. In the current paper, in tandem with Chicco et al. (2021), we have taken cognizance of deploying the orthodox confusion matrix structure in conjunction with the Mathews Correlation Coefficient or the φ coefficient widely used in machine learning interventions to measure the quality of binary classifications. This procedure was induced to maintain accurate predictions of α error and β error.

The main objective of this study is to determine the best proxy of firm performance by using financial ratios. To the best of our knowledge, not much is known about the financial ratios among Czech firms. There are a few studies on financial ratios, but this study is different due to the following reasons.

- This is the first study about Czech firms using financial ratios and measuring the firm performance through the decision tree approach using secondary data from the Albertina database.
- This study provides empirical evidence about the financial ratios of the Czech economy, specifically in the automobile, construction, and manufacturing sectors. These sectors are preferred as the sectors are playing a significant role in the Czech economy. Moreover, it is logical and easy to compare the proxies of the firm performance by using the financial ratios of the sectors.

Following the introductory section of this research paper, section 2 discusses the literature review, section 3 presents the methodology, section 4 synthesizes empirical findings, and section 5 is about the conclusion, theoretical and practical implications, further research, and limitations of the study.

Literature Review

In the literature, there is no standard rule to select the financial ratios that are used to determine the firm performance as each firm practices different ratios to analyse performance. However, determining firm performance using different financial ratios has been a challenging and interesting problem for many managers and researchers. Different scholars used different financial ratios with different methodologies in the literature. For example, Habibi & Iqbal (2021) used 47, Yousaf & Bris (2021a) used 15, Zaini & Mahmuddin (2019) used 24, Valaskova et al. (2018) used 14, Hua et al. (2007) used 22, Wang & Chen (2006) used 11, and Bose (2006) used 24 financial ratios. Zaini & Mahmuddin (2019) concluded that there is no standard financial ratio analysis rule in the previous literature. This is also confirmed by Delen et al. (2013). Therefore, we used 16 financial ratios to analyze firm performance by the decision tree approach in the current study.

The analysis of financial ratios has become a more important topic in practice and academic research in the present coronavirus disease (COVID-19) situation. Many authors have employed financial ratios as bankruptcy risks. For example, Yousaf & Bris (2021a) examined the financial risks of Czech firms by using the stepwise regression technique. The study's results revealed that the current ratio, return on capital, current assets turnover rate and net working capital turnover rate positively influence the firm's financial health. On the other hand, asset turnover rate, return on capital employed, fixed assets turnover rate, debt to equity ratio, and inventory turnover rate have a negative impact on the firm's financial health. Valaskova et al. (2018) used profitability ratios, activity ratios, liquidity ratios, and indebtedness ratios to examine the financial risks in Slovak firms by using multiple regression techniques. In the same way, Zuardi (2021) used financial ratios to examine the financial health of Islamic banks.

Firm performance can be examined based on financial ratios as these ratios are convenient for managers, investors, and policy-makers to classify and predict future earnings and costs. Many authors discussed the importance of financial ratios to examine the firm's performance. For instance, Yalcin et al. (2012) divided financial ratios into eight main criteria to measure financial performance by implementing value-based financial performance measures and accounting-based financial performance measures of Turkish manufacturing industries. Rezaie et al. (2014) used four main financial ratios to measure the performance of Iranian cement firms. Abdel-Basset et al. (2020) employed four main financial ratios to measure the performance of the top ten steel Egyptian firms.

Firm performance is ongoing, essential, and an important topic in the literature and practice. Numerous authors used many proxies to measure firm performance. For example, Listiorini & Putri (2022); Yousaf et al. (2021); Dwilita & Mingka (2022); Kayani et al. (2020) have used return on assets (RoA); Lubis & Alfiyah (2021); Al-Zararee et al. (2021); Islamiyati & Diana (2021) used return on equity (RoE); Sharma et al. (2020); Senan et al. (2021); Gayathri & Vijayalakshmi (2021) used return on capital employed (RoCE) as a proxy to measure firm's performance. Some scholars employed two proxies (Yousaf, 2022a; Jyoti & Khanna, 2021; Bawazir et al., 2021; Yousaf & Bris, 2021b; Akgün & Karataş, 2020). Several authors have employed more than two proxies of the firm's performance (Almaskati, 2022; Bayraktaroglu et al., 2019; Soewarno & Tjahjadi, 2020; Yousaf, 2021; Xu & Li, 2020).

Numerous studies have contributed to the existing literature on firm performance using different proxies and methodologies. Most of the studies have employed the generalized method of moments (Yousaf & Bris, 2021b; Sharma et al., 2020; Kayani et al., 2020) and multiple regression techniques (Yousaf, 2021; Soewarno & Tjahjadi, 2020; Xu & Li, 2020). Only a few authors have used the decision tree approach (Delen et al., 2013; Abdel-Basset et al., 2020; Zaini & Mahmuddin, 2019). However, examining the best proxy of the firm performance in the prior literature is limited. Therefore, the current research will fill this gap and contribute to the literature and practice in many contexts.

Based on the previous literature, we used five following main financial ratios to measure the performance of the Czech firms.

- (i) Profitability ratios: Profitability ratios explore how firms use their current assets to make a profit.
- (ii) Liquidity ratios: These ratios present how fast can firms attain their commitments. The ratios also present a firm's capacity to meet its immediate requirement in the short run.
- (iii) Coverage ratios: Coverage ratios are used to estimate the firm ability to pay its financial responsibilities.
- (iv) Turnover ratios: Turnover ratios can be used for how a firm is using its assets to generate revenues. So, the ratios measure the value of a firm's revenues or sales relative to the value of its assets.
- (v) Leverage ratios: These ratios are also called debt ratios because they measure the ability of a firm to recover its long-term debt.

A detailed description of the above financial ratios is given in Table 1.

Methodology

Source of Data

In the current research, the data considered 287 Czech firms, which were chosen randomly from three sectors: construction, manufacturing, and automobile. The three sectors are selected in the present research as it is easy and logical to compare the financial ratios within the selected sectors instead of the banking, services, agriculture, or educational sectors. The secondary data was obtained from the Albertina database. Numerous scholars obtained data from the database, such as Vrbka et al. (2022); Blažková & Dvouletý (2022); Yousaf (2022b); Camska et al. (2021); Yousaf (2021); Kucera et al. (2021); Vrbka (2020); Pivoňková & Tepperová (2021); Yousaf et al. (2021). Chandrapala & Knapkova (2013) notified that Albertina data deals with the expansion, processing, delivering, and distribution of the databases. Činčalová & Hedija (2020) claimed that the Albertina database covers the data of more than 2.7 million subjects. So, the sample covered the time period of 2016-2020, and 1374 total observations.

Variables

Due to the availability of data from the Albertina database and based on the previous literature, three proxies are considered as dependent variables to measure the firm performance, RoA, RoE, and RoCE. Many financial ratios are used as independent variables. The complete list of the variables/financial ratios is given in Table 1.

Table 1: Variable description

Variables	Abbreviation	Measurements
Profitability Ratios		
Return on equity ratio	RoE	Net income / shareholder's equity
Return on assets ratio	RoA	Net income / total assets
Return on common equity	RoCE	Earnings before interest and taxes (EBIT) / capital employed
Liquidity Ratios		
Current ratio	CR	Current Assets / Current liabilities
Quick ratio	QR	(Current assets – inventory) / current liabilities
Operating cash flow ratio	OCF	Operating cash flow / current liabilities
Coverage Ratios		
Interest coverage ratio	ICR	EBIT / interest expense
Cash flow coverage ratio	CFC	Opertating cash flow / total debt
Turnover Ratios		
Receivable turnover rate	RTR	Sales / accounts receivables
Inventory turnover rate	ITR	Cost of sold goods / inventory
Asset turnover rate	ATR	Sales / total assets
Net working capital turnover rate	WCT	Sales / (current assets – current liabilities)
Fixed assets turnover rate	FAT	Sales / fixed assets
Current assets turnover rate	CAT	Sales / current assets
Leverage Ratios		
Debt ratio	DR	Total debt / total assets
Debt to equity ratio	DER	Total liabilities / equity

Source: Authors' calculations

Solution for Missing Data

To solve for the missing values in the dataset, we activated the Multivariate Imputation by Chained Equations (MICE) package (van Buuren & Groothuis-Oudshoorn, 2011) in R Studio version 4.0.3. The package generates multiple imputations (replacement values) for multivariate missing data. Fully conditioned specified separate models impute incomplete variables. The algorithm can impute various classes of continuous, binary, unordered categorical and ordered categorical data. Since our dependent variables (RoA, RoE, and RoCE) have been binary coded, it is recommended to deploy Logistic regression (logreg) with bootstrap for imputation of the null values, and predictive mean matching (pmm) for the regressors (Azur et al., 2011).

Exploratory Factor Analysis (EFA)

An exploratory analysis was conducted to validate the ratio variables' unrepresented dimensions. Furthermore, two tests (Kaiser–Meyer–Olkin Measure of Sampling Adequacy and Bartlett's test of sphericity) were executed to solve for data adequacy.

Table 2: KMO and Bartlett's Test of Sphericity

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.612
Bartlett's Test of Sphericity	Approx. Chi-Square	6595.128
	df	78
	Sig.	.000

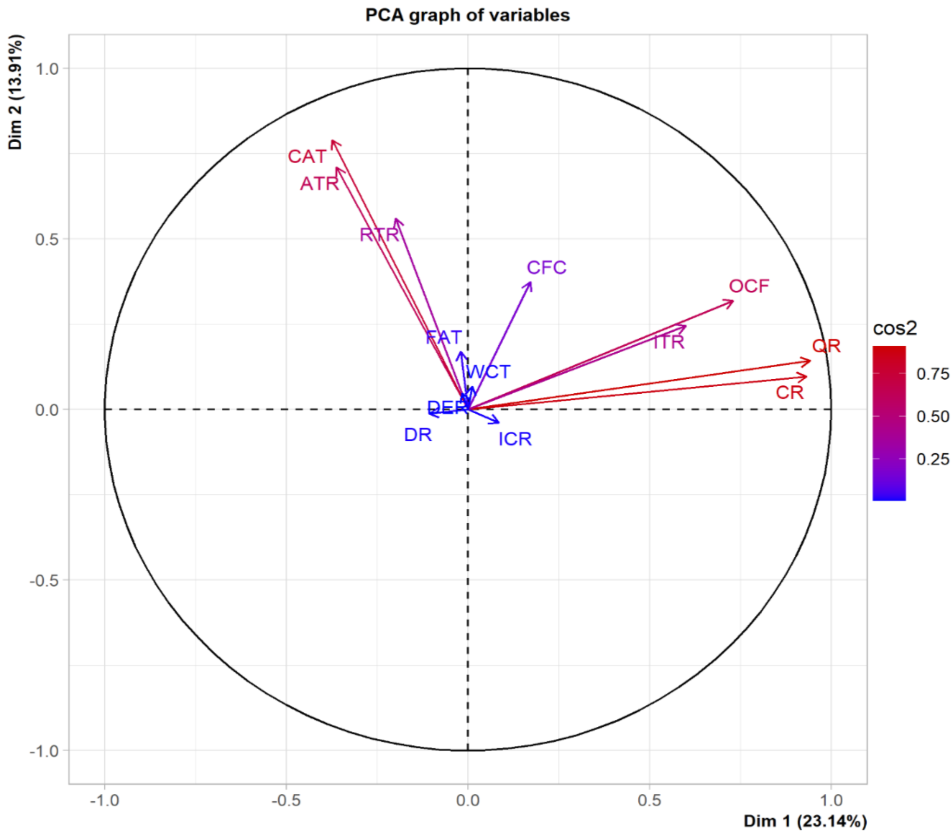
Source: Authors' processing from SPSS

The KMO measure in Table 2 describes the value being > 0.5, which is under the admissible range recommended by Kaiser (1974), giving the authors a signal of data adequacy for further analysis. Bartlett's test of sphericity is a check of the nature of the matrix being investigated. If an identity matrix is detected, it would imply that the variables in question are measuring different exogenous variables and thus, poorly correlated with each other; this is precisely contradictory to the essence of the entire EFA exercise. The matrix components need to have a certain degree of correlation for the investigation to proceed, as an experiment cannot be conducted on identity matrices. In this study, Bartlett's Test of Sphericity has achieved the desired Chi-squared threshold at $p<.001$ significance level (Table 2), thereby indicating correlation among the potential predictors (Field, 2005).

Principal Component Analysis (PCA) reduces the dimensions of a given data set in order to partition them into a smaller set of artificial variables known as principal components which represent the most variances in the original values (Nilashi et al., 2017). This procedure designates how a variable contributes to that component, whereas factor analysis establishes a mathematical model from which factors are estimated (Dunteman, 1989). When the PCA algorithm is executed on a dataset, a matrix structure representing the relationship among the variables is established. Linearity within the matrix is computed by determining the eigenvalues of the matrix. Eigenvectors indicate the loading of a particular variable on a particular factor (Field, 2005).

In concurrence with Delen et al. (2013), the study as a part of the exploratory factor analysis (EFA) procedure conducted PCA to optimize the research model. The authors executed PCA based on Varimax Rotation utilizing R Program's Factoshiny() (Lê et al., 2008). Figure 1 displays the PCA scatter graph on a 2-D vector space. In order to determine the quality of the contribution at the variable level, and to remove redundancy in the model the authors have deployed Cos^2 (Eigenvalue squared in geometric terms) values in order to retain the best fit variables that may considerably boost the explanatory capacity of the model (Abdi & Williams, 2010).

Figure 1: PCA Variable Graph



Source: Authors' Processing from R ver 4.0.3

In Figure 1, as the variables are expressed in the 2D vector space they undergo SVD (Singular Value Decomposition) which inflates the data matrix and creates artificial sets of variables; after the varimax orthogonal rotation has taken place, the correlation matrix, transforms into the covariance matrix thereby producing the best fit target variables (Lê et al., 2008). The squared cosine shows the importance of a component for a given observation (Harman, 1976).

The squared cosine is calculated as follows.

$$Cos_{i,l}^{2} = \frac{f_{i,l}^{2}}{\sum_{l} f_{i,l}^{2}} = \frac{f_{i,l}^{2}}{d_{i,g}^{2}}$$

Where $Cos^2_{i,\ g}$ contributes a relatively large portion to the total distance, these components are important for that observation. $d^2_{i,\ g}$ is the squared distance of a certain observation to the origin. According to the Pythagorean theorem, the squared distance, d^2

$_{i, g}$, is measured as the sum of the squared values of this observation's factor scores. The distance to the centre of gravity is defined for supplementary observations and the squared cosine can be calculated and is meaningful. Therefore, the value of Cos^2 can assist in extracting the components that are significant to understand both active and supplementary observations (Abdi & Beaton, 2021). According to Figure 1, the quality of the variance contribution can be comprehended according to the heat-map themed hues (warmer hues for higher contributions and lighter hues for lower contributions) and primarily, the cosine departure from the centre of gravity. On the basis of the above premise, the authors have extracted CAT, ATR, RTR, FAT, WCT, CFC, OCF, ITR, QR and CR in conformity with Johnson & Wichern (2013). Cos^2 values greater than 0.30 have been retained for this study, in accordance with Korenius et al. (2007). This implies that each predictor must possess at least 0.30 and above Cos^2 value in either dimension to be included in the study. The factors extracted are given in Table 3 with their respective values.

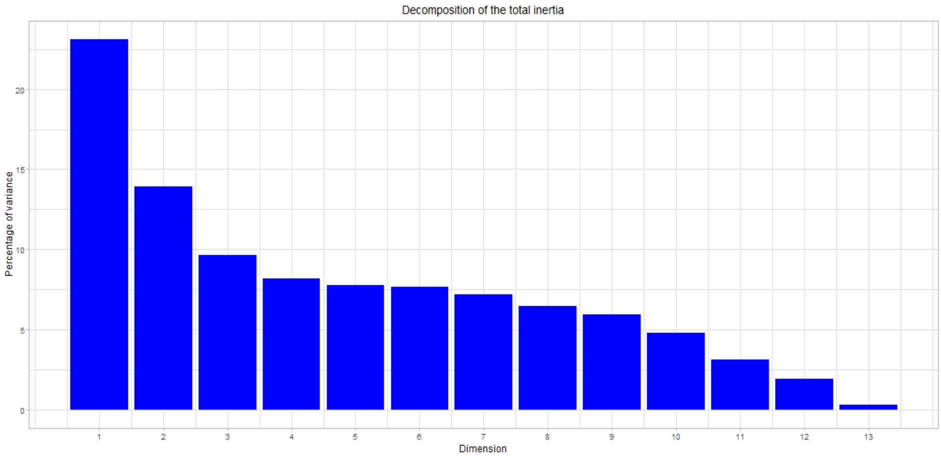
Table 3: Dimensions of the Principal Components

Variable	Dimension 1	Ctr	Cos^2	Dimension 2	Ctr	Cos^2
CAT	0.930	27.869	.800	0.250	12.521	0.300
ATR	-0.362	4.366	0.131	0.710	27.883	0.504
RTR	-0.199	1.315	0.040	0.559	17.297	0.313
FAT	0.567	25.615	0.322	0.169	10.587	0.290
WCT	0.412	25.544	0.337	0.150	11.220	0.361
CFC	0.665	35.147	0.442	0.374	7.715	0.140
OCF	0.731	17.747	0.534	0.319	5.626	0.102
ITR	0.600	11.949	0.359	0.245	3.315	0.060
QR	0.942	29.493	0.887	0.143	1.126	0.020
CR	0.932	28.869	0.868	0.097	0.517	0.009
DER	-0.150	0.221	0.027	-0.081	0.007	0.040
DR	0.078	1.230	0.002	-0.005	0.002	0.010
ICR	-0.012	0.012	0.001	0.003	0.010	0.005

Source: Authors' Processing from R ver 4.0.3

As suggested by Husson et al. (2017), we have maintained two dimensions of the Principal Components, Dimension 1 representing 23.138% and Dimension 2 representing 13.912% of all variances, respectively. We provide the Decomposition of the total inertia histogram to justify the dimension reduction procedure adhered to in this study (refer to Figure 2).

Figure 2: Decomposition of Total Inertia- Factor Extraction



Source: Authors' Processing from R ver 4.0.3

From Figure 2, it can be understood that Dimension 1 and Dimension 2 on the 'x' axis represent the maximum variance in the CFA model ('y' axis).

CHAID Decision Tree Algorithm

CHAID stands for Chi-Squared Automatic Interaction Detector, an effective statistical intervention developed by Kass (1980). It was primarily used for segmentation or tree growth. In this method, a tree is developed on the basis of adjusted Bonferroni significance testing. The same can be utilised for regression analysis and clustering, it can also be used to observe interactions among variables. CHAID tree has the capability to generate more than two decision branches by producing highly visualized multi-way splits. It has a higher degree of comprehension as it can execute both case weights and frequency variables. There are certain pros of using the CHAID square as a predicting tool, as Díaz-Pérez & Bethencourt-Cejas (2016) observe: (1) Chi-square is a non-parametric statistic, (2) Nominal type and interval class variables can be deployed as predicting variables and, (3) Continuous variables can be chosen as criterion variables. The above premises have led the study in adopting CHAID as our decision tree algorithm as it merges the variables best to explain the outcome in the dependent variable, which is CHAID's advantage over other decision tree algorithms (Miller et al., 2014).

Performance Measures

Overall Accuracy: Accuracy may be defined as the percentage of observations that are accurately being estimated, it can also be called the ratio of correctly predicted observations against the total number of observations.

Precision: It may be defined as the ratio of the number of correctly predicted true positive cases against the sum of true positive and false positive.

Recall: Also known as the sensitivity or True Positive rate. It is computed as the ratio of True Positive (the number of correctly predicted cases) to the sum of false negative and true positive.

F-Measure: The harmonic mean value of Precision and Recall measures. (Witten & Frank, 2000) expressed this as:

$$2 \times \frac{p \times R}{p + R}$$

Where; Recall = R
Precision = p

Specificity: It is defined as the True Negative, and it is the ratio of the number of the True Negative to the sum of False Positive and True Negative.

Cut-off: According to certain cut-off values in the features, tree-based models split the data multiple times. This procedure creates various subsets of the dataset, where individual instances belong to one particular subset. The concluding subsets are known as leaf or terminal nodes, and the intermediate subsets are called split nodes or internal nodes. The average outcome of the training data set is deployed to predict the outcome in each leaf node. Decision Trees are a robust tool for classification and regression. Table 4 provides the individual cut-off values against their predicted accuracy that have been plotted in Figures 3, 4, and 5.

Table 4: Cut Off Values and Accuracy

Outcome Variable	*x =Cut off*	y= Accuracy
RoA	.459	.889
RoCE	.379	.879
RoE	.6	.82

Source: Authors' Processing from R ver 4.0.3

Figure 3: Cut Off Plot for RoA

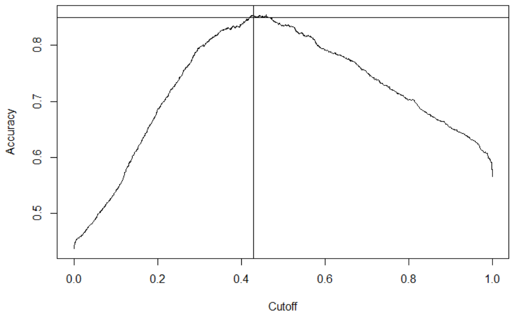


Figure 4: Cut Off Plot for RoCE

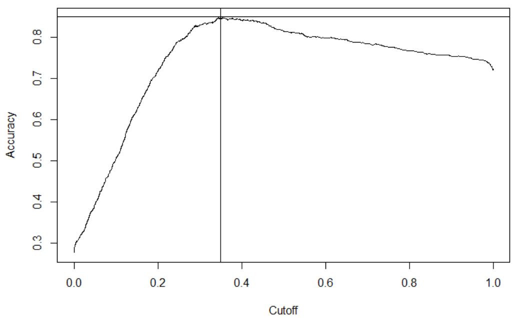
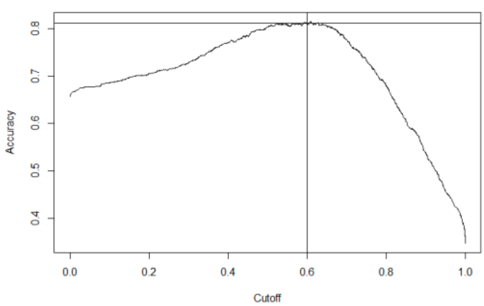


Figure 5: Cut Off Plot for RoE



Source: R Studio ver 4.0.3

Table 5: Performance Measures of Model

Split Criterion	Precision	Specificity	Sensitivity	Accuracy	Misclassification Rate	F-Measure
RoA	0.81	0.70	0.61	0.83	0.17	0.69
RoE	0.61	0.82	0.77	0.80	0.20	0.52
RoCE	0.95	0.78	0.82	0.82	0.18	0.88

Source: Authors' Processing from R ver 4.0.3

Table 6: The Confusion Matrix

Split Criterion		Successful (1)	Unsuccessful (0)
RoA	Successful (1)	746	167
	Unsuccessful (0)	74	470
RoE	Successful (1)	305	92
	Unsuccessful (0)	195	865
RoCE	Successful (1)	1002	217
	Unsuccessful (0)	51	187

Source: Authors' Processing from R ver 4.0.3

Comparative Analysis of CHAID Performance

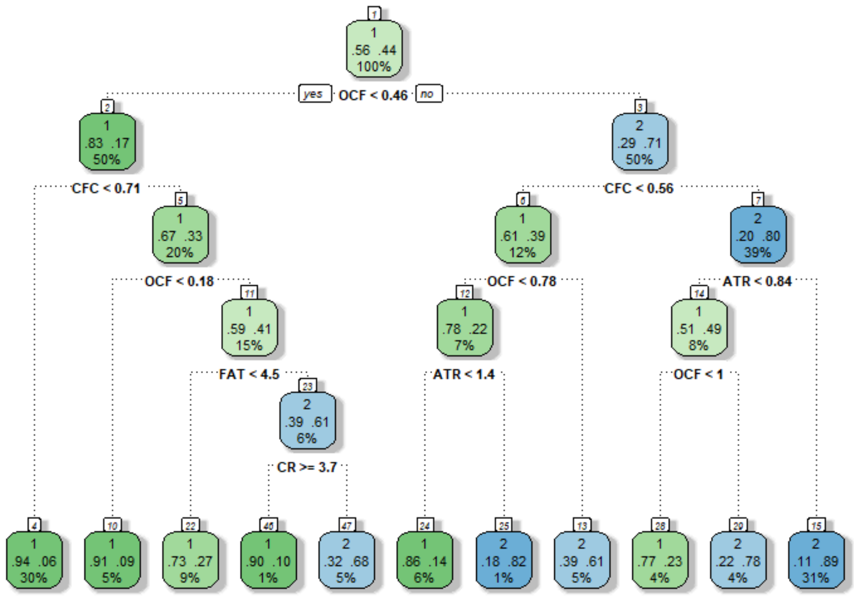
For the three given target variables; RoE, RoCE, and RoA, it is understood from Table 5 and Table 6 that all the variables portray robust accuracy levels and admissible misclassification rates (Husson et al., 2017). In terms of detecting the occurrence of most true positive cases, RoCE achieves a precision of 95% and a robust F-Measure of 0.88 (Tharwat, 2018).

The RoE variable could sense 77% of the true positive cases and a staggering 82% of the true negative cases with a precision of 61%. The variable RoA achieved a predicting precision of 81% while capturing a robust true negative rate of 70% and reports a true positive rate of 61%. Further, it is to be noted that the entire dataset was partitioned as 70% for the training set and the remaining 30% for the test set.

RoA as the target variable predicts the OCF (value being < 0.46) to be the split criterion in developing the decision tree. Therefore, if the value of OCF is below 0.46, the next best alternative is the CFC at < 0.71. If the CFC is > 0.46 then, the branches of the tree will reach out to CFC levels at <0.56 which will further decide the tree's progress. Here the last branch at CR ≥ 3.7 is the most optimized variable.

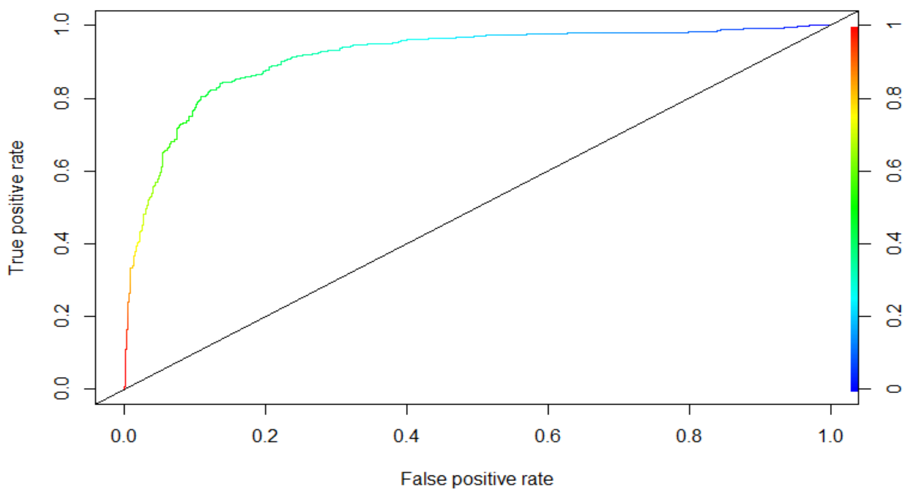
Based on the training and testing of the dataset, it is observed from the above decision tree that has used RoE as a predictor variable that OCF at <0.32 is the decision criterion. If the condition of OCF at <0.32 is met (having only 34% chances), then the tree branches into CFC at <0.58 value, and simultaneously, it is to be noted that there is a 66% chance of the condition not being met. If the condition is not met, the tree branches out to the next available alternative to support RoE as a robust performance indicator.

Figure 6: Decision Tree Plot for RoA



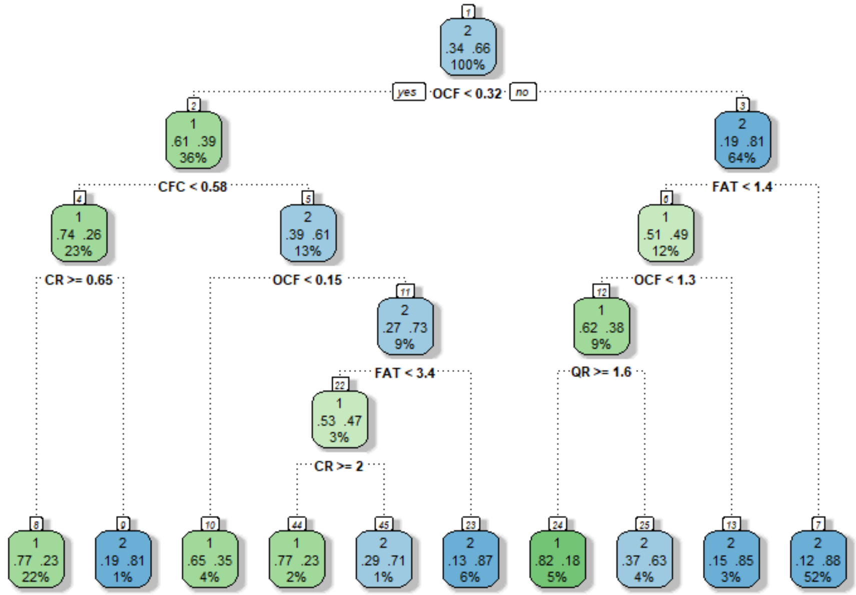
Source: Authors' Processing from R ver 4.0.3

Figure 7: Area under receiver operating characteristic (ROC) Plot for RoA



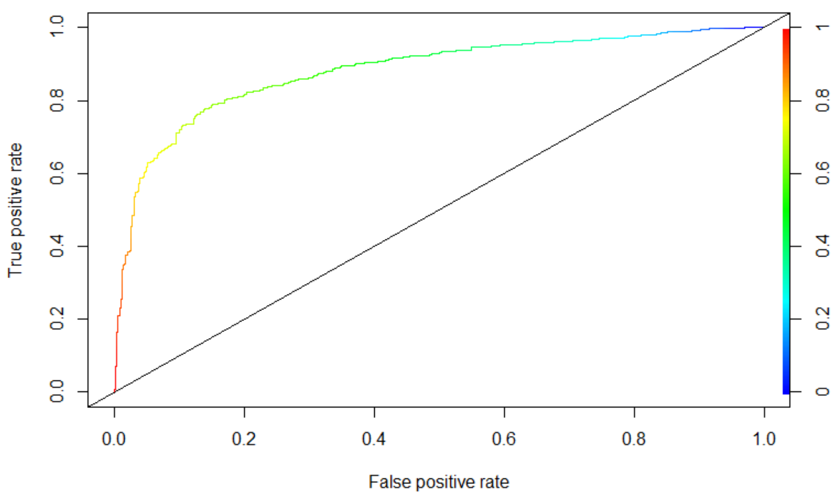
Source: Authors' Processing from R ver 4.0.3

Figure 8: Decision Tree Plot for RoE



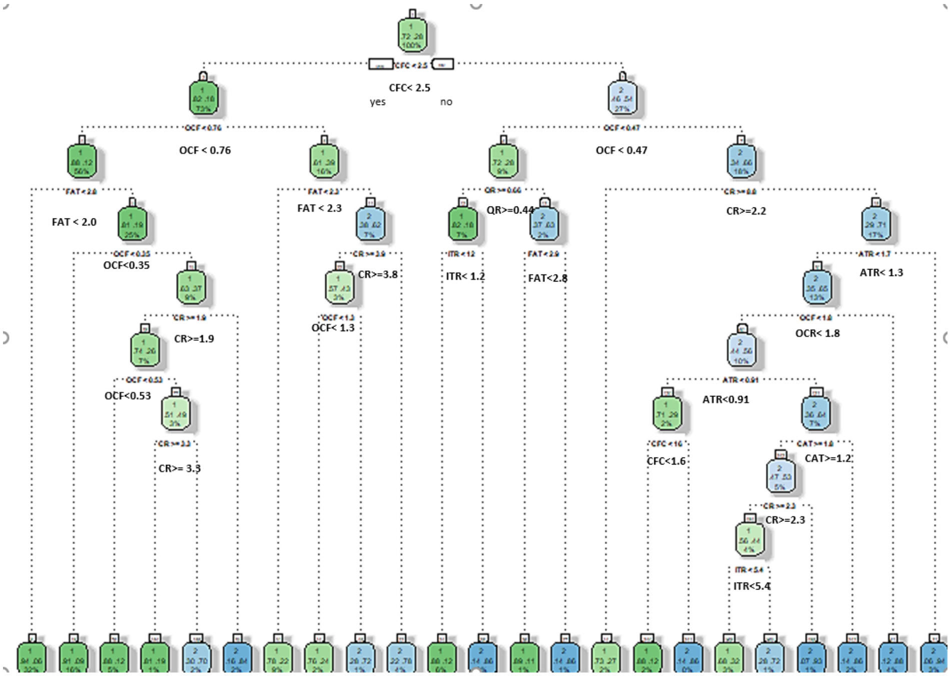
Source: Authors' Processing from R ver 4.0.3

Figure 9: Area under receiver operating characteristic (ROC) Plot for RoE



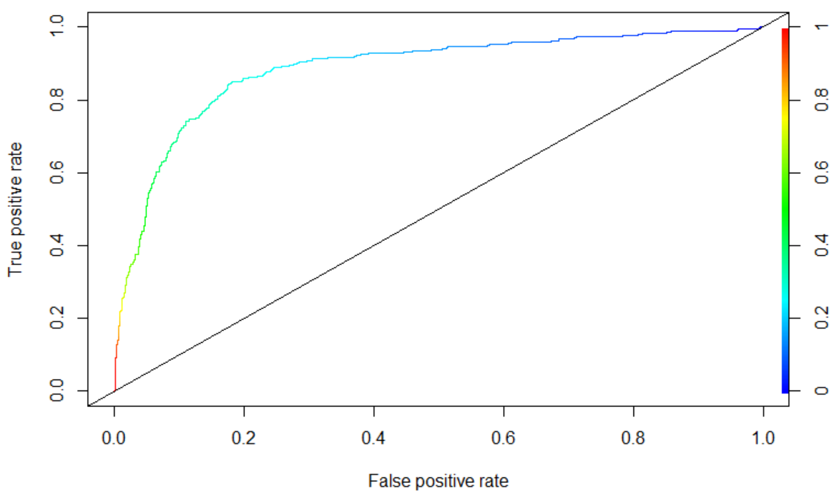
Source: Authors' Processing from R ver 4.0.3

Figure 10: Decision Tree Plot for RoCE



Source: Authors' Processing from R ver 4.0.3

Figure 11: Area under ROC Plot for RoCE



Source: Authors' Processing from R ver 4.03

When targeting RoCE as the dependent variable, the best alternative or the root lies at CFC variable being at < .25 and rests at CR being at >= 3.3 as the most optimized variable to decide upon as it contains 51% chances of occurring. From Figure 10, it can be established that this tree creates many sub-trees, indicating higher chances of misjudgements as the split criterion is yet to be satisfied given all the available combinations. Furthermore, in Figure 10, the area representing the false positive rate is larger in the ROC curve.

Conclusion and Discussion

The main aim of the current study is to examine the best predictor among different proxies of the firm's performance. About 287 Czech firms were selected from three sectors: construction, automobile, and manufacturing. The data was gained from the Albertina database from 2016 to 2020. Three different proxies of firm performance; RoA, RoE, and RoCE were used as dependent variables. After employing the decision tree CHAID algorithm, we closely observed the area under the receiver operating characteristic curve (ROC) for all cases of the dependent variables (Figures 7, 9, and 11). It appears that when RoA was deployed as the split measuring, it induces an improved curve, thereby qualifying RoA as the best predictor of firm performance. Parameters attained for all split measures (RoA, RoCE, and RoE) portray robustness for RoA as the dependent variable with high accuracy and low misclassification rate. From all the decision trees, it could be deciphered that RoA as the dependent variable contains the most robust parameters for successfully predicting the financial performance among Czech firms.

The present study's findings offer practical and theoretical implications. Not much is known about the financial ratios and firm performance, specifically for Czech firms. Theoretically, the study's results will extend the literature not only on financial ratios but also on the proxies of the firm's performance. The results will contribute to the existing literature on which proxy of the firm's performance is the best by using the CHAID approach. Practically, the findings are significant for investors as they consider the firm's performance before making investment decisions on bonds or equity, and creditors. The results of the study will help managers, directors, and policymakers as they make decisions and policies to improve firms' financial performance. Additionally, the findings are useful for researchers and academics as we briefly explained how and why RoA is the best proxy for measuring firm performance.

The results of the current research proposed a new and deeper perception of how financial ratios impact the firm's performance. However, much could be done in the future on the financial ratios and firm performance. Further research can be conducted through different sectors of the economy, increasing time periods, and by including different regions or countries. In future research, we also suggest that the academics and researchers use other financial ratios, non-linear methodologies, and more proxies of the firm's performance for a comprehensive study.

Some limitations warrant consideration to the present research. (i) Only three proxies of the firm's performance were included based on the previous literature and available data. (ii) The corona virus disease (COVID-19) impacts were excluded in the present study. (iii) Due to the availability of data, a short time period was considered (from 2016 to 2020).

Disclosure statement: The authors declare no conflict of interest.

Funding: This work was supported by Internal Grant Agency (IGA) in Tomas Bata University in Zlin, Czech Republic, under the project No. IGA/FAME/2022/012.

References

ABDEL-BASSET, M., DING, W., MOHAMED, R., & METAWA, N. (2020). An integrated plithogenic MCDM approach for financial performance evaluation of manufacturing industries. *Risk Management*, *22*(3), 192–218. https://doi.org/10.1057/s41283-020-00061-4

ABDI, H, & BEATON, D. (2021). *Principal component and correspondence analyses using R*. Springer International Publishing.

ABDI, HERVÉ, & WILLIAMS, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433–459. https://doi.org/10.1002/wics.101

AKGÜN, A. İ., & MEMIŞ KARATAŞ, A. (2020). Investigating the relationship between working capital management and business performance: evidence from the 2008 financial crisis of EU-28. *International Journal of Managerial Finance*. https://doi.org/10.1108/IJMF-08-2019-0294

AL-ZARAREE, A., AL-SAWALHAH, J., & SAMARA, A. (2021). The relationship between the return on equity (ROE) and the capital structure of the Jordanian public shareholding industrial companies. *Academy of Accounting and Financial Studies Journal*, *25*(1), 1–10.

ALMASKATI, N. (2022). The determinants of bank profitability and risk: A random forest approach. *Cogent Economics & Finance*, *10*(1), 2021479. https://doi.org/10.1080/23322039.2021.2021479

AZUR, M. J., STUART, E. A., FRANGAKIS, C., & LEAF, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *Wiley Online Library*, *20*(1), 40-49. https://doi.org/https://doi.org/10.1002/mpr.329

Bawazir, H., Khayati, A., & AbdulMajeed, F. (2021). Corporate governance and the performance of non-financial firms: the case of Oman. *Entrepreneurship and Sustainability Issues*, *8*(4), 595–609. https://doi.org/10.9770/jesi.2021.8.4(36)

BAYRAKTAROGLU, A. E., CALISIR, F., & BASKAK, M. (2019). Intellectual capital and firm performance: an extended VAIC model. *Journal of Intellectual Capital*, *20*(3), 406–425. https://doi.org/10.1108/JIC-12-2017-0184

BISNODE, 2021; Bisnode Albertina, https://www.bisnode.cz/produkty/albertina/. Gold Edition

BLAŽKOVÁ, I., & DVOULETÝ, O. (2022). Relationship between firm total factor productivity and performance: case of the Czech high-tech industry. *International Journal of Entrepreneurial Venturing*.

BOSE, I. (2006). Deciding the financial health of dot-coms using rough sets. *Information & Management*, *43*(7), 835-846.

CAMSKA, D., KLECKA, J., & SCHOLLEOVA, H. (2021). Influence of age on selected parameters of insolvent companies. *Problems and Perspectives in Management*, *19*(2), 77–90. https://doi.org/10.21511/ppm.19(2).2021.07

CHANDRAPALA, P., & KNÁPKOVÁ, A. (2013). Firm-specific factors and financial performance of firms in the Czech Republic. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, *61*(7), 2183–2190. https://doi.org/10.11118/actaun201361072183

CHICCO, D., TÖTSCH, N., & JURMAN, G. (2021). The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Mining*, *14*(1), 13. https://doi.org/10.1186/s13040-021-00244-z

ČINČALOVÁ, S., & HEDIJA, V. (2020). Firm characteristics and corporate social responsibility: the case of Czech transportation and storage industry. *Sustainability*, *12*(5), 1992. https://doi.org/10.3390/su12051992

DELEN, D., KUZEY, C., & UYAR, A. (2013). Measuring firm performance using financial ratios: A decision tree approach. *Expert Systems with Applications*, *40*(10), 3970–3983. https://doi.org/10.1016/j.eswa.2013.01.012

DÍAZ-PÉREZ, F. M., & BETHENCOURT-CEJAS, M. (2016). CHAID algorithm as an appropriate analytical method for tourism market segmentation. *Journal of Destination Marketing and Management*, *5*(3), 275–282. https://doi.org/10.1016/j.jdmm.2016.01.006

DUNTEMAN, G. H. (1989). Sage university paper series on quantitative applications in social sciences. *Principal components analysis*, Newbury Park, CA: Sage.

DWILITA, H., & MINGKA, A. Y. (2022). The effect of Nim, Npl, firm size and exchange rate on profitability with car as an intervening variable in the company conventional banking sector registered in Indonesia stock exchange 2015-2019 period. *Accounting and Business Journal*, *4*(1), 28–37. https://doi.org/10.54248/ABJ.V4I1.4052

FIELD, A. P. (2005). Intraclass Correlation. *Encyclopedia of Statistics in Behavioral Science*. Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/0470013192.bsa313

GAYATHRI, C. S., & VIJAYALAKSHMI, S. (2021). A study on the effect of cost of capital on profitability of a company. *International Research Journal on Advanced Science Hub*, *3*(0), 54–59. https://doi.org/10.47392/IRJASH.2021.165

HABIBI, A. & IQBAL, M. (2021). Benefits of financial ratios for financing sharia banking Indonesia. *Jurnal Ekonomi Dan Keuangan Syariah*, *5*(1), 1–12. https://doi.org/10.29313/amwaluna.v5i1.5299

Harman, H. H. (1976). *Modern Factor Analysis.* University of Chicago Press.

HUA, Z., WANG, Y., XU, X., ZHANG, B., & LIANG, L. (2007). Predicting corporate financial distress based on integration of support vector machine and logistic regression. *Expert Systems with Applications*, *33*(2), 434-440.

HUSSON, F., LÊ, S., & PAGÈS, J. (2017). *Exploratory multivariate analysis by example using R*. Routledge New York, NY.

ISLAMIYATI, G. M., & DIANA, N. (2021). Effect of Mudarabah and Musharakah financing on return on equity (ROE) case studies on Islamic commercial banks in Indonesia for the period 2015-2019. *Jurnal Ekonomi Syariah Teori Dan Terapan*, *8*(3), 255–263. https://doi.org/10.20473/VOL8ISS20213PP255-263

JOHNSON, R. A. & WICHERN, D. W. (2013). *Applied Multivariate Statistics analysis* (6th ed.). New Jersy: Pearson Education Inc.

JYOTI, G., & KHANNA, A. (2021). Does sustainability perforcmance impact financial performance? Evidence from Indian service sector firms. *Sustainable Development*, *29*(6), 1086-1095. https://doi.org/10.1002/sd.2204

KAISER, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/BF02291575

KASS, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, *29*(2), 119. https://doi.org/10.2307/2986296

KAYANI, U. N., DE SILVA, T. A., & GAN, C. (2020). Working capital management and firm performance relationship: An empirical investigation of Australasian firms. *Review of Pacific Basin Financial Markets and Policies*, *23*(3). https://doi.org/10.1142/S0219091520500265

KORENIUS, T., LAURIKKALA, J., & JUHOLA, M. (2007). On principal component analysis, cosine and Euclidean measures in information retrieval. *Information Sciences*, *177*(22), 4893–4905. https://doi.org/10.1016/j.ins.2007.05.027

KUČERA, J., VOCHOZKA, M., & ROWLAND, Z. (2021). The ideal debt ratio of an agricultural enterprise. *Sustainability (Switzerland)*, *13*(9). https://doi.org/10.3390/su13094613

LARASATI, C. I., & PURWANTO, P. (2022). How financial ratios and firm size affect profitability: evidence from food and beverages industry in Indonesia. *The Winners*, *23*(1), 43–50. https://doi.org/10.21512/TW.V23I1.7099

LÊ, S., JOSSE, J., & HUSSON, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 1–18. https://doi.org/10.18637/jss.v025.i01

LISTIORINI, L., & PUTRI, R. A. (2022). The effect of leverage, return on asset (ROA), and company size on company value with good corporate governance as moderating variables (study on Manufacturing Companies). *Accounting and Business Journal*, *4*(1), 38–49. https://doi.org/10.54248/ABJ.V4I1.4053

LUBIS, I., & ALFIYAH, F. N. (2021). Effect of return on equity and debt to equity ratio to stock return. *Indonesian Financial Review*, *1*(1), 18–32. https://doi.org/10.55538/ifr.v1i1.3

MILLER, B., FRIDLINE, M., LIU, P. Y., & MARINO, D. (2014). Use of CHAID Decision Trees to formulate pathways for the early detection of metabolic syndrome in young adults. *Computational and Mathematical Methods in Medicine*.

https://doi.org/10.1155/2014/242717

NILASHI, M., BAGHERIFARD, K., RAHMANI, M., & RAFE, V. (2017). A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques. *Computers and Industrial Engineering*, *109*, 357–368. https://doi.org/10.1016/j.cie.2017.05.016

PIVOŇKOVÁ, A., & TEPPEROVÁ, J. (2021). Interest limitation rule under ATAD: case of the Czech Republic. *DANUBE*, *12*(2), 121–134. https://doi.org/10.2478/DANB-2021-0009

REZAIE, K., RAMIYANI, S. S., NAZARI-SHIRKOUHI, S., & BADIZADEH, A. (2014). Evaluating performance of Iranian cement firms using an integrated fuzzy AHP-VIKOR method. *Applied Mathematical Modelling*, *38*(21–22), 5033–5046. https://doi.org/10.1016/j.apm.2014.04.003

SEAN, N. A. M., ANAGREH, S., AL-DALAIEN, B. O. A., ALMUGARI, F., KHALED, A. S. D., & AL-HOMAIDI, E. A. (2021). Working capital management and banks' performance: Evidence from India. *The Journal of Asian Finance, Economics and Business, 8*(6), 747–758. https://doi.org/10.13106/JAFEB.2021.VOL8.NO6.0747

SHARMA, R. K., BAKSHI, A., & CHHABRA, S. (2020). Determinants of behaviour of working capital requirements of BSE listed companies: An empirical study using co-integration techniques and generalised method of moments. *Cogent Economics & Finance*, *8*(1), 1720893. https://doi.org/10.1080/23322039.2020.1720893

SOEWARNO, N., & TJAHJADI, B. (2020). Measures that matter: an empirical investigation of intellectual capital and financial performance of banking firms in Indonesia. *Journal of Intellectual Capital*, *21*(6), 1085–1106. https://doi.org/10.1108/JIC-09-2019-0225

THARWAT, A. (2018). Classification assessment methods. *Applied Computing and Informatics*. https://doi.org/10.1016/j.aci.2018.08.003

TRACY, A. (2012). *Ratio analysis fundamentals: how 17 financial ratios can allow you to analyse any business on the planet*. CreateSpace Independent Publishing Platform.

VALASKOVA, K., KLIESTIK, T., & KOVACOVA, M. (2018). Management of financial risks in Slovak enterprises using regression analysis. *Oeconomia Copernicana*, *9*(1), 105–121. https://doi.org/10.24136/oc.2018.006

VAN BUUREN, S., & GROOTHUIS-OUDSHOORN, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

VRBKA, J. (2020). The use of neural networks to determine value based drivers for SMEs operating in the rural areas of the Czech Republic. *Oeconomia Copernicana*, *11*(2), 325–346. https://doi.org/10.24136/OC.2020.014

VRBKA, J., KALINOVÁ, E., & DVOŘÁKOVÁ, Z. (2022). Optimization of the capital structure of an agricultural company in the Czech Republic. *SHS Web of Conferences*, *132*, 1008. https://doi.org/10.1051/SHSCONF/202213201008

WANG, T. C., & CHEN, Y. H. (2006). Applying rough sets theory to corporate credit ratings. In *2006 IEEE International Conference on Service Operations and Logistics, and Informatics* (pp. 132-136). IEEE. https://doi.org/10.1109/SOLI.2006.329050

WITTEN, I. H., & FRANK, E. (2000). Machine Learning Algorithms in Java. *Machine Learning*, 265–320.

XU, J., & LI, J. (2020). The interrelationship between intellectual capital and firm performance: evidence from China's manufacturing sector. *Journal of Intellectual Capital*. https://doi.org/10.1108/JIC-08-2019-0189/FULL/HTML

YALCIN, N., BAYRAKDAROGLU, A., & KAHRAMAN, C. (2012). Application of fuzzy multi-criteria decision making methods for financial performance evaluation of Turkish manufacturing industries. *Expert Systems with Applications*, *39*(1), 350-364. https://doi.org/10.1016/j.eswa.2011.07.024

YOUSAF, M. (2022a). Bankruptcy risks, firm size, and firm profitability: A dynamic panel data approach. *International Journal of Sustainable Economy*. https://doi: 10.1504/IJSE.2023.10045805

YOUSAF, M. (2022b). Labour productivity and firm performance: evidence from certified firms from the EFQM Excellence Model. *Total Quality Management and Business Excellence*. https://doi.org/10.1080/14783363.2022.2054319

YOUSAF, M. (2021). Intellectual capital and firm performance: evidence from certified firms from the EFQM Excellence Model. *Total Quality Management and Business Excellence*. https://doi.org/10.1080/14783363.2021.1972800

YOUSAF, M., & BRIS, P. (2021a). Assessment of bankruptcy risks in Czech companies using regression analysis. *Problems and Perspectives in Management*, *19*(3), 46–55. https://doi.org/10.21511/PPM.19(3).2021.05

YOUSAF, M., & BRIS, P. (2021b). Effects of working capital management on firm performance: evidence from the EFQM certified firms. *Cogent Economics and Finance*, *9*(1). https://doi.org/10.1080/23322039.2021.1958504

YOUSAF, M., BRIS, P., & HAIDER, I. (2021). Working capital management and firm's profitability: Evidence from Czech certified firms from the EFQM excellence model. *Cogent Economics and Finance*, *9*(1). https://doi.org/10.1080/23322039.2021.1954318

ZAINI, B. J., & MAHMUDDIN, M. (2019). Classifying firms' performance using data mining approaches. *International Journal Supply Chain Management*, *8*(1), 690.

ZUARDI, M. H. (2021). Measuring healthiness of islamic banks using solvabilitas financial ratios. *International Journal of Islamic Economics*, *3*(1), 17–36. https://doi.org/10.32332/IJIE.V3I1.3279