

Critical Analysis of Financial Ratios through Integrated Principle Component Analysis (PCA) and Canonical Correlation Analysis (CCA)

G. Anupama

*Part time PhD scholar, Department of Mechanical Engineering,
College of Engineering(A), Andhra University, Visakhapatnam, India*

V.V.S. Kesava Rao

*Professor, Department of Mechanical Engineering,
College of Engineering(A), Andhra University, Visakhapatnam, India*

Abstract- The aim of this study is to examine the multivariate associate between various financial ratios and identify the critical financial ratios. The segregation of financial ratios into input and output ratios are useful to determine the business insolvency/failure and financial efficiency of the business organizations. A total of 18 software companies are considered with nine financial ratios. PCA is applied in order to prove its usefulness in reducing financial information expressed as ratios. The achieved results, in terms of variable selection, are next justified by the application of a Canonical Correlation Analysis (CCA). The use of both methods shows they are complementary, since the ratios correlated to the extracted factors in PCA are also important in defining the canonical variates in CCA, showing the relationship between them.

Key words – Principal Component analysis, Canonical Correlation Analysis

I. INTRODUCTION

Advances in the information and communications technologies over the last decades have fostered the generation of huge amounts of data, and their processing involves a high cost in terms of timing and resources. Some of the financial ratios are usually being quoted in many companies balance statements and it is better if we can group them into certain number of groups, so that we can more clearly define the relationship existing between the groups.

1.1 Stockholders' equity ratio (IP1) –

The shareholder equity ratio shows how much money shareholders will receive if a company has to liquidate its assets. The ratio is expressed as a percentage and is calculated by dividing a company's total shareholder equity by its total assets. A company's total shareholder equity is the combined money its shareholders have invested, as well as any subsequent investments, plus its retained earnings.

1.2 Turnover rate of accounts receivable (Debtor Turnover Ratio) (IP2) –

The receivables turnover ratio is an activity ratio measuring how efficiently a firm uses its assets. Receivables turnover ratio can be calculated by dividing the net value of credit sales during a given period by the average accounts receivable during the same period

= net value of credit sales / average accounts receivable

Debtor turnover ratio is the relationship between net sales and average debtors. It is also called account receivable turnover ratio because we debtor and bill receivables' total is used for following formula

1.3 Turnover rate of inventory (IP3) –

The inventory turnover ratio is defined as ratio of cost goods sold to average inventory maintained.

1.4 Return of stockholder equity (OP1) –

Shareholders' equity is equal to a company's assets minus its debt, ROE could be thought of as the return on net assets. ROE is expressed as a percentage and can be calculated for any company if net income and equity are both positive numbers. The return on equity ratio or ROE is a profitability ratio that measures the ability of a firm to generate profits from its shareholders investments in the company. In other words, the return on equity ratio shows how much profit each dollar of common stockholders' equity generates.

1.5 Quick ratio (OP2) –

The quick ratio is an indicator of a company's short-term liquidity position and measures a company's ability to meet its short-term obligations with its most liquid assets.

1.6 Ratio of cash flow (OP4) –

The operating cash flow ratio is calculated by dividing cash flows from operations by current liabilities. The operating cash flow ratio is a measure of the number of times a company can pay off current debts with cash generated within the same period

Cash flow ratio = Operating cash flow /current liabilities

1.7 Return of assets (OP6) –

Return on assets (ROA) is an indicator of how profitable a company is relative to its total assets. ROA gives a manager, investor, or analyst an idea as to how efficient a company's management is at using its assets to generate earnings. Return on assets is displayed as a percentage and its calculated as:

ROA = Net Income / Total Assets

1.8 Market share (MS) –

Market share represents the percentage of an industry, or market's total sales, that is earned by a particular company over a specified time period. Market share is calculated by taking the company's sales over the period and dividing it by the total sales of the industry over the same period. This metric is used to give a general idea of the size of a company in relation to its market and its competitors.

Market share = Company's sales /Total industry's sales

1.9 Annual total fixed assets (IP4) –

Fixed assets as an accounting term refers to the property that is used by a company for producing of goods and services. They include machinery and tools, buildings, vehicles, and land, of course. Fixed assets, as a rule, have a useful time or useful life in excess of one year. Useful time implies this period of the asset's exploitation during which it is in good technical condition, is able to produce reliably, is cost effective (i.e. a tool that consumes reasonable cash for regular maintenance only, and not for major or overhaul repairs).

Complementary, a Canonical Correlation Analysis (CCA) is applied, whose objective is to analyze whether the group of selected ratios by PCA is related to the group of removed ratios, which would mean that the information supplied for both sets of variables is redundant and the elimination of the latter would be appropriate. Finally, a study of the relationship between both methods is carried out, by comparing the results obtained with each analysis in our empirical study.

II. PROBLEM STATEMENT

The aim is to group financial ratios that characterizes business failure/success in IT companies by selecting them through the application of PCA (principal component analysis) and CCA(Canonical correlation analysis). The objective is to divide them into two groups which are mostly correlated with each other and isolated with other variables. That is to consider the Dependent and independent variables and their inter relationships.

2.1 Data source –

In this case, Data has been collected from 18 IT companies. Last 5 years financial data is considered for the analysis and subsequently these financial ratios are separated as 9 variables based on different parameters.

III. LITERATURE REVIEW

Hart (2017) [1] made a study to examine the multivariate association between physical activity (PA) parameters and body composition (BC) measures in college students. In the study, the data on 60 college students is collected through a questionnaire. The study resulted that there is a correlation between PA and BC constructs.

Beaver.W.H (1966) [2] initiated the statistical models to make financial forecasts made a turning point in the survey of the business failure symptoms. One of the most important applications in the field of financial analysis, where data reduction methods are applied because of the huge amount of available quantitative information is always difficult to tackle with ease on the evaluation of business failure or insolvency by the stake holders. A firm's disappearance has on a diversity of agents acting in the economic system and involved in the firm success or failure (Wu, 2010) [4]. To anticipate the financial difficulties firms can face, a large list of business failure prediction models have been

developed, after Beaver's (1966) [2] and Altman's (1968) [3] models, which are considered to be pioneering in this field (Wu, 2010) [4].

Most of them have used financial ratios relating different accounting entries in financial statements as explanatory variables to be included in the models, as they allow processing all financial information in an easier way and making comparisons both among different firms and different periods (Hernández, 2011) [5].

Nevertheless, due to the lack of a theory of business failure that could be a guide for the selection of ratios to be included in the models (Scott, 1981) [6], it has been basically empirical, based on their popularity in literature and their predictive success in previous research, as Beaver (1966) [2] did (Balcaen & Ooghe, 2006) [15]. This criterion has resulted in a large list of ratios potentially explanatory of business failure (García-Gallego et al., 1996) [7], which makes it necessary the previous application of a data reduction method, such as PCA.

Chiu and Wang (2011) [8] proposed an evaluation model to assess the financial performance of construction companies. In the study, the authors combined the methods of Strength Weakness Opportunity and threats (SWOT), Canonical Correlation Analysis (CCA) and Data Envelopment Analysis (DEA). From the study, it was concluded that the proposed model is an effective tool to determine performance evaluation of construction companies in general and will be useful to other companies also. Chiu and Wang has reviewed some significant studies related to failure performance evaluation in order to understand the indicators and evaluation methods used for evaluation performance. They provide the foundation to construct a new model for assessing the financial performance of (software companies similar to) construction companies (Chiu and Wang, 2011) [8] considered for evaluation of "stock holders equity ratio", "Turnover rate of accounts receivables", "Turnover rate of inventory", "Return of stock holder equity", "Quick ratio". "Annual fixed assets", "Operating cash flow" and "return of assets ratio".

Lan Yang (2017) [9] used SPSS software and statistical method of PCA (Principal component analysis) to establish the financial early warning model that predicts and measures risks associated in various Chinese companies using its financial information.

Comrey & Lee (1992) [10] mentioned that Principal component analysis involves extracting linear composites of observed variables. And also, Factor analysis (FA) is based on a formal model predicting observed variables from theoretical latent factors. In psychology these two techniques are often applied in the construction of multi-scale tests to determine which items load on which scales. They typically yield similar substantive conclusions (for a discussion see Comrey & Lee (1992) [10] Factor-Analytic Methods of Scale Development in Personality and Clinical Psychology).

Hair et al., (1998) [11], justified When faced with the choice between PCA and Factor analysis, it was decided that PCA would be the most appropriate technique for our study; the following points justify this choice. The application of PCA is more popular than factor analysis in the ratio classification literature as well as in many other areas in the financial and banking literature.

Although factor analysis might appear appealing, given that it uncovers the dimensions that exist in the dataset, it involves some complications which have, to an extent, contributed to the popularity of PCA (Hair et al., 1998, pp. 102-103) [11]. PCA is more appropriate given that the study is more driven by data reduction than by the purpose of uncovering the latent dimensions in the data.

Thus, there is a better chance that the results of this study can be utilized by researchers who seek to choose financial ratios as inputs for their statistical models. By choosing one ratio to represent each of the identified components, researchers can be assured that they have covered all the dimensions that exist in the larger set of bank ratios; Hair et al., 1998, [11].

PCA is a data reduction statistical method that is widely applied when dealing with a large amount of quantitative information, in order to reduce the number of variables to consider. Its objective is to obtain new orthogonal and linearly independent variables, which are named factors or principal components, as linear combinations of the original variables. Regarding the number of factors to be extracted, it is necessary to take into account the percentage of variance which is explained by each factor and which also measures their relative importance (García-Gallego et al., 1996) [7].

Mardia et al. (1979) [12] point out as a practical rule to extract an enough number of factors so that they explain at least 90% of the original variance. However, Hair et al. (2010) [9] indicate this criterion should not be used 'in the social sciences, where information is often less precise', so they consider as satisfactory 'a solution that accounts for 60 percent of the total variance (and in some instances even less)'. In any case, as García-Gallego et al., (1996) [7] notes, a balance between the highest percentage of total variance and the summary of this variance in a reduced number of components must be achieved. Actually, in most research papers a percentage between 70 and 80% is usually enough, and this percentage is usually explained by 3 or 4 components². The extracted factors are interpreted by their correlations with the original variables, which allow giving a description to the components.

Jolliffe(2002) [13] pointed out that “the central idea of principal component analysis is to reduce the dimensionality of a data set in which there are a large number of interrelated variables, while retaining as much as possible of the variation present in the data set”. The reduction will result in a new set of variables called the principal components which are not correlated and where the first few components retain most of the variation that was present in all the original variables

PCA-CCA methodology –

Nevertheless, due to the lack of a theory of business failure that could be a guide for the selection of ratios to be included in the models (Scott, 1981) [14], it have been basically empirical, based on their popularity in literature and their predictive success in previous research, as Beaver (1966) [2] did (Balcaen&Ooghe, 2006) [15]. This criterion has resulted in a large list of ratios potentially explanatory for business failure (Garcia-Gallego et al., 2016) [7], which makes it necessary the previous application of a data reduction method, such as PCA. In order to analyze the possibilities of PCA in the processing of financial information, in this paper we try to identify those characteristics which better explain business failure, due to their correlation with the extracted factors by PCA. The subsequent application of a CCA allows enriching the results, in the sense of analyzing the relationship between the group of selected ratios by PCA and the group of removed ratios.

In this paper, a valuable hybrid analysis tool followed to assess the critical financial ratios for evaluation of performance of software companies, but also offers useful information for practical operations. It will be helpful for investors, financial institutions, software companies and people who are interested in engaging in the industry of software. The Proposed Methodology is discussed in the following steps.

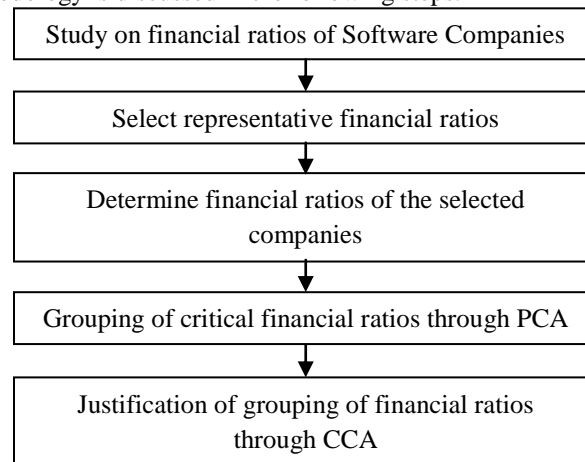


Figure 1. Flowchart of Critical Financial Ratios for Performance Evaluation of Software Companies

3.1 Step 1: Applying principal component analysis –

On the initial list of financial ratios, a PCA was applied in order to reduce their number to a smaller set of factors which reflect the main business activity issues, thus by reducing the number of variables and subsequently considering the least possible number of variables, which explain a high percentage of the original variance, least number of variable size implies to removal of the redundant information they might contain in the original data set. The procedure was carried out in several steps using the statistical software SPSS 17.

3.1.1 Step 1.1: Correlation matrix –

Before conducting a principal components analysis, one has to check the correlations between the variables. If any of the correlations are too high (say above .9), you may need to remove one of the variables from the analysis, as the two variables seem to be measuring the same thing. Another alternative would be to combine the variables in some way (perhaps by taking the average). If the correlations are too low, say below 0.1, then one or more of the variables might load only onto one principal component (in other words, make its own principal component).

3.1.2 Step 1.2: Conduct KMO Test–

The KMO measures the sampling adequacy (which determines if the responses given with the sample are adequate or not) which should be close than 0.5 for a satisfactory factor analysis to proceed. Kaiser (1974) recommend 0.5 (value for KMO) as minimum (barely accepted), values between 0.7-0.8 acceptable, and values above 0.9 are superb.

3.1.3 Step 1.3: Conduct Bartlett's Test of Sphericity –

Bartlett's test is another indication of the strength of the relationship among variables. This tests the null hypothesis that the correlation matrix is an identity matrix.

3.1.4 Step 1.4: Findout Communalities –

This is the proportion of each variable's variance, explained by the Principal components. It can be defined as the sum of squared factor loadings. The initial value of the communality in a principal component analysis is 1. The next item from the output is a table of communalities which shows how much of the variance (i.e. the communality value which should be more than 0.65 to be considered for further analysis. Else these variables are to be removed from further steps factor analysis) in the variables accounted for by the extracted factors. The objective of the principal component analysis is to reduce the dimensionality by reducing the number of components with which we can minimize the variable size. The reduction of number of components is achieved by the formation of total variance table.

3.1.5 Step 1.5: Formation of Total variance –

The total variance explains the percentage of variability in the original variables. The variance table represents the initial Eigen values, percentage of variance and cumulative percentage of Eigen values.

3.1.5.1 Step 1.5.1: Scree plot –

A plot is drawn for Eigen values as Y-axis and no of components as X-axis

3.1.6 Step 1.6: Formulate component matrix –

This table contains component loadings, which are the correlations between the variable and the component. Because these are correlations, possible values range from -1 to +1. Component matrix shows the correlation coefficients between each PC and each variable. The sign of the coefficients indicates if the relation between the coefficient and the variable is direct or inverse, while the numeric value indicates the correlation strength between the variables.

3.1.7 Step 1.7: Formulate Rotated Component Matrix –

The idea of rotation is to reduce the number factors on which the variables under investigation have high loadings. Rotation does not actually change anything but makes the interpretation of the analysis easier. Rotation maintains cumulative percentage of variation explained by the extracted components, but that variation is now spread more evenly over the components. Clustering is possible in rotated matrix.

There are ratios which have been selected (first set) and other ones that have been removed from the analysis (second set). Since they measure the same issue in a different way, ratios in each group are supposed to be related, that is, they contain similar or redundant information (LeClere, 2006).

3.2 Step 2: Applying Canonical Correlation Analysis –

CCA is useful to study the relationships between both sets of variables, by obtaining linear combinations of the original variables, called canonical variates, which present the highest correlation with each other and, simultaneously, are in correlation to the rest of linear combinations that can be obtained.

3.2.1 Step 2.1: Classification of dependent and Independent variables –

PCA is implemented as discussed in previous section to identify the financial ratios as dependent and independent variables. Apply the data in SPSS as provided.

3.2.2 Step 2.2: Multivariate test –

Multivariate tests were performed to test whether the maximal correlation and those orthogonal successive correlations that follow were significantly different from zero. All the p values of the three tests, that is, the Pillai's criterion, Hotellings' trace and Wilks' test were less at .05 level of significance. The null hypothesis may be rejected, if we can conclude that canonical correlations were significantly different from zero.

3.2.3 Step 2.3: Determination of number of roots for the given set of data –

The canonical analysis will select the minimum canonical roots based on the consideration of minimum number of variables set.

3.2.4 Step 2.4: determine Canonical correlation coefficients and the eigenvalues –

Eigen values are calculated using the squared canonical correlation. The largest Eigen value is equal to the largest squared correlation divided by first largest squared correlation. The relative size of the Eigen values reflects how much of the variance in the canonical variates can be explained by the corresponding canonical correlation. Always the Eigen value corresponding to the first correlation is the greatest.

3.2.5 Step 2.5: Test the significance of each root –

The test of significance of each root is verified by the p value. If p value is less than 0.05 then that particular root is significant.

3.2.6 Step 2.6: Determine raw canonical coefficients for dependent variables –

The raw canonical coefficients above are used to generate the canonical variates, represented by the columns (1 2 3) in the coefficient tables, for each set. They are interpreted in a manner analogous to interpreting regression coefficients.

3.2.7 Step 2.9: determination of raw canonical coefficients for independent canonical variables (note: covariates) –

Raw canonical coefficient defines the relationship between the variables in a given group and the canonical variates. They can be interpreted in the same manner as regression coefficients by assuming the canonical variate as the response variable.

3.2.8 Step 2.7 Determine Standardized canonical coefficients for dependent variables –

In the standardized canonical coefficients, the dependent variables are rescaled to have a mean of zero and a standard deviation of 1. These coefficients would indicate how a 1 standard deviation increase in the variable would change in the variate.

When the variables in the model have very different standard deviations, the standardized coefficients allow for easier comparisons among the variables. The standardized canonical coefficients are interpreted in a manner analogous to interpreting standardized regression coefficients.

3.2.9 Step 2.10: Determine Standardized canonical coefficients for independent canonical variables (covariates) –

In the standardized canonical coefficients, the independent variables are rescaled to have a mean of zero and a standard deviation of 1. These coefficients would indicate how a 1 standard deviation increase in the variable would change in the variate.

When the variables in the model have very different standard deviations, the standardized coefficients allow for easier comparisons among the variables. The standardized canonical coefficients are interpreted in a manner analogous to interpreting standardized regression coefficients.

3.2.10 Step 2.8: Determine Correlations between dependent and canonical variables –

These are the correlations between each dependent variable in a group and the group's canonical variates.

3.2.11 Step 2.11 Determination of Correlations between Independent canonical variables (Covariates) and dependent canonical variables (canonical variables) –

These are the correlations between each independent variable in a group and the group's canonical variates.

IV. RESULTS AND ANALYSIS

The data on nine financial ratios of 18 software companies from the financial year 2013-14 to 2017-18 is compiled from the financial reports and is presented in Appendix -A. The average data on the financial ratios in average perspective is shown in table 1.

Table-1 Average values of Financial Ratios

Software Company	Stock-holders equity ratio (FR1=IP.1)	Turnover rate of accounts receivables (FR2=IP.2)	Turnover rate of inventory (FR3=IP.3)	Return of stockholder equity (FR4=OP.1)	Quick Ratio (FR5=OP.2)	Operating cash flow ratio (FR7=OP.4)	Return of Assets (FR8=OP.6)	Market Share (FR9=M.S)	Annual Fixed assets (FR6=IP.4)
SWC1	0.4835	5.8657	1071.0390	-0.5958	0.4487	0.1244	-0.1191	0.0038	-0.0368
SWC2	0.0063	5.6527	153.4263	30.7891	2.3799	0.6032	0.1867	0.1252	0.2270

SWC3	0.0112	5.9329	3178.4408	19.5916	4.1768	0.7303	0.1842	0.1935	0.2722
SWC4	0.0159	4.5723	363.1890	6.6993	1.8083	0.5084	0.1066	0.0101	0.1180
SWC5	0.0170	6.6350	1284.0541	3.2239	2.5205	0.3096	0.0499	0.0025	0.0753
SWC6	0.0389	5.6359	1228.7885	5.2961	3.0064	0.7537	0.1712	0.0138	0.1692
SWC7	0.0287	8.0097	1813.4911	3.1501	3.4431	0.5657	0.0898	0.0169	0.1563
SWC8	0.0274	4.8101	3451.3026	3.7416	2.1626	0.4399	0.1005	0.0084	0.1615
SWC9	0.0065	5.6263	1542.9972	28.4322	4.3278	0.6142	0.1832	0.0133	0.3908
SWC10	0.0342	6.0779	1253.5642	4.2276	3.6714	0.7130	0.1398	0.0074	0.1873
SWC11	0.0177	2.6234	183.4787	0.9567	0.4572	0.1419	0.0188	0.0103	0.3149
SWC12	0.0315	5.5766	505.0253	5.8860	4.9653	1.0348	0.1759	0.0015	0.0479
SWC13	0.0113	6.4123	757.7908	13.7595	1.9409	0.4018	0.1460	0.0063	0.0874
SWC14	0.0023	5.0137	4190.0039	118.2106	3.9594	1.0799	0.2648	0.3333	0.2773
SWC15	0.0611	5.4071	3408.2780	4.0310	2.7842	0.7857	0.2321	0.0033	0.2178
SWC16	0.0179	5.2446	813.7534	8.0653	2.0701	0.3880	0.1376	0.0807	0.1731
SWC17	0.0085	5.4578	103.8384	15.6016	2.0906	0.5367	0.1263	0.1606	0.2093
SWC18	0.0236	6.0394	20.5325	5.6972	2.2980	0.5424	0.1350	0.0090	0.1384

From the table 1, it is found that Stockholders equity ratio is high for the compaySWC1 (48.35%) followed by SWC5(6.11%) and SWC6(3.89%), Whereas the companies SWC2 and SWC14 bottoms the chart with a least average stock holders equity ratio of 0.0063 and 0.0023 respectively.

In case of Turnover rate of accounts receivables, SWC7 is in the first position with highest average of 8.0097 followed by SWC5 (6.635) and SWC13 (6.4123) in second and third position. Company SWC11 is at the bottom most position with the least average of 2.6234. In terms of Turnover rate of inventory SWC14 is at the top position with high average of 4190.004 Followed by SWC8 (3451.303) and SWC15 (3408.278).

Return of stock holder’s equity is highest for SWC14 (118.2106) followed by SWC2 (30.7891) and SWC9 (28.4322). Whereas SWC1 is having a negative Return of Stock holder’s equity of -0.5958 bottoms the last in the chart.

Another important aspect of financial ratios is Quick ratio. Company SWC12 is the highest among all the others with a highest average of 4.9653 tops the lead followed by SWC9 and SWC3 with a Quick ratio’s of 4.3278 and 4.1768 respectively, Whereas SWC1 bottoms the chart with a least average ratio of 0.4487.

Company SWC9 is having a highest Operating income ratio (0.3908) among all other companies, alongside with that, SWC11 and SWC14 takes the 2nd and 3rd lead positions with OIR of 0.3149 and 0.2773 respectively, while SWC1 is bottom the chart with an average of negative Operating income ratio of -0.0368.

In context of operating cash flow ratio, the company SWC14 is in the first position (1.0799), SWC12 is in the 2nd position (1.0348) and SWC15 is in the 3rd (0.7857) leader board position. SWC14 is having a highest (0.2648) Return of assets ratio(RAR) when compared with the rest of the companies, followed by SWC15(0.2321) and SWC2(0.1867) topped with 2nd and 3rd position, but on contrary, company SWC1 bottoms the chart with a negative RAR of -0.1191.

While considering the Market share of the companies, SWC14 tops the leader board position with highest market share of 33% followed by SWC3 and SWC17 with a market share of 19.35% and 16.06% respectively. Around 90% of total market share is held with 5 companies (SWC14, SWC3, SWC17, SWC2 and SWC16) whereas the company SWC12 market penetration is the least, with an average market share of 0.15% and bottoms the chart.

4.1 Step 1: Applying Principal component analysis –

4.1.1 Step 1.1 correlation matrix –

The table 2 gives the correlations between the original variables. From the correlation matrix it is observed that the correlation coefficients are not too high (>0.9). Using the correlation, its evident that all the nine variables are considered for further analysis

Table 2: Correlation Matrix

	Stock-holders Equity Ratio	Turnover Rate Of Accounts Receivables	Turnover Rate of Inventory	Return Of Stock-holder Equity	Quick Ratio	Operating Cash Flow Ratio	Return of Assets	Market Share	Annual Fixed Assets	
Correlation	Stock-holders	1.000	-0.029	-0.030	-0.166	-0.297	-0.235	-0.343	-0.171	-0.164

Equity Ratio										
Turnover Rate Of Accounts Receivables	-0.029	1.000	0.147	-0.066	0.273	0.181	0.029	-0.067	-0.155	
Turnover Rate Of Inventory	-0.030	0.147	1.000	0.373	0.233	0.264	0.240	0.322	0.050	
Return Of Stock-holder Equity	-0.166	-0.066	0.373	1.000	0.307	0.446	0.463	0.830	0.467	
Quick Ratio	-0.297	0.273	0.233	0.307	1.000	0.606	0.478	0.227	0.017	
Operating Cash Flow Ratio	-0.235	0.181	0.264	0.446	0.606	1.000	0.640	0.353	0.127	
Return Of Assets	-0.343	0.029	0.240	0.463	0.478	0.640	1.000	0.400	0.177	
Market Share	-0.171	-0.067	0.322	0.830	0.227	0.353	0.400	1.000	0.805	
Annual Fixed Assets	-0.164	-0.155	0.050	0.467	0.017	0.127	0.177	0.805	1.000	

4.1.2 Step 1.2: Result: Conduct KMO test –

Looking at the table 3 below, the KMO measure is 0.631, which is more than 0.5 and therefore the total sample size of 90 with a total element size of 810 can be satisfactorily accepted.

Table-3 KMO and Bartlett’s Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.631
Bartlett's Test of Sphericity	Approx. Chi-Square	396.005
	Df	36
	Sig.	0.000

4.1.3 Step 1.3: Conduct Bartlett’s Test of Sphericity –

From the same table 3, it is observed that Bartlett’s Test of Sphericity is significant (0.000). That is, significance is less than 0.05. In fact, it is actually 0.000, i.e. the significance level is small enough to reject the null hypothesis. This means that correlation matrix is not an identity matrix.

4.1.4 Step 1.4: Findout Communalities –

For instance, each of the variable’s(return of stockholder equity,quick ratio,operating cash flow ratio,return of assets,market share and annual fixed assets)variance’s explained by the principal components are having loading values more than 0.65. Whereas, stockholders equity ratio,turnover rate of accounts receivables and turnover rate of inventory extraction loadings are accounted by less than 0.65. Since there are total 9 variables, there will be nine components. And these components can be reduced by identifying the variables loadings on the component.

Table-4 Community Matrix

variables	Initial	Extraction
Stockholders Equity Ratio	1.000	0.614
Turnover Rate Of Accounts Receivables	1.000	0.489
Turnover Rate Of Inventory	1.000	0.621
Return Of Stockholder Equity	1.000	0.784
Quick Ratio	1.000	0.687
Operating Cash Flow Ratio	1.000	0.705
Return Of Assets	1.000	0.665
Market Share	1.000	0.942
Annual Fixed Assets	1.000	0.754

4.1.5 Step 1.5: Formation of Total variances –

Table-5 Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.464	38.485	38.485	3.464	38.485	38.485	2.562	28.465	28.465
2	1.723	19.140	57.625	1.723	19.140	57.625	2.290	25.444	53.909
3	1.076	11.957	69.582	1.076	11.957	69.582	1.411	15.673	69.582
4	0.832	9.239	78.821						
5	0.723	8.033	86.854						
6	0.458	5.084	91.938						
7	0.370	4.109	96.047						
8	0.308	3.425	99.472						
9	0.048	0.528	100.000						

There are 9 components extracted during a PCA analysis as there are 9 variables that are put into it. Eigenvalues are the variances of the principal components. The number of rows reproduced on the right side of the table is determined by the number of principal components whose eigenvalues are 1 or greater. In this case 3 components are generated based on the Eigen value which is greater than 1. And hence only 3 components are selected and 69.582 % of variance is being explained by them.

4.1.5.1 Step 1.5.1: Screen plot –

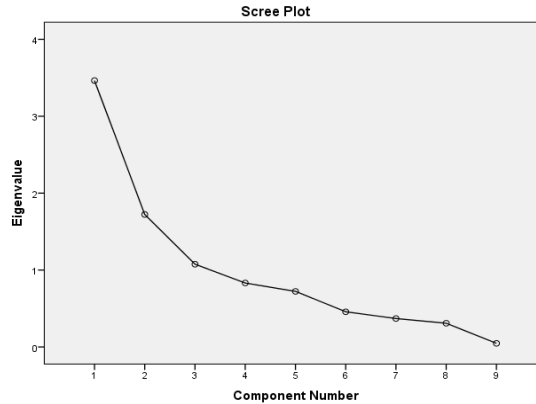


Figure 2. Scree Plot

The scree plot is a graph of the eigenvalues on the y-axis against all the factors/components on the x-axis. The graph is useful for determining how many factors to retain. The point of interest is where the curve starts to flatten. It can be seen that the curve begins to flatten between factors 3 and 4. Note also that factor 4 onwards have an eigen value of less than 1, so only three factors have been retained.

4.1.6 Step 1.6: Formulate component matrix –

Table-6 Component Matrix

	Component		
	1	2	3
Stockholders Equity Ratio	-.400	-.184	.648
Turnover Rate Of Accounts Receivables	.083	.576	.388
Turnover Rate Of Inventory	.460	.126	.628
Return Of Stockholder Equity	.821	-.289	.166
Quick Ratio	.605	.562	-.076
Operating Cash Flow Ratio	.728	.415	-.052

Return Of Assets	.735	.266	-.233
Market Share	.821	-.503	.124
Annual Fixed Assets	.549	-.668	-.080

There are 3 components extracted as the 3 components had an Eigen value greater than 1. Almost 5 variables have higher correlation but not have sufficient details and rotated component matrix give a better clear picture of component loadings.

4.1.7 Step 1.7: Formulate Rotated Component Matrix –

Looking at the table below, it is observed that return of stockholder equity, market share and annual fixed assets are substantially loaded on component 1) while quick ratio, operating cash flow ratio and return of assets are substantially loaded on component 2) all the remaining variables (stockholders equity ratio, turnover rate of accounts receivables and turnover rate of inventory) are not substantially loaded on any component.

Previous studies have only considered the practical significance of the loadings without any reference to statistical significance, the number of factors or the number of variables. Only loadings of $|\geq 0.70|$ [or greater] were considered significant in Pinches et al. (1973, 1975), Johnson (1979), and Ezzamel (1987a), amongst others. The current study, however, uses the statistical significance guidelines given in Hair et al. (1998) [9]; it also considers the number of components, variables and sample size when judging the significance of factor loadings.

Thus we are considering 0.72 as the highest possible value to differentiate the value set between the variables.

Table-7 Rotated Component Matrix

	Component		
	1	2	3
Stockholders Equity Ratio	-.056	-.702	.343
Turnover Rate Of Accounts Receivables	-.287	.164	.616
Turnover Rate Of Inventory	.326	.040	.716
Return Of Stockholder Equity	.801	.293	.240
Quick Ratio		.742	.369
Operating Cash Flow Ratio	.212	.730	.357
Return Of Assets	.295	.748	.138
Market Share	.944	.198	.106
Annual Fixed Assets	.840	.038	-.220

PCA The idea of rotation is to reduce the number factors on which the variables under investigation have high loadings. Rotation does not actually change anything but makes the interpretation of the analysis easier. Looking at the table below, it is observed that return of stockholder equity, market share and annual fixed assets are substantially loaded on component 1) while quick ratio, operating cash flow ratio and return of assets are substantially loaded on component 2) all the remaining variables (stockholders equity ratio,turnover rate of accounts receivables and turnover rate of inventory) are not substantially loaded on any component.

4.2 Step 2: Applying Canonical Correlation Analysis –

4.2.1 Step 2.1: Classification of dependent and independent variables –

In the principal component analysis, total 9 variables are divided into 6 dependent/output variables (return of stockholder equity,quick ratio,operating cash flow ratio,return of assets,market share,annual fixed assets) and 3 independent variables (stockholders equity ratio ,turnover rate of accounts receivables,turnover rate of inventory) based on the result obtained by principal component analysis.

4.2.2 Step 2.2: Multivariate test –

Table-8 Multivariate Test

Test Name	Value	Exact F	Hypoth.DF	Error.DF	Sig.of.F
Pillais	0.81057	57.76522	6	81	0.00
Hotellings	4.27891	57.76522	6	81	0.00
Wilks	0.18943	57.76522	6	81	0.00
Roys	0.81057				

In this case from the above table the value of P value for all above tests proved to be 0.00 and hence there is an evidence that there exists maximum correlation and also orthogonal succession of correlations are different from zeros.

4.2.3 Step 2.3: Determination of number of roots for the given set of data –

In earlier case, Using PCA, all variables divided into 2 sets of variable. Out of which 6 variables (dependent variables) grouped as one set and 3 are independent variables of other set. The number of canonical roots are considered based on the set with least number of variable. In our case, the canonical roots are 3.

4.2.4 Step 2.4: determine Canonical correlation coefficients and the eigenvalues –

The table reports the canonical correlation coefficients and the eigenvalues of the canonical roots. The first canonical correlation coefficient is .0.518 with an explained variance of the correlation of 53.84% and an eigenvalue of 0.367. This implies that, how much is the the canonical correlations as well how much variance of the dependent variables is explained by the dimensions.

Table-9 Canonical Correlation Coefficients and Eigen values

Root No	Eigenvalue	Pct.	Cum. Pct.	Canon Cor.	Sq. Cor
1	0.367	53.840	53.840	0.518	0.268
2	0.170	24.966	78.807	0.381	0.145
3	0.144	21.193	100.000	0.355	0.126

4.2.5 Step 2.5: test the significance of each root –

The table * shows the test of the significance of each of the roots. It was found that of the three possible roots all the first root is significant with $p < .05$. Since the model contains six dependent variables (return of stockholder equity, quick ratio, operating cash flow ratio, return of assets, market share and annual fixed assets) and three independent variables. SPSS extracts three canonical roots or dimensions. The first test of significance tests all three canonical roots of significance, the second test excludes the first root and tests roots two to three, the last test tests root three by itself.

Table-10 Test of Significance

Roots	Wilks L.	F	Hypoth. DF	Error DF	Sig. of F
1 TO 3	0.54646	3.03806	18	229.59	0.000
2 TO 3	0.74686	2.5769	10	164	0.006
3 TO 3	0.87386	2.99528	4	83	0.023

4.2.6 Step 2.6: Determine raw canonical coefficients for dependent variables –

The raw canonical coefficients are interpreted in a manner analogous to interpreting regression coefficients i.e., for the variable return of stockholder equity, a one unit increase in return of stockholder equity leads to a 1.098 increase in the first canonical variates of set 1 when all of the other variables are held constant.

For the variable quick ratio, a one unit increase in quick ratio leads to a 2.393 decrease in the first canonical variate of set 1 when all of the other variables are held constant. For the variable operating cash flow ratio, a one unit increase in operating cash flow ratio leads to 0.599 decrease in the first canonical variate of set 1 when all of the other variables are held constant. For the variable return of assets, a one unit increase in return of assets leads to a 1.146 decrease in the first canonical variate of set 1 when all of the other variables are held constant. For the variable market share, a one unit increase in market share leads to 4.925 decrease in the first canonical variate of set 1 when all of the other variables are held constant. For the variable annual fixed assets, a one unit increase in annual fixed assets leads to a 3.129 increase in the first canonical variate of set 1 when all of the other variables are held constant.

Table-11 Raw Canonical Coefficients for Dependent Variables

Variable	1	2	3
Return Of Stockholder Equity	1.098	-3.316	4.392
Quick Ratio	-2.393	-1.869	-4.412
Operating Cash Flow Ratio	-0.599	1.991	-1.904
Return Of Assets	-1.146	-5.818	1.894
Market Share	-4.925	7.773	-0.264
Annual Fixed Assets	3.129	-6.285	0.257

4.2.7 Step 2.9: *determination of raw canonical coefficients for independent canonical variables (note: covariates) –*
 The raw canonical coefficients are interpreted in a manner analogous to interpreting regression coefficients i.e., for the variable stockholders equity ratio , a one unit increase in stockholders equity ratio leads to a 3.292 increase in the first canonical variate of covariate set (set 2) when all of the other variables are held constant.

Similarly, for the variable turnover rate of accounts receivables, a one unit increase in turnover rate of accounts receivables leads to a 1.324 decrease in the first canonical variate of covariate set (set 2) when all of the other variables are held constant. For the variable turnover rate of inventory, a one unit increase in turnover rate of inventory leads to a 4.256 decrease in the first canonical variate of covariate set (set 2) when all of the other variables are held constant.

Table-12 Raw Canonical Coefficients for Independent Canonical Variables

Covariate	1	2	3
Stockholders Equity Ratio	3.292	6.616	1.054
Turnover Rate Of Accounts Receivables	-1.324	1.695	-5.483
Turnover Rate Of Inventory	-4.256	1.885	2.434

4.2.8 Step 2.7: *Determine Standardized canonical coefficients of dependent variables –*

For example, consider the annual fixed assets, a one standard deviation increase in annual fixed assets leads to a 0.829 standard deviation increase in the score on the first canonical variate for the dependent variable set when the other variables in the model are held constant. For return of stockholder equity, a one standard deviation increase in return of stockholder equity leads to a 0.217 standard deviation increase in the score on the first canonical variate of the dependent variable set when the other variables in the model are held constant. consider quick ratio or operating cash flow ratio or return of assets or market share, a one standard deviation increase in quick ratio or operating cash flow ratio or return of assets or market share leads to 0.388 or 0.106 or 0.166 or 1.282 standard deviation decrease in the score on the first canonical variate for the dependent variable set when the other variables in the model are held constant respectively.

Table-13 Standardized Canonical Coefficients of Dependent Variables

Variable	1	2	3
Return Of Stockholder Equity	0.217	-0.656	0.869
Quick Ratio	-0.388	-0.303	-0.715
Operating Cash Flow Ratio	-0.106	0.353	-0.338
Return Of Assets	-0.166	-0.841	0.274
Market Share	-1.282	2.024	-0.069
Annual Fixed Assets	0.829	-1.665	0.068

4.2.9 Step 2.10: *Determine Standardized canonical coefficients for independent canonical variables(covariates) –*

The standardized canonical coefficients are interpreted in a manner analogous to interpreting standardized regression coefficients. For example, consider the stockholders equity ratio, a one standard deviation increase in stockholders equity ratio leads to a 0.441 standard deviation increase in the score on the first canonical variate for the covariate variable set when the other variables in the model are held constant. Similarly, consider the Turnover rate of accounts receivables or turnover rate of inventory, a one standard deviation increase in turnover rate of accounts receivables or turnover rate of inventory leads to a 0.227 or 0.819 respective standard deviation decrease in the score on the first canonical variate for the covariate variable set when the other variables in the model are held constant.

Table-14 Standardized Canonical Coefficients for Independent Canonical Variables

Covariate	1	2	3
Stockholders Equity Ratio	0.441	0.887	0.141
Turnover Rate Of Accounts Receivables	-0.227	0.291	-0.941
Turnover Rate Of Inventory	-0.819	0.363	0.469

4.2.10 Step 2.8: *Determine Correlations between dependent and canonical variables –*

For example, it can be seen in the "dependent" variables that return of stockholder equity has a Pearson correlation of -0.703 with the first variate, -0.080 with the second variate, and 0.600 with the third variate. Except annual fixed

assets (-0.151) other dependent variables show high negative correlations (return of stockholder equity-0.703, quick ratio-0.741, operating cash flow ratio-0.698, return of assets-0.685 and market share 0.626) with first canonical variate.

Table-15 Correlations between Dependent and Canonical Variables

Variable	1	2	3
Return Of Stockholder Equity	-0.703	-0.080	0.600
Quick Ratio	-0.741	-0.262	-0.536
Operating Cash Flow Ratio	-0.698	-0.158	-0.224
Return Of Assets	-0.685	-0.548	0.103
Market Share	-0.626	-0.142	0.535
Annual Fixed Assets	-0.151	-0.451	0.412

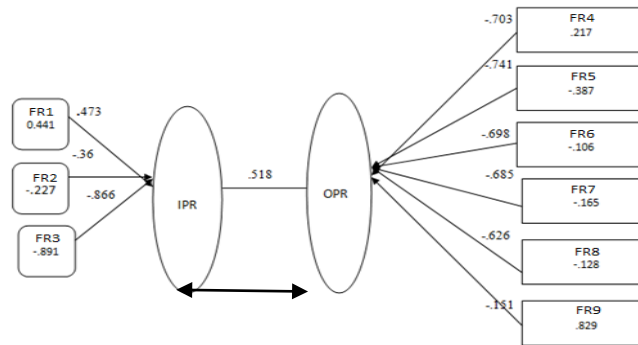
4.2.11 Step 2.11: Determination of Correlations between Independent canonical variables (Covariates) and dependent canonical variables (canonical variables) –

These are the correlations between each variable in a covariate group and the group’s canonical variates. For example, it can be seen in the independent variable that stockholders equity ratio has a Pearson correlation of 0.473 with the first variate, 0.868 with the second variate, and 0.154 with the third variate. Turnover rate of inventory variables show high negative correlation of -0.866 with first canonical variate.

Table-16 Correlations between Covariates and Canonical Variables

Covariate	1	2	3
Stockholders Equity Ratio	0.473	0.868	0.154
Turnover Rate Of Accounts Receivables	-0.360	0.319	-0.877
Turnover Rate Of Inventory	-0.866	0.379	0.326

V. FINAL CONCLUSION



FR1=Stockholders equity ratio, FR2= Turnover rate of accounts receivables, FR3=Turnover rate of inventory, FR4=Return of stockholder equity, FR5=Quick Ratio, FR6=Operating cash flow ratio, FR7=Return of Assets, FR8=Market Share, FR9= Annual Fixed assets

Figure 3. Relation between Covariates and Canonical Variables through Canonical root 1

Figure 3 shows a pictorial diagram which displays associations of the two sets of variables. It is meant for easy visualization of the input financial ratios and output financial ratio model. Out of the 3 roots, the First canonical correlation root is sufficient to consider the relationship between the set of variables, thus omitted the rest of the roots for explanation.

This coefficient has the same properties as any other correlation: it ranges between minus one and one, a value near zero indicates low correlation, and an absolute value near one indicates near perfect correlation.

The strength of the relationship between the pairs of variates is reflected by the CCA coefficient (Rc). For the first function, Rc = 0.518 and the first canonical function is significant. (Alissa Sherry and Robin K. Henson(2005) and TandanaiRwizi (2015).

The single one directional arrow indicates a correlation of variables with the respective canonical variates. In this case, stockholders equity ratio (stock holders equity ratio) shows positive correlation (0.473) with input ratio (IPR). Where as, turnover rate of accounts receivables and turnover rate of inventory having a negative correlation (-.36 and

-.866) with Input ratio respectively. Similarly, return of stockholder equity have a negative correlations of -.703, with Output ratio(OPR). In the same fashion rest of the variables quick ratio, annual fixed assets,return of assets,market share and annual fixed assets are also negatively correlated.

The double arrow represents a reciprocal inter-relationship between input ratios and output ratios. In this case the canonical correlation coefficient of 0.518 indicates that there is a moderate relationship present between input ratios and Output ratios.

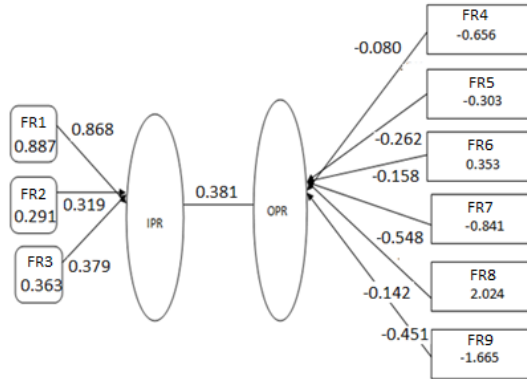


Figure 4.Relation between Covariates and Canonical Variables through Canonical root 2

Figure 4 shows a pictorial diagram which displays correlation of the variables with the 2nd set of canonical variates. In this case, stockholders equity ratio (0.868), turnover rate of accounts receivables(0.319) and turnover rate of inventory(0.379) having a positive correlation with input ratio (IPR). Whereas, return of stockholder equity, quick ratio,annual fixed assets,return of assets,market share ad annual fixed assets are negatively correlated with output ratio (OPR). In this case the canonical correlation coefficient of 0.381 denotes that there is a weak relationship between input ratios and out put ratios exists.

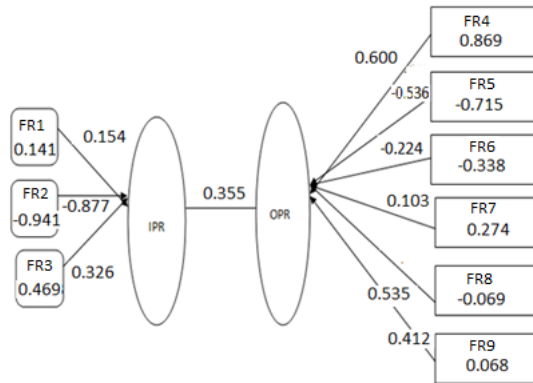


Figure 5.Relation between Covariates and Canonical Variables through Canonical root 3

Similarly, Figure 5 shows a correlation of the variables with 3rd set of canonical variates. Where, stockholders equity ratio (0.154) and turnover rate of inventory (0.326) having a positive correlation with IPR, and turnover rate of accounts receivables(-0.877) variable have negative correlation with IPR. Similarly, quick ratio (-0.536) and operating cash flow ratio (-0.224) have a negative correlation with the OPR. Where as return of stockholder equity,return of assets,market share ad annual fixed assets are positively correlated with the OPR. The canonical correlation coefficient of 0.355 for the canonical root 3 indicates that there is a weak correlation between IPR and OPRexists.

VI. CONCLUSION

PCA was carried out on the 9 selected financial ratios, resulting in a percentage of variance explained by the first three factors of 69.58%. Based on the Factor loading in the rotated component matrix the ratios are grouped in this case study based on the 3 factors. Accordingly, the group of 6 ratios is selected since the rotated component matrixes

of these ratios are showing more than 0.725. Thus Return of Stockholder Equity, Quick Ratio, Operating Cash Flow Ratio, Return of Assets, Market Share and Annual Fixed Assets are obtained as important factors. In this case, the group of 6 ratios selected by PCA, due to their correlation with the extracted factors, and the group of 3 ratios removed from the analysis, because they were not correlated to the factors and these ones did not explain them well, so their communalities were low.

There are ratios which have been selected (first set) and other ones that have been removed from the analysis (second set). Since they measure the same issue in a different way, ratios in each group are supposed to be related, that is, they contain similar or redundant information (LeClere, 2006), which would justify the selection of the 6 ratios, to the detriment of the 3 removed ones.

Though the three canonical roots are significant ($p < 0.05$), it is observed that for the first canonical root the canonical correlation coefficient is high (0.518) and the first canonical root is considered. Also from the raw canonical coefficients of output variables for first canonical root are return of stockholder equity (1.098), quick ratio (-2.393), operating cash flow ratio (-0.599), return of assets (-1.146), market share (-4.925) and annual fixed assets (3.129). But market share shows strong impact. Similarly, from the raw canonical coefficients of input variables for the first canonical root are stockholders equity ratio (3.292), turnover rate of accounts receivables (-1.324) and turnover rate of inventory (-4.256) where turnover rate of inventory shows strong impact.

In conclusion the research made in this paper exhibit a co-relation of Input financial ratios and Output Financial ratios with canonical correlation coefficient of 0.518. It was also noticed that Stockholders Equity Ratio, turnover rate of accounts receivables and turnover rate of inventory are identified as the main components of input ratios that affects the output ratios. And also main components, namely (return of stockholder equity, quick ratio, operating cash flow ratio, return of assets, market share and Annual fixed assets are emerged as Output financial ratios as achieved. This categorization of financial ratios is useful to determine the financial efficiencies of business organizations, with the study of minimum number of financial ratios instead of considering the complete list of those 9 ratios.

VII. REFERENCES

- [1] Peter D. Hart (2017), "A Canonical Correlation Analysis of Physical Activity Parameters and Body Composition Measures in College Students", *American Journal of Sports Science and Medicine*, Vol. 5, No. 4, pp. 64-68
- [2] Beaver, W.H. (1966). "Financial ratios as predictors of failure", *Journal of Accounting Research*, Vol. 6, No. 2, pp. 179 – 192.
- [3] Altman, E. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance* 23(4): 589-609.
- [4] Wu, W. (2010), *Beyond Business Failure Prediction; Expert Systems with Applications*, 37(3): 2371-2376.
- [5] Hernández, Salvador & Gras-Gil, Ester & Antón Renart, Marcos. (2011). Financial information and restructuring of Spanish savings banks in a context of crisis. Changes in the regulation; content and evolution of FROB. *CIRIEC-España, revista de economía pública, social y cooperativa*. 99-126.
- [6] Scott, James. (1981). The probability of bankruptcy: A comparison of empirical predictions and theoretical models. *Journal of Banking & Finance*. 5. 317-344. 10.1016/0378-4266(81)90029-7.
- [7] Garcia-Gallego, A. (1996). Principal components and canonical correlation analyses as complementary tools. Application to the processing of financial information. *Economic computation and Economic Cybernetics studies and Research*, Issue 4/2016, Vol.50
- [8] Chiu, Ching-Yi & Wang, Mei-Wei (2011). "An integrated DEA-based model to measuring financial performance of construction companies", *WSEAS transactions on business and Economics*, Vol.8.
- [9] Yang, Lan. (2017). "Research on application of principal component statistical analysis in the financial early-warning", *Research on Modern Higher Education*. Vol. 3, PP. 61-68. 10.24104/rmhe/2017.03.01010.
- [10] Comrey, A. L. & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Lawrence Erlbaum, 2nd edition. (1st edition Comrey alone: 1973, New York: Academic Press).
- [11] Hair, J.F., Tatham, R.L., Anderson, R.E. and Black, W., 1998. *Multivariate Data Analysis*. 5th ed. Upper Saddle River (NJ): Prentice-Hall.
- [12] Mardia, K. V., Kent, J. T., Bibby, J. M. (1979), *Multivariate Analysis*; Academic Press Inc. Ltd.;
- [13] Jolliffe T. *Principal component analysis*. New York: Springer; 2002.
- [14] Scott, James. (1981). The probability of bankruptcy: A comparison of empirical predictions and theoretical models. *Journal of Banking & Finance*. 5. 317-344. 10.1016/0378-4266(81)90029-7.
- [15] Balcaen, Sofie & Ooghe, Hubert. (2006). 35 Years of Studies on Business Failure: An Overview of the Classic Statistical Methodologies and Their Related Problems. *The British Accounting Review*. 38. 63-93. 10.1016/j.bar.2005.09.001.