Profitability on the Financial Markets; A Discriminant Analysis Approach

S.A. Ionescu, C.S. Murgoci, C.M. Gheorghe, E. Ionescu

Abstract**—** In this paper we tried to group in three classes the companies listed without interruption for 6 years from Bucharest Stock Exchange. We used cluster analysis, namely an iterative method of clustering, the k-means algorithm. Using data results, we have made tests for the three classes of prediction using discriminant analysis. Fisher's functions have helped us to make predictions on the affiliation of a new listed company on one of the 3 classes of risk. In this study, emphasis was placed on the liquidity of companies, but also on how efficient are used the raw materials, the basic elements in the current financial crisis. This should give us a clearer picture of companies that are ready to get over this difficult time.

Keywords— Discriminant analysis, Cluster analysis, Pattern recognition, Stock exchange, Portfolio analysis, Classifiers.

I. INTRODUCTION

In most human activities appears the need to surround, to make the difference, to group or classify certain objects in the form of categories whose determination must be very clear and very natural.

Research aimed at structuring and differentiation of many items on specific categories or classes, depending on the fundamental properties of objects, are known under various names, such as grading, clustering, group or discrimination.

In general, we can say that discrimination and clustering represent activities of arranging objects, individuals or observations, in the form of groups, classes or categories, depending on the degree of similarity or contrast between them.

The overall aim of the pattern recognition theory is identifying at the level of many complex and heterogeneous forms or objects of structures, groups, classes or clusters existing at the latent level in that crowd and those who shape in a natural way, depending on the similarities and differences between many of these items.

II. USING DISCRIMINANT ANALYSIS AS A PATTERN RECOGNITION TECHNIQUE.

Frequently in data analysis appears the necessity of studying populations that are *heterogeneous* in terms of features analyzed, which complicate the understanding process of knowing these populations and requires making a specific scientific approach.

The most significant expression of a heterogeneous population is found in particular in statistics, data analysis and econometrics, being represented even by very large quantities of information to be processed, summarized and interpreted.

In case of the survey of populations like this, in order to have consistency and relevance the investigation results, it is necessary to make a division of these populations in sub-populations with a certain degree of homogeneity, following that the analysis and modelling process involved in the study to be made differentiated for each sub-population basis.

Formulation of an accurate and robust conclusion on the expression of the populations characterized by a greater or less heterogeneity is possible only if the analysis takes into account the population structure by categories.

In other situations, such as those that are analyzed in various economic and social entities, considered to come from populations with very different characteristics, there is the interest to identify and to recognize the origin of these entities, and to obtain a correct classification of these in certain classes representative for the population of origin.

Situations like these exceed economic and financial sphere, they are meeting frequently in a wide variety of other important areas of science such as computer science, biology, anthropology, medicine, sociology, geology, meteorology, etc.

In the field of economic-financial, entities that are subject to problems of determining affiliation to a group or class can be firms, customers of banks, buyers of a product, the administrative-territorial units, markets, goods or services, etc.

The general setting procedure, based on defining characteristics and using specific methods and techniques, about some objects belonging to certain groups or classes known before is called *discriminant analysis. Discriminant analysis* represents the process of using a variety of methods, techniques and algorithms to determine which features of certain objects have the greatest relevance in terms of recognition of these objects belonging to certain aprioric classes and to define and set most likely belonging of objects to the different classes.

Establishment of objects belonging to a population to specific classes is based upon properties or characteristics of objects, which are represented at formal through *variables*. Variables from the optimal set of features are called *descriptor variables* and they can be represented either by the whole set of variables that describe the objects, or just by a subset of it. This means that the multitude of descriptor variables is a set verifying the relationship:

$$\{x_1, x_2, \dots, x_n\} \subseteq \{x_1, x_2, \dots, x_N\}. \qquad (1)$$

Descriptor variables are not used in the classification itself, as such, but in a transformed form, represented by discriminant variables. The criteria to be deducted for the separation of classes of the population analysis are used to build equations or functions, which define *points, curves or surfaces* separating these classes.

Equations and functions used to separate classes and are known as the *classifiers.* Functions based on which is the separation of classes are called *discriminant functions, classification functions or score functions*, and are defined in relation to the *descriptor variables* of objects and they are used in the determination of new variables, called *discriminated variable* or *score variables*.

As we see below, the overwhelming majority of cases the use of discriminate analysis, discriminatory functions are *linear functions* of the form:

$$d_i = \beta_1^{(i)} \cdot x_1 + \beta_2^{(i)} \cdot x_2 + \dots + \beta_n^{(i)} \cdot x_n = \beta^{(t)} \cdot x \qquad (2)$$

Where i=1,2,...p.

Number of discriminant features, namely p, is determined by the number of descriptor variables and the number of classes existing in the population studied.

A. Defining the discriminant analysis problem

Seen in a very general way, solving a problem of classification using discriminant analysis assumes the deduction of rules or criteria so that, after knowing the vector x of properties of an object belonging to a population Ω, it can take a decision on the classification of that object in one of the possible K classes as that can be structured population Ω.

Fundamental assumption of the discriminant analysis is that the Ω set is composed by heterogeneous elements and that, by default, in the Ω set there are a number of K classes, marked with $\omega_1, \omega_2, \dots, \omega_k$ and called *real classes* or *initial classes*, whose composition is not completely known and who have the following properties:

$$\omega_k \subset \Omega, \qquad k = 1, 2, \dots, K \qquad (3)$$

$$\omega_1 \cup \omega_2 \cup \dots \cup \omega_K = \Omega \qquad (4)$$

Recall that, in general, the initial classes of the Ω set are considered to be not disjunct, meaning there is the possibility that:

$$\omega_i \cap \omega_j \neq \Phi. \qquad (5)$$

Statistical distribution of objects in each real class ω_k is described using multidimensional densities of conditional probability of classes, ie with functions ω , whose form is assumed to be known.

In discriminant analysis, most important, both from the theory and the practice, are no real classes, but classes of prediction, which we define below. We consider the Ω set and real classes $\omega_1, \omega_2, \dots, \omega_k$ from who it is initially formed. The main purpose of the discriminant analysis is to identify an efficient way of structuring the Ω set in a number of K classes or regions. The second property is that any object in the Ω set should be classified. Third property require that any objects of the set to be classified *only in one class*.

A classification can be considered as perfect, that is unaffected by the errors, if and only if there is a perfect coincidence between any class of prediction $\widetilde{\omega}_k$ and real class counterparts ω_k. This is not always possible because of the consequences which they entail ownership of disjunctive classes of prediction.

As you can see, unlike the real classes $\omega_1, \omega_2, \dots, \omega_k$, which may have some overlapping, classes of prediction $\widetilde{\omega}_1, \widetilde{\omega}_2, \dots, \widetilde{\omega}_k$ must be disjunctive two by two, that is to not have objects in common. Because classes of prediction are disjunctive two by two, they appear to be some fragmentations of real classes, which means that they can be regarded as being defined in the form of restrictions imposed on the real classes. As a result of the real counterparts classes fragmentation, classes of prediction will differ more or less of real classes, so that, between a class of prediction and a real counterparts class it will be the relationship:

$$\widetilde{\omega}_k \neq \omega_k, \qquad k=1,2,\dots,K \qquad (6)$$

The differences that exist between classes of prediction and real classes, resulted from the fact that a class of prediction is achieved by a real class fragmentation, represents the possibility expression that certain objects could be classified incorrectly.

An immediate consequence of the way classes of prediction are defined is that each prediction class is, in fact, a real counterparts class subset, namely:

$$\widetilde{\omega}_k \subseteq \omega_k, \qquad k=1,2,\dots,K \qquad (7)$$

On the other hand, because the classes of prediction should include all the Ω set objects,verify:

$$\widetilde{\omega}_1 \cup \widetilde{\omega}_2 \cup \dots \cup \widetilde{\omega}_K = \omega_1 \cup \omega_2 \cup \dots \cup \omega_K = \Omega \qquad (8)$$

Under these circumstances, it is obvious that as long as each real class is a full-field events, any class of prediction, which is a real counterparts class subset, appears to be an incomplete field of event.

III. The Separation of Classes in Forms Space

The first and most difficult problem to be solved in the discriminant analysis is the separation of classes in the prediction of the Ω set. The most direct way of separating Ω set classes is represented by *defining the space of separation surfaces or decision surfaces*. These areas of separation are those which cause the offset of the classes of prediction $\widetilde{\omega}_1, \widetilde{\omega}_2, ..., \widetilde{\omega}_k$ and it pass, necessarily, by the set of objects belonging to the intersection of the classes that separate them.

For reasons like simplification of the classification process, usually, there are used linear separation surfaces, like straight lines, planes or hyperplanes[1].. Separation surfaces are defined by functions known as discriminant functions. The information utilized for construction of separating areas of prediction classes are represented by a sample volume of T objects extracted from the population Ω, objects whose affiliation of classes is accurately known $\widetilde{\omega}_1, \widetilde{\omega}_2, ..., \widetilde{\omega}_k$.

Finding an effective way to separate the set elements on disjunctive classes is a difficult problem, especially because of the existence in the set Ω of some objects that belong simultaneously to two different real classes. Affecting of this kind of objects to a class or another could be possible only through the probabilistic calculus

The main problem to be solved in the discriminant analysis is that of constructing criteria or rules of classification, and based on it, we can make predictions about affiliation of new forms, with initially unknown affiliation.

Criteria for classification are known as *classifiers*, and the deduction of these criteria is called *training* of the classifier. The classifier is actually an algorithm which determines the most likely affiliation of a form to a prediction class. The training of the classifier is based on the information contained in a sample form whose affiliation is known before and which is called *training set*.

The sample that represents the training set is extracted from the population and has been analyzed and it contains the primary data used in any discriminant analysis. In some situations, for the training of the classifier can be used, effectively, only one part of the sample available, the other part being used for testing and validating the ability of the classifier obtained through the training set, to properly classify forms whose affiliation is known. In this way, the training set could represent only a part of the sample available. Part of the sample used for testing and validating the power of discrimination of the classifier is called *set of prediction*. Often, the entire sample available can be used both as a set of formation, and the set of prediction, which means that the two sets may coincide.

There are several ways to approach that can be used for the classifier formation. We can mention: criterion of minimizing the cost of classification, Bayes criteria or criteria of the posterioric probabilities, Fisher's criteria of the linear discriminant functions, metric criteria or Mahalanobis distance criteria, the verisimilitude report criteria etc.. Using each of the criteria mentioned before, establish a new classifier, whose essence is the same for the majority of the criteria mentioned.

A. Linear Classifier

The first way of approaching problems with classification of discriminant analysis techniques dates from 1933 and it was proposed by Fisher. Subsequently such approaches have developed steadily, and applications based on discriminant analysis were extended to even more areas of activity and have diversified increasingly more.

Most of them and the most useful application of discriminant analysis based on Fisher's criterion are met in the financial-banking field, area in which these kind of techniques are called *credit-scoring techniques* and they are the most important tools to substantiate decisions on granting loans.

Method of discriminant analysis proposed by Fisher is a parametric method, characterized by simplicity and robustness, and offers possibilities of interpretation very useful for analysis. The simplicity of this method stems from the fact that using it does require only the evaluation of estimations of population and its classes parameters, parameters represented by averages, variants or covariants. This is a very important advantage of Fisher's discriminant analysis, in comparison, for example, with Bayes's techniques, whose use involves knowing of the aprioric probabilities.

The theoretical basis of Fisher's discriminant analysis is represented by the variant analysis. Fisher's Criterion defines a way to deduct the discriminant functions on the basis of comparative analysis between intragroup variability and intergroup variability, at the level of classes or analyzed population groups. The discriminant functions deducted on the basis of Fisher's criteria are called also *score functions* and they are *linear* functions.

As we mentioned before, the fundamental criterion which on is based the division set Ω in subsets $\omega_1, \omega_2, ..., \omega_k$ is a mixed criterion, which aims to *minimize the intragroups variability and maximize intergroups variability.* Using this combined criterion provide the best differentiation of classes or of population groups $\mathbf{\Omega}$. The idea behind Fisher's criterion is the determination of directions or axes, so that, the classes of $\mathbf{\Omega}$ set to be differentiate as much as between them and, at the same time, each class to have a higher degree of homogeneity. In other words, Fisher's criterion is to determine some of the directions along which intergroup variability is higher and intragroup variability is the smaller. Projections of objects on the axes defined

From a certain point of view, the discriminant analysis can be considered as similar to principal components analysis, which aims to identify general axes relative to the variability of objects to be maximum[2]. The main difference between discriminant analysis and principal components analysis is related to the fact that in principal components analysis the causal space is considered in its entirety, without making any differentiation between its elements in terms of a specific criteria.

In case of principal components analysis, the variability is viewed as a general characteristic of the population analyzed, without taking into account the existence of any structure on this population group or class. Consequently, the variability which is the object of principal components analysis is considered as a whole, without any possibility of decomposition in relation to a certain structure of causal space analyzed.

In contrast, in case of discriminant analysis it is considered that population is divided into groups or classes, and the variability of this population can be split in two components: intergroup variability and intragroup variability.

In addition to the difference mentioned before, in the discriminant analysis the new directions to be identified should not necessarily be orthogonal, unlike principal components analysis in which the directions of maximum variability should check the orthogonal property.

The most important issue of the Fisher's criterion of discrimination between classes of a population is related to the decomposition of variability of this population[3]. We will detail how to split the population variability in relation to the two meanings of it: *simple variability* ¬ expressed through the total amount of square deviations and *mixed or composed variability* ¬ measured through mixed products of deviations matrix. It is obvious that mixed variability can be defined only for multidimensional objects.

As we mentioned before, the analysis of the discriminant function is equivalent to finding directions, or vectors, in relation to the intragroup variability to be minimal, and the intergroup variability to be maximal. These directions will define the discriminant space axes and they can be identified in the form of linear combinations of descriptive variables selected for analysis.

Therefore, the procedure for building a discriminant function is reduced to establishing the vector β, that is weighting β_1, β_2,.. β_n. We need to specify that the linear nature of the function is imposed as initial assumption and it should not be seen as resulting from the imposition of a specific performance criterion on separability classes.

B. Defining Fisher's discriminant functions

We presented earlier the way it can be inferred a Fisher discriminant function. The criterion on which it was inferred discriminant function of this type is a mixed criterion, which aims simultaneously in two aspects: minimizing intragroups variability and maximizing intergroup variability.

A Fisher discriminant function is determined as a linear combination of discriminant variables, whose combination coefficients are components of the eigenvector of the matrix $\sum_w^{-1} \cdot \sum_b$. From this way of defining, result, by default, that it can be identified more discriminant functions. The maximum number of possible discriminant functions that can be identified on Fisher's is equal to the number of the distinct and strictly positive values of the matrix $\sum_w^{-1} \cdot \sum_b$.

Since this matrix has the size $n \times n$, when it is strictly positive defined and it has the maximum rank, the result is that the total number of discriminant functions that can be determined is equal to ″n″. We will present next the way it can be determined all possible discriminant functions.

For this we will note the ″n″ values of the matrix $\sum_w^{-1} \cdot \sum_b$ with $\lambda_1,\lambda_2,\ldots,\lambda_n$ and we will assume that they are ordered in terms of values that they have as follows:

$$\lambda_1 > \lambda_2 > \ldots > \lambda_n > 0 \qquad (9)$$

We note with $\beta^{(1)},\beta^{(2)},\ldots,\beta^{(n)}$, its own ″n″ vectors of the matrix $\sum_w^{-1} \cdot \sum_b$, associates, in order, with their own values $\lambda_1,\lambda_2,\ldots,\lambda_n$. The first discriminant function is defined using the vector itself, which corresponds to the higher own value, and has the following form:

$$D_1(x) = \beta_0^{(1)} + \beta_1^{(1)} \cdot x_1 + \beta_2^{(1)} \cdot x_2 + \ldots + \beta_n^{(1)} \cdot x_n \qquad (10)$$

Since this function corresponds to the highest possible value of the report between the intergroup variant and intragroup variant, it provides the best separability of the classes, in terms of mixed criterion mentioned above. This means that the object projections on the new axe determined by the vector of coefficients $\beta^{(1)}$ can be separated into classes that differentiate in the greatest degree possible and that has the highest possible degree of homogenity.

Similarly, the second discriminant function is defined using the eigenvector which corresponding to the second eigenvalues, namely:

$$D_2(x) = \beta_0^{(2)} + \beta_1^{(2)} \cdot x_1 + \beta_2^{(2)} \cdot x_2 + \ldots + \beta_n^{(2)} \cdot x_n \qquad (11)$$

Being determined on the basis of the second eigenvalues of the matrix $\sum_w^{-1} \cdot \sum_b$, this discriminant function corresponds to a smaller value of the report between the intergroup variant and intragroup variant. Consequently, it provides a smaller resolution in terms of separability leave of the set. From this point of view, it is possible that projections of objects on the new axe which has the vector as support to match the classes that are less homogeneous and differentiate less between them.

Finally, with eigenvector associated with the lower eigenvalue, that is vector $\beta^{(n)}$, it determines the last discriminant function, namely:

$$D_n(x) = \beta_0^{(n)} + \beta_1^{(n)} \cdot x_1 + \beta_2^{(n)} \cdot x_2 + \ldots + \beta_n^{(n)} \cdot x_n \qquad (12)$$

By comparison with other discriminant functions, this latter discriminant function ensures the poorest separability between classes of the Ω set. The power of separability from low to lower that have discriminant functions $d_1,d_2,\ldots,d_n$, leads to the idea of the need to select for analysis only a certain number of discriminant functions, in order of their power of discrimination.

The effective number of discriminant functions that must be retained in the analysis depends directly on the number of classes and the number of discriminant variables.

Discriminant functions (Fisher) are linear combinations of descriptive variables of the form:

$$D(x) = \beta_0 + \beta^t \cdot x \tag{13}$$

where x is variables descriptor vector and β is the eigenvector of the matrix $\sum_w^{-1} \cdot \sum_b$. Discriminant functions values are called *discriminant scores.*

Discriminant variables are linear combinations of descriptive variables of the form:

$$d = \beta_0 + \beta^t \cdot x \tag{14}$$

where x and β have the significance of the previous definition. The average and the variance of discriminant variables are:

$$E(d) = \beta_0 + \beta^t \cdot \mu \tag{15}$$

$$Var(d) = \beta^t \cdot \Sigma \cdot \beta = \beta^t \cdot \Sigma_W \cdot \beta + \beta^t \cdot \Sigma_b \cdot \beta \tag{16}$$

Once discriminant functions were estimated, they can be used to make predictions about the affiliation of new objects to classes of prediction.

IV. The Bucharest Stock Exchange Case

An analysis of 45 companies listed permanently on Bucharest Stock Exchange (BSE) in 2002-2006 has already been done[4]. In an attempt to group them using cluster analysis techniques, it were used hierarchical clustering methods, like single linkage method or Ward 's method and iterative methods of clustering. The results were satisfactory for k-means algorithm, an iterative refinement heuristic, invented in 1956 [5] and later developed by Loyd [6]. K means algorithm is an algorithm that tries to group n objects in k clusters, where k<n. The objective is to minimize total intracluster variance or to maximize the intercluster variance.

A. Financial Ratio Analysis

For each case were used 8 variabiles, defined as:

1. Liquidity ratios

The liquidity of a company is measured by its ability to satisfy its short-term obligations as they come due. Liquidity refers to the solvency of the company's overall financial position – the ease with which it can pay for its bills. Because a common precursor to financial distress and bankruptcy is low or declining liquidity, these ratios should be viewed as good leading indicators of cash flow problems. The most used liquidity ratios are current ratio, quick ratio, cash ratio and net working capital to total assets ratio.

1.1. Current ratio

This ratio measures a company's general ability to meet its short-term obligations. It is computed as follows:

$$\text{Current ratio} = \frac{\text{Current assets}}{\text{Current liabilitie s}} \tag{17}$$

Generally, a higher value of this ratio shows a higher liquidity for the company. Although a benchmark value of 2.0 for this ratio is considered acceptable, this should not be regarded as a general rule, since the value of the current ratio is dependent on the industry. Of a higher relevance is considered the relationship between this ratio and the company's ability to predict its cash flow: the higher the predictability power of the company regarding the future cash flows, the lower the value it can have for the current ratio.

2. Solvability ratios

Solvability ratios are related to the debt position of a firm. More specifically, the long-term debt of a firm is of interest, since this commits the company to a stream of payments over the long run. In general, the more debt a firm uses in relation to its total assets and equity, the greater its indebtedness and, consequently, the greater its financial leverage. Therefore, a higher proportion of debt indicates greater financial risk.

There are two general types of solvency measures: measures of the degree of indebtedness and measures of the company's ability to service its debts. The degree of indebtedness measures the amount of debt relative to other significant balance sheet items – typically total assets and equity, the most used ratios being the long-term debt ratio, the debt-to-equity ratio and the total debt ratio. On the other hand, the ability to service debt reflects the company's ability to make the payments required on a schedules basis over the life of a debt. The ratio that reveals the company's ability to service its debt is times interest earned or coverage ratio.

2.2. Debt-to-equity ratio

This is a complement ratio to the previous one, and shows the importance of debt financing as opposed to equity financing. The higher the value of this ratio is, the higher the weight of debt as financing tool for the company, and the higher the risk of financial distress. At the same time, and in a similar manner to the previous ratio, a low value point out a under-use of debt, with the direct consequence of a lower rate of return for the owners of the business. The ratio is computed as follows:

$$\text{DER} = \frac{\text{Long - term debt} + \text{Value of leases}}{\text{Owners' equity}} \tag{18}$$

3.3. Total debt ratio

This ratio is especially revealing for a firms that derives substantial capital from short-term liabilities. Given the low level of long-term debt for Euro Food, this ratio offers a better perspective on the degree of risk of the company than the previous ones. It is computed as follows:

$$\text{TDR} = \frac{\text{Currentliabilities} + \text{Long-termdebt} + \text{Valueof leases}}{\text{Totalassets}} \tag{19}$$

3. Efficiency ratios

These ratios measure the speed with which various accounts are converted into sales or cash – inflows or outflows. With regard to current accounts, measures of liquidity alone are generally inadequate because differences in the structure of a firm's current assets and current liabilities can significantly affect its "true" liquidity. It is therefore important to look beyond measures of overall liquidity and to assess the efficiency of specific current accounts, as well as the efficiency of total assets employed by a company.

3.4. Total assets turnover

This ratio indicates the efficiency with which the firm uses its assets to generate sales. It is computed as follows:

$$\text{Total asset turnover} = \frac{\text{Operating revenues}}{\text{Total assets}} \tag{20}$$

It is essential to note that the value of this ratio is extremely dependent on the industry. For example, total asset turnover range from about 1 (or even less) for large, capital-intensive industries (steel, autos, heavy manufacturing companies) to over 10 for some retailing or service operations.

4. Profitability ratios

The ratios that are included into this category indicate two aspects of profitability: (1) the rate of profit on sales; (2) the percentage return on capital employed. Generally, the higher the value of these ratios, the higher the company's profitability.

4.5. Net profit margin

This ratio relates net income (profit after taking into account financial expenses and taxes, but before dividend payments) to sales, and is commonly considered as a measure of a firm's success with respect to earnings on sales. The value of this ratio differs considerably across industries but, generally, the higher it is, the higher the firm's profitability. It is computed as follows:

$$\text{Net profit margin} = \frac{\text{Net profit}}{\text{Operating revenues}} \tag{21}$$

4.6. Return on assets (ROA)

This ratio, also called the *return on investments (ROI)*, measures the overall effectiveness of management in generating profits with its available assets. The higher the firm's return on total assets, the better. The ratio is computed as follows:

$$\text{ROA} = \frac{\text{Net profit}}{\text{Total assets}} \tag{22}$$

4.7. Return on equity (ROE)

This ratio measures the return earned on the common stockholders' investment in the firm. Generally, the higher this return, the better off are the owners. ROE is computed as follows:

$$\text{ROE} = \frac{\text{Net profit}}{\text{Equity}} \tag{23}$$

5. Indicators of Market Value

58. Earnings per Share (EPS)

The company's EPS is generally of interest to present or prospective stockholders and management. EPS represents the number of Monetary Units earned during the period on behalf of each outstanding share of common stock. EPS is calculated as follows:

$$EPS = \frac{\textit{Earnings available for common stockholders}}{\textit{Number of shares of common stock outstanding}} \tag{24}$$

B. What is new

In today's global crisis, we thought it would be interesting to add two new differentiation variables. The first would be *cash ratio*, which is the most conservative liquidity ratio, which relates only the cash items (cash and short-term financial assets) to current liabilities as follows:

$$\text{Cash ratio} = \frac{\text{Cash} + \text{Short - term financial assets}}{\text{Current liabilities}} \tag{25}$$

With regard to current accounts, measures of liquidity alone are generally inadequate because differences in the structure of a firm's current assets and current liabilities can significantly affect its "true" liquidity so the second variable is *Inventory turnover*. This ratio indicates the efficiency with which the firm uses its inventory. At the same time, the implied ratio indicates the processing time of the company's products. Generally, the higher the value of the ratio, the higher the efficiency of using inventory, but, as in the case of the previous ratio, this ratio is strictly dependent on the industry. Also, it is important to note that a low value for this ratio is problematic, since it might be a sign of capital being tied up in inventory and problems with selling the final products. The ratio is computed as follows:

$$\text{Inventory turnover} = \frac{\text{Operating expenses}}{\text{Inventory}} \tag{26}$$

Thus, having 10 variables, we standardized and we applied again k-means algorithm. Noting that splitting into 4 classes offers 2 classes almost similar, we tried a partitioning of the objects in only 3 categories for 38 companies listed permanently between 2002-2007.

C. Results obtained

Thus, in 2007 we got the first class that is composed of Financial Investment Companies (FICs) and several other companies. Members of this class are characterized by the following aspects: solvability rates have the lowest values which indicate a low level of indebtedness, normally for financial investment companies, total debt ratio has the lowest value, return on assets and return on equity values are the most closer to 0, namely those close to the general average, so natural for these companies that own shares in several companies. Also, net profit margin and earnings per share recorded the highest values. As you see, investments in FIC sites are the most profitable.

The 2nd Cluster is a clusters of average. Net profit margin, earnings per share and total assets turnover have comparable values in standardized variables.

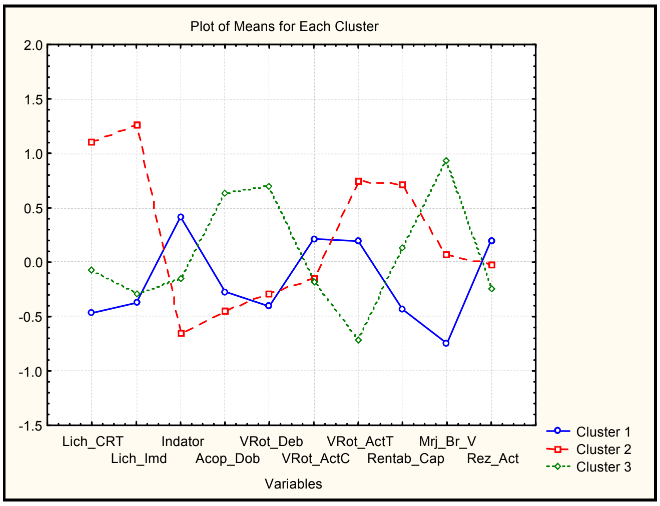

Figure 1

As the difference between the first two classes would be that the first comprises less indebted companies, which have a noticeably high liquidity, profitability and a slightly higher. In other words, we can find in the second group companies with good levels of liquidity and profitability, but under the FICs. Obviously the second option would be investing in companies of the 2nd cluster.

The last category includes companies that have real problems. Companies in the latter cluster were the lowest solvability ratios, lowest liquidity ratios and lowest profitability. In fig.1 you can see the attributes of clusters determined with the k-means algorithm.

We have repeated all these operations for the last 5 years and tried to see how companies have evolved as they migrated from one cluster to another. We could see that FICs were part of the same class every year, even if previous years were joined by other companies.

D. Discriminant analysis

After finding a proper way to separate the elements of the Ω set (those 38 companies) on classes of prediction ω_1, ω_2, ω_3 (the 3 classes found), the main task of discriminant analysis is to decide on the membership of the 3 classes of new objects from the Ω set or to make predictions concerning the affiliation of these objects. This means that the problem of classification using discriminant analysis can be made as follows:

Giving an object that is known vector x of values of its characteristics, is required to determine the object belonging to one of the 3 classes possible, ω_1, ω_2, ω_3 of the set Ω.

We are trying to verify the data obtained by the K-means algorithm. Classification Matrix (generated by statistics 7.1) checks how many cases were good predictioned and how many were wrong. Thus, for each element A_{ij}, we can interpret: Item A was calculated as belonging to the class i and actually belong to the class j. Figure exposed below (table 1) noted that besides the main diagonal, all the matrix elements are equal to 0. As seen, prediction is considered 100% correct.

	Classification Matrix (Stefan_IND_11) Rows: Observed classifications Clolumns: Predicted classifications			
Group	Percent Correct	G_1:1 p=.44737	G_2:2 p=.21053	G_3:3 p=.34211
G_1:1	100,000	17	0	0
G_2:2	100,000	0	8	0
G_3:3	100,000	0	0	13
Total	100,000	17	8	13

Table 1

E. Mahalanobis distances:

In table 2 appear the distances between centroids of the 3 classes. Obviously, the main diagonal is zero.

	Squared Mahalanobis Distances (Stefan_IND_11)		
Var_Clas	G_1:1	G_2:2	G_3:3
G_1:1	0.00000	23.46597	30.84500
G_2:2	23.46597	0.00000	27.93896
G_3:3	30.84500	27.93896	0.00000

Table 2

We could also generate a table (table 3) in which in the first two columns are showed the cases and belonging to the groups determined by the K-means algorithm. Mahalanobis distances of the following 3 columns are the distances from each firm to centroid of each class. We can notice that for all those 38 companies, the smallest distance corresponds the cluster centroid of which belongs the company.

Interesting is that for any case, a new company, can predict belonging to a class calculating the for Mahalanobis distances. Obviously, the minimum will show which class belongs to the newcomer.

F. Discriminant Functions

The main problem to be solved within discriminant analysis is the construction of criteria or rules of classification, from which it can make predictions concerning the affiliation of new forms, with initial affiliation unknown. Criteria for classification are known as the classifiers, and the deduction of these criteria is called training of the classifiers.

The classifier is actually using an algorithm which determines the most likely belong to a form to a certain class of prediction. The training of the classifier is based on the information contained in a sample form whose affiliation is known aprioric which is called training set.

Determination of discriminant function is equivalent to finding some directions, or vectors, in relation with whom the intragroup variability would be minimal, and intergroup variability to be high.

	Observed Classif.	1 Mah. Distance	2 Mah. Distance	3 Mah. Distance
OtelInox	2	67,24908	23,58100	68,67287
Azomure	1	10,08497	27,23722	47,59443
Compa	1	4,71148	20,82651	38,99490
Oil	1	6,42258	27,03342	21,66523
Turbome	3	38,10748	42,20765	5,17626
Argus	3	16,51596	18,44893	5,36003
OltChim	1	6,69162	14,92195	10,84655
Sicomed	3	57,29044	46,21570	10,32575
Armonil	3	46,48856	53,21035	17,79527
Terapia	3	37,35062	33,42294	9,72282
SIF_BT	1	12,05725	51,93501	53,54130
ZimTub	1	7,92608	36,15070	41,68277
VelPitar	1	12,42725	30,65832	24,29838
Bermas	3	30,93517	32,95997	6,50884
Aerostar	1	2,96811	27,94058	30,06770
VaeApoc	1	1,98110	24,93895	32,51925
AstraRom	1	2,56938	28,01788	30,12886
Mefin	1	4,08341	20,46803	40,39430
SN_Orsov	1	4,33937	27,20846	47,12793
Antibiotic	3	30,76574	41,27384	10,77339
Automati	3	33,69122	23,60456	4,44909
Electroput	1	32,92170	58,51367	64,50474
Sofert	3	52,21094	42,70212	8,76706
Siretul	2	28,11264	5,07846	24,03093
THR	2	21,93913	9,24703	16,28507
SeverNav	3	28,20683	27,20665	13,41986
RomCarb	1	7,29635	20,33439	34,14575
Rulmentu	1	7,22442	11,89560	28,74135
Mechel	3	34,03632	28,96974	8,92617
MjMaillis	3	67,64886	63,61442	31,05068
IATSA_P	2	27,85791	4,58925	21,14615
Flaros	2	20,72841	11,49128	45,34490
Ilefor	2	17,38662	4,05994	38,13593
Mittal	2	33,56190	4,23653	35,82074
ElPreco	1	12,91507	54,87127	47,65153
Electrom	3	37,82378	22,43910	9,46848
Bega Tehn	2	22,97248	4,61754	23,33045
ComElf	1	4,73514	25,83105	30,41801

Table 3

These directions will define the axes of discriminant space against and can be identified in the form of linear combinations of descriptions variables selected in the analysis.

In conclusion, we can say that Fisher's discriminant functions are linear functions with the following form:

$$D(X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n \quad (27)$$

Where $\beta_0 = -(\beta_1 \cdot \mu_1 + \beta_2 \cdot \mu_2 + \ldots + \beta_n \cdot \mu_n)$ is the free element, and coefficients $\beta_1, \beta_2, \ldots, \beta_n$ are components of an eigenvector of the matrix $\sum_w^{-1} \cdot \sum_b$.

Classification Function parameters, calculated using Statistica 7.1 appear in table 4:

Variable	Classification Functions; Grouping Var_Clas(Stefan_IND_11) G_1:1 p=.44737	G_2:2 p=.21053	G_3:3 p=.34211
Lich_CRT	-2.42295	0.72274	2.72371
Lich_Imd	0.68960	3.00506	-2.75105
Indator	1.79983	-2.46766	-0.83506
Acop_Dob	-1.63140	-0.84026	2.65045
Vrot_Deb	-1.83965	-1.87618	3.56026
Vrot_ActC	1.12591	-0.94167	-0.89285
Vrot_ActT	1.14229	2.86493	-3.25680
Rentab_Cap	-1.98692	1.05747	1.94753
Mrj_Br_V	-2.07831	1.25443	1.94583
Rez_Act	1.74793	-0.62868	-1.89887
Constant	-3.82054	-6.67595	-6.01773

Table 4

The discriminant functions will be:

$$D_1(X) = -3,82 - 2,42X_1 + 0,69X_2 + \ldots +1,75X_{10} \quad (28)$$
$$D_2(X) = -6,68 + 0,72X_1 + 3,01X_2 + \ldots - 0,63X_{10} \quad (29)$$
$$D_3(X) = -6,02 + 2,72X_1 - 2,75X_2 + \ldots - 1,90X_{10} \quad (30)$$

The coefficients for the first discriminant function are derived so as to maximize the differences between the group means. The coefficients for the second function are also derived to maximize the difference between the group means, but the values of the functions are not correlated. The second function is orthogonal to the first and the third is orthogonal to the second. Variables will be even values of the 10 indicators normalized.

What is most interesting about the functions of classification is that it may set belonging to a set of classes for any new company whose indicators are known but unknown membership.

G. The a prioric and posterior probabilities:

To estimate the aprioric probabilities P_1, P_2, P_3, is calculated the number of cases or the number of firms in each class using the information on the K-means algorithm. Then we determine T_i which represents the number of firms in class i and then calculate the relative frequencies $P_i = T_i / T$.

	Prior Probabilities of Classifications (Stefan_IND_11)		
Class	1 N=17,000	2 N=8,000	3 N=13,000
Probability	0,447368	0,210526	0,342105

Table 5

We will assume that the probability density of the classes are normal, in other words are like the following:

$$f_{\omega_i}(x) = \frac{1}{(\sqrt{2\cdot\pi})^n \cdot \left|\Sigma_i\right|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}\cdot(x-\mu^{(i)})^t \cdot \Sigma_i^{-1} \cdot (x-\mu^{(i)})}$$

$i = 1,2$ (31)

a posterior probability will be:

$$P(\omega_i / x) \approx \frac{P(\omega_i) * f_{\omega_i}}{f_\Omega} \qquad (32)$$

where i = 1, 2, 3.

In table 6 we have in the first two columns firms and classes predicted using the K-means algorithm. Thus, the next 3 columns are presented as a posterior probability object to belong to the class 1, 2 or 3 and in the last three columns can be found standings probability, for example classes which belongs to the highest probability.

Note that the algorithm K has been very precisely, in most cases with a probability of over 99% respectively as subject to belong to the predicted class of K-means algorithm.

V. Conclusion

With discriminant analysis model assumptions we have checked K-means algorithm and we have succeeded in calculating the classification of features that may help in future predictions.

The most useful application of discriminant analysis are seen in the banking area in which techniques are called credit-scoring techniques wich are the most important tools for the decision on the granting of loans. Such firms can be divided into classes of trust and credit decision to make depending on membership class.

Another area would be marketing, clients can be divided into different classes of interest for those who do studies. Last but not least the establishment of development areas can be based on algorithms presented in this paper.

This is an important step towards the following researches and also represents an efficient instrument in the context of the global financial crisis. given that the last period were recorded depreciation on all markets, BSE has not been an exception, remains to be seen if the clusters found will remain approximately similar to or will change radically.

Remains a bad opinion of the authors for the Romanian economy is not fully reflected in BSE, many companies are not listed.

VI. What to do

As a continuation of this work, with application on the capital market, we can see the following points:

- broadening the base of indicators, selecting a larger and more representative of financial-economic indicators;
- determining the cluster for each year in part;
- an analysis of evolution in the dynamic; observation of clusters changing from one year to another;
- explanations of migration cases from one cluster to another;
- combination of companies with indicators on the evolution of capital market. to view the indicators that most influence the decision to invest;
- identification of the type characteristic performance of Romanian companies;
- determining the structural movement of the Romanian economy depending on size and performance of firms;
- identification of significant trends in terms of economic and financial level of Romanian companies;
- deduction of certain classes of risk and that the risk classes of companies listed on BSE;
- construction of models for character developments phenomenon Romanian scholar; building portfolios based on clusters, instead of the classical.
- making predictions on the evolution of the Romanian stock ticker and financial assets of the courses listed on the BSE.
- making predictions on the evolution of the economic performance of financial firms in the Romanian economy

References

[1] Ruxanda Gh., *Analiza datelor*, Bucuresti, Editura ASE, 2001

[2] Chris Ding and Xiaofeng He. "K-means Clustering via Principal Component Analysis". *Proc. of Int'l Conf. Machine Learning* (ICML 2004), pp 225-232.

[3] Fisher R.A . The use of multiple measurements in taxonomic problems, *Ann. of Eugenics*,7, 1936, p179-188.

[4] Ionescu, S.A.; Gheorghe, C.M. & Zaheu, I. Using Pattern Recognition Techniques To Determine Profitability on Financial Markets. 0641-0642, *Annals of DAAAM for 2008 & Proceedings of the 19th International DAAAM Symposium*, pp 321, 2008

[5] H. Steinhaus. *Sur la division des corp materiels en parties*. Bull. Acad. Polon. Sci., C1. III vol IV, 1956. :801– 804,

[6] S. Lloyd, Last square quantization in PCM's. Bell Telephone Laboratories Paper (1957). *Published in journal much later: S. P. Lloyd. Least squares quantization in PCM. Special issue on quantization, IEEE Trans. Inform*. Theory, 28, 1982.:129–137.

[7] Arnott, Robert D. Cluster Analysis and Stock Price Comovement," *Financial Analysts Journal*, vol. 37, no. 6 (November/December):56-62, 1980, ISSN : 0015-198X

[8] Farrell, James L. Analyzing Covariation of Returns to determine Homogeneous Stock Groupings, *Journal of Business,* vol. 48, no. 2 (April): 186-207, 1974, ISSN:0021-9398

[9] Farrell, James L. Homogeneous Stock Groupings: Implications for Portfolio Management, *Financial Analysts Journal*, vol. 31, no. 3 May/June:50-62, 1986, ISSN : 0015-198X

	Observed Classif.	1 prob.	2 prob.	3 prob.	Highest Prob.	Second Highest	Third Highest
OtelInox	2	0,000000	1,000000	0,000000	2	1	3
Azomures	1	0,999911	0,000089	0,000000	1	2	3
Compa	1	0,999851	0,000149	0,000000	1	2	3
Oil	1	0,999610	0,000016	0,000374	1	3	2
Turbomecanica	3	0,000000	0,000000	1,000000	3	1	2
Argus	3	0,004915	0,000880	0,994205	3	1	2
OltChim	1	0,906241	0,006961	0,086798	1	3	2
Sicomed	3	0,000000	0,000000	1,000000	3	2	1
Armonil	3	0,000001	0,000000	0,999999	3	1	2
Terapia	3	0,000001	0,000004	0,999994	3	2	1
SIF_BT	1	1,000000	0,000000	0,000000	1	2	3
ZimTub	1	1,000000	0,000000	0,000000	1	2	3
VelPitar	1	0,997931	0,000052	0,002017	1	3	2
Bermas	3	0,000006	0,000001	0,999992	3	1	2
Aerostar	1	0,999997	0,000002	0,000001	1	2	3
VaeApocarom	1	0,999995	0,000005	0,000000	1	2	3
AstraRom	1	0,999998	0,000001	0,000001	1	2	3
Mefin	1	0,999870	0,000130	0,000000	1	2	3
SN_Orsova	1	0,999995	0,000005	0,000000	1	2	3
Antibiotice	3	0,000060	0,000000	0,999940	3	1	2
Automatica	3	0,000001	0,000043	0,999957	3	2	1
Electroputere	1	0,999999	0,000001	0,000000	1	2	3
Sofert	3	0,000000	0,000000	1,000000	3	2	1
Siretul	2	0,000021	0,999854	0,000125	2	3	1
THR	2	0,003543	0,950685	0,045772	2	3	1
SeverNav	3	0,000803	0,000623	0,998573	3	1	2
RomCarbon	1	0,999305	0,000694	0,000001	1	2	3
Rulmentul	1	0,956437	0,043548	0,000016	1	2	3
Mechel	3	0,000005	0,000027	0,999968	3	2	1
MjMaillis	3	0,000000	0,000000	1,000000	3	2	1
IATSA_Pit	2	0,000019	0,999569	0,000412	2	3	1
Flaros	2	0,020537	0,979463	0,000000	2	1	3
Ilefor	2	0,002706	0,997294	0,000000	2	1	3
Mittal	2	0,000001	0,999999	0,000000	2	1	3
ElPreco	1	1,000000	0,000000	0,000000	1	3	2
Electromagnetica	3	0,000001	0,000938	0,999061	3	2	1
Bega_Tehnomet	2	0,000220	0,999640	0,000140	2	1	3
ComElf	1	0,999986	0,000012	0,000002	1	2	3

Table 6

[10] Jensen, Robert E. A Cluster Analysis Study of Financial Performance of Selected Business Firms, *The Accounting Review*, vol.46, no.1 (January): 36-56, 1971, ISSN 0001-4826

[11] King, Benjamin F. Market and Industry Factors in Stock Price Behavior, *Journal of Business*, vol. 39, no. 1, 1966, ISSN:0021-9398

[12] Livingston, Miles Industry Movements of Common Stocks, *Journal of Finance*, vol.32, no. 3 (June):861-874, 1977, ISSN: 0022-1082

[13] D.Richard, Can We Afford to Grow Older?: A perspective on the *Economics of Aging*. 1998-Cambridge, London: The MIT Press;

[14] Marius Radulescu, Sorin Radulescu, Constanta Zoie Radulescu, Gheorghita Zbaganu, The efficient frontier of the minimum-variance portfolio selection problems, *Proceedings of the 9th WSEAS Int. Conf. on MATHEMATICS AND COMPUTERS IN BUSINESS AND ECONOMICS (MCBE'08),* Bucharest, Romania, June 24-26, 2008, p39-43

[15] Mihaela Cristina Dragoi, Ecaterina Ionescu, Irina Eugenia Iamandi, Alina Chiciudean, Laura Gabriela Constantin, Economic and Social Determinants of Health. Modelling the Case of Romania, *Proceedings of the 9th WSEAS Int. Conf. on MATHEMATICS AND COMPUTERS IN*

BUSINESS AND ECONOMICS (MCBE'08), Bucharest, Romania, June 24-26, 2008, p91-96

[16] Gabriela Prelipcean,. Mircea Boscoianu, Some Aspects regarding the Dynamic Correlation between Different Types of Strategic Investments in Romania after Integration, *Proceedings of the 9th WSEAS Int. Conf. on MATHEMATICS AND COMPUTERS IN BUSINESS AND ECONOMICS (MCBE'08),* Bucharest, Romania, June 24-26, 2008, p162-166

[17] Zeljko Panian, A Break-Through Approach to Real-Time Decisioning in Business Management, *Proceedings of the 2nd WSEAS International Conference on MANAGEMENT, MARKETING and FINANCES (MMF'08),* Cambridge, Massachusetts, USA, March 24-26, 2008, p25-30

[18] Begumhan Ozdincer, Cenktan Ozyildirim, Determining The Factors of Bank Performance with a Focus on Risk and Technical Efficiency, *Proceedings of the 2nd WSEAS International Conference on MANAGEMENT, MARKETING and FINANCES (MMF'08),* Cambridge, Massachusetts, USA, March 24-26, 2008, p31-39

[19] Gh.Ruxanda, A. Stoian, Modelling of the Relationship between Foreign Direct Investmentand Economic Growth, *ECONOMIC COMPUTATION AND ECONOMIC CYBERNETICS STUDIES AND RESEARCH*, vol3-4, 2008, ISSN: 0424-267X, p49-62

[20] Fengming M. Chang, Characteristics Analysis for Small Data Set Learning and the Comparison of Classification Methods, *Proceedings of the 7th WSEAS International Conference on ARTIFICIAL INTELLIGENCE, KNOWLEDGE ENGINEERING and DATA BASES (AIKED '08)* Cambridge, UK, February 23-25, 2008, p122-127

[21] Chang-Rui Yu, Yan Luo, Research on Models of Decision-Making Problem Recognition, *WSEAS TRANSACTIONS on BUSINESS and ECONOMICS*, Issue 3, Volume 5, March 2008, p37-43, ISSN: 1109-9526

[22] Mihaela Cristina Dragoi, Ecaterina Ionescu, Irina Eugenia Iamandi, Alina Chiciudean, Laura Gabriela Constantin, An Economic Analysis of the Romanian Healthcare System based on an European Comparative Approach, *WSEAS TRANSACTIONS on BUSINESS and ECONOMICS,* Issue 6, Volume 5, June 2008, p330-340, ISSN: 1109-9526

[23] Bi-Huei Tsai, Clustering Evolution of Taiwan IC Industry in China, *WSEAS TRANSACTIONS on BUSINESS and ECONOMICS,* Issue 10, Volume 5, October 2008, p471-480, ISSN: 1109-9526

[24] Loukeris N. and Matsatsinis N. Corporate financial evaluation and bankruptcy prediction implementing artificial int elligence methods, *WSEAS Transactions on Business and Economic* 4 (3), April, 2006, p. 338-343. ISSN: 1109-9526

[25] Jabbour, G. and Maldonado, L. Prediction of financial time series with time-line hidden Markov experts and ANNs, *WSEAS Transactions on Business and Economic* 9 (4), September, 2007, p. 140-145, ISSN: 1109-9526

[26] Bach, M. P. and Simicevic, V. Financial structure and the profitability of Croatian banks: Modelling and Forecasting, *WSEAS Transactions on Business and Economic* 3 (3), March, 2006, p. 142-149. ISSN: 1109-9526

[27] *** *Annual Statement Studies, Financial Ratio Benchmarks*, The Risk Management Association, 2007-2008

[28] *** *Introductory Manual on Project Finance for Managers of PPP Projects*, Version 1 – National Treasury/PPP Manual January 2001