Audit financiar, XV, Nr. 3(147)/**2017**, *385-394*
ISSN: 1583-5812; ISSN on-line: 1844-8801

Using fuzzy c-means clustering algorithm in financial health scoring

Pınar OKAN GOKTEN,
Gazi University, Turcia,
E-mail: pinar.okan@gazi.edu.tr

Furkan BASER,
Ankara University, Turcia,
E-mail: furkan.baser@ankara.edu.tr

Soner GOKTEN,
Baskent University, Turcia,
E-mail: sgokten@baskent.edu.tr

Abstract

Classification of firms according to their financial health is currently one of the major problems in the literature. To our knowledge, as a first attempt, we suggest using fuzzy c-means clustering algorithm to produce single and sensitive financial health scores especially for short-term investment decisions by using recently announced accounting numbers. Accordingly, we show the calculation of fuzzy financial health scores step by step by benefit from Piotroski's criteria of liquidity/solvency, operating efficiency and profitability for the firms taken as a sample. The results of correlation analysis indicate that calculated scores are coherent with short-term price formations in terms of investors' behavior and so fuzzy c-means clustering algorithm could be used to sort firm in a more sensitive perspective.

Keywords: *Accounting numbers, financial analysis, financial classification, Fuzzy c-means (FCM) clustering algorithm*

JEL Classification: *G30, M49*

To cite this article:

Gokten, P.O., Baser, F. and Gokten, S. (2017), Using fuzzy c-means clustering algorithm in financial health scoring, *Audit Financiar*, vol. XV, nr. 3(147)/2017, pp. 385-394, DOI: 10.20869/AUDITF/2017/147/385

To link to this article:

http://dx.doi.org/10.20869/AUDITF/2017/147/385

Received: 28.02.2017
Revised: 16.03.2017
Accepted: 20.03.2017

Introduction

Investors' decisions represent the expectations derived from cumulative beliefs that include past experiences and formation of recent reasonable differences in prior beliefs (Ball and Brown, 1968; Morris 1996; Fama, 1998; Core et al., 2003; Cajueiro and Tabak, 2004; Brimble and Hodgson, 2007). In this sense, announced accounting numbers as quantized signals play an important role on changes in beliefs and can cause rapid stock price fluctuations mostly in weak efficient or inefficient markets where investors must actively manage their portfolios in order to expect a proper return in the frame of speculative investment behavior (Fama et al., 1969; Malkiel and Fama, 1970; Harrison and Kreps, 1978; D'Ambrosio, 1980; Harvey, 1993; Urrutia, 1995; Aitken, 1998; Grieb and Reyes, 1999). In other words, especially in short-term, investors buy or sell stocks based on changes in financial health of firms which become clear by recently announced accounting numbers (Core et al., 2003). Therefore investors need summarized indicators to make investment decisions in post-announcement short-term.

Financial classification is useful tool for market participants to compare differentiation in financial situations. Although there are lots of general accepted scores, summarizing the large amount of valuable data is currently one of the major problems in the literature. For instance, F-Score is widely accepted benchmark developed by Piotroski (2000) to show financial performance of the firms as single summarized indicator and provides many useful insights to identify financially healthier firms. However, the numerical characteristic makes F-Scoring (between '0': the lowest and '9': the highest qualification) insensitive to sort and classify firms especially to explain the price formations ans short-term investors' decisions.

In the literature, there are various studies which utilize a clustering algorithm for classification problem. On the other hand, while most part of these studies have tried to integrate clustering techniques into portfolio management (Pattarin et al., 2004; Tola et al., 2008; Nanda et al., 2010 etc.), there are limited number of studies that concentrate on classifying the firms based on their announced accounting numbers.

Wang and Lee (2008) suggest a clustering method based on a fuzzy relation to classify the financial ratios of different companies and they stated that the clustering method can be applied in conditions where the cluster number is not determined. On the other side, their study does not mention the benefit of using this kind of clustering.

The main contribution of this study is to suggest a systematic alternative to sort firms sensitively according to changes in their financial health based on recently announced accounting numbers. In other words, we show how Fuzzy c-means (FCM) clustering algorithm could be used in order to produce single and more sensitive numerical indicator, hereupon called 'fuzzy financial health score (F-FHS: between '0' and '1')', that show the changes in financial health compared to previous year. To our knowledge, our study is first to provide a methodological perspective under this point of view.

We present this methodological perspective through an implementation on selected sample. Since the reaction level of markets with low efficiency on recently announced accounting numbers is high, we select the data of 166 active firms listed and traded on National Market of Istanbul Stock Exchange[1] as a sample in model implementation. We use delta determinants of F-Score to calculate F-FHSs of selected firms: ∆ROA (change in return on assets), ∆CFO (change in cash flow from operations), ∆LEV (change in leverage), ∆CR (change in current ratio), ∆MARGIN (change in gross margin) and ∆TURN (change in asset turnover).

2013 and 2014 annual announced accounting numbers were used because in that period Turkey initialized its position against IMF and ranked as the sixth biggest economy in Europe and the sixteenth in the world. Therefore, these years can more clearly reflect firm specific performance under smooth economic conditions.

In order to see if F-FHSs are meaningful summarized single indicators or not, correlation analysis is executed between calculated scores and realized returns of firms for given short period. Ten trading days (n) are used as pre and post terms of announcement time of financial statements and three different indicators are used as return inputs, 'rA', 'rB' and 'rC'.

'r_A' denotes the price changes of stock in percentages by using 'P_{t+n}' and 'P_{t-1}', indicate the stock prices at

[1] Studies in the literature such as Balaban (1995), Kawakatsu and Morey (1999), Buguk and Brorsen (2003) etc. found that Turkish market is weak form efficient.

the end of post announcement term and the one trading day before from 't' respectively, while 't' indicates the announcement date of financial statements.

$$\mathrm{r_A} = \frac{P_{t+n} - P_{t-1}}{P_{t+n}} \qquad (1)$$

In order to make return input more explanatory from the view of investors' active behavior, trading volumes are taking into account for pre and post terms via calculating their weights '$w_{t\mp i}$' and '$\mathrm{r_B}$' that denotes the weighted average price changes of stock in percentages is added into analysis as second return input.

$$r_B = \frac{\left[\frac{\sum_{i=1}^{n} w_{t+i}P_{t+i}}{\sum_{i=1}^{n} w_{t+i}}\right] - \left[\frac{\sum_{i=1}^{n} w_{t-i}P_{t-i}}{\sum_{i=1}^{n} w_{t+i}}\right]}{\frac{\sum_{i=1}^{n} w_{t+i}P_{t+i}}{\sum_{i=1}^{n} w_{t+i}}} \qquad (2)$$

More return or less loses results compared to market's return also perceived as win situation by investors. In this sense, '$\mathrm{r_C}$' is added as another return input via calculating the spread between '$\mathrm{r_B}$' and market return ($\mathrm{r_M}$) which indicates percentage change in market index value (IV) between the post n[th] trading day and announcement date.

$$\mathrm{r_C} = r_B - \frac{IV_{t+n} - IV_t}{IV_{t+n}} \qquad (3)$$

The structure of the paper is as follows. In the next section, a brief overview of the FCM algorithm is provided. In section 2, data sources are mentioned, calculation of F-FHSs is shown step by step and the results of correlation analysis are given. In Section 3, conclusions are mentioned.

1. FCM Clustering Algorithm

Clustering algorithms based on its structure are generally divided into two types: fuzzy and nonfuzzy (crisp) clustering. Crisp clustering algorithms give better results if the structure of the data set is well distributed. However, when the boundaries between clusters in data set are ill defined, the concept of fuzzy clustering becomes meaningful (Nefti and Oussalah, 2004). Fuzzy methods allow partial belongings (membership) of each observation to the clusters, so they are effective and useful tool to reveal the overlapping structure of clusters (Zhang, 1996). Fuzzy c-means (FCM) clustering algorithm is one of the most widely used method among fuzzy associated models (Bezdek and Pal, 1992).

Fuzzy clustering methods are used for calculating the membership function that determines to which degree the objects belong to clusters and used for detecting overlapping clusters in the data set. FCM clustering algorithm, one of the commonly used clustering method, is initially proposed by Dunn (1973) and developed by Bezdek (1981).

Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ denote a set of n objects and each i object ($i = 1, 2, \dots, n$) is represented with d dimensional vector $\mathbf{x}_i = [x_{1,i}\ x_{2,i}\ \dots\ x_{d,i}]^T \in \Re^d$. So, $n \times d$ dimensional data matrix, composed of a set of n vectors is

$$\mathrm{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix}. \qquad (4)$$

A fuzzy clustering algorithm separates data matrix, X into c overlapping clusters in accordance with the design of a fuzzy partition matrix, U. Fuzzy partition matrix, U is composed of the degrees of memberships of objects, $\mathbf{x}_i$ ($i = 1, 2, \dots, n$) in every cluster k ($k = 1, 2, \dots, c$). The degree of membership of i. vector in cluster k is represented by $\mu_{k,i} \in$ U. Accordingly, the partition matrix is given by

$$\mathrm{U} = \begin{bmatrix} \mu_{1,1} & \mu_{2,1} & \cdots & \mu_{c,1} \\ \mu_{1,2} & \mu_{2,2} & \cdots & \mu_{c,2} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{1,n} & \mu_{2,n} & \cdots & \mu_{c,n} \end{bmatrix}. \qquad (5)$$

In fuzzy clustering method, each cluster is represented with a vector of cluster centers which is usually identified as the centroids of d objects, e.g., average of all the datum of the corresponding cluster (Celikyilmaz and Turksen, 2009). The algorithm calculates c number of cluster center vectors $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c\} \in \mathcal{R}^{c \times d}$ where each cluster center is denoted as $\mathbf{v}_k \in \Re^d$, $k = 1, 2, \dots, c$.

FCM clustering algorithm is a simple and convenient method. In this method, the number of clusters, c is assumed to be known or at least fixed. Because this assumption is considered to be unrealistic in many data analysis problems, the method for determining the number of clusters such as Cluster Validity Index (CVI) analysis has been developed in FCM clustering algorithm (Pal and Bezdek, 1995; Kim and Ramakrishna, 2005; Celikyilmaz and Turksen, 2008).

FCM clustering method is based on a constrained optimization problem reaching the optimum solution with the minimum of the objective function. The mathematical model of this optimization problem with two prior information such as number of cluster, c and fuzziness parameter, m is identified as:

$$\min J(\mathrm{X};\mathrm{U},\mathrm{V}) = \sum_{k=1}^{c}\sum_{i=1}^{n}\left(\mu_{k,i}\right)^{m} d^{2}(\mathbf{x}_i,\mathbf{v}_k)$$
$$0 \le \mu_{k,i} \le 1\ ,\forall i,k \qquad (6)$$
$$\sum_{k=1}^{c}\mu_{k,i} = 1\ ,\forall i > 0$$
$$0 < \sum_{i=1}^{n}\mu_{k,i} < n\ ,\forall k > 0,$$

where each cluster is represented by a prototype, $\mathbf{v}_i$ (Bezdek, 1981). The value of $m \in (1,\infty)$ in objective function is expressed as the degree of fuzziness or fuzzifier, and it determines the degree of overlapping of clusters. The situation of "$m = 1$" which means that the clusters are not overlapping represents the crisp clustering structure (Hammah and Curran, 1998). Here, $d^{2}(\mathbf{x}_i,\mathbf{v}_k)$ is the measure of distance between i. object and k. cluster center. FCM clustering algorithm specifically uses Euclidean distance. Quadratic distance ensures that the objective function is not negative definite, > 0 .

Optimum membership values and cluster centers derived from the solution of optimization problem in (6) with the method of Lagrange multipliers are calculated as,

$$\mu_{k,i}^{(t)} = \left[\sum_{l=1}^{c}\left(\frac{d\left(\mathbf{x}_i,\mathbf{v}_k^{(t-1)}\right)}{d\left(\mathbf{x}_i,\mathbf{v}_l^{(t-1)}\right)}\right)^{\frac{2}{m-1}}\right]^{-1} \qquad (7)$$

$$v_k^{(t)} = \frac{\sum_{i=1}^{n}\left(\mu_{k,i}^{(t)}\right)^{m}\mathbf{x}_i}{\sum_{i=1}^{n}\left(\mu_{k,i}^{(t)}\right)^{m}}\ ,\forall k = 1,2,\ldots,c \qquad (8)$$

In eq. (7), $\mathbf{v}_k^{(t-1)}$ denotes cluster center vector for cluster i obtained in $(t-1)$th iteration. $\mu_{k,i}^{(t)}$ in eqs. (7) and (8) denotes optimum membership values obtained at t. iteration. According to this operation, the membership values and cluster centers seem to be dependent on each other. Therefore, Bezdek (1981) proposed an iterative formula for determining membership values and cluster centers. Accordingly, at each iteration t, objective function $J^{(t)}$ is determined by

$$J^{(t)} = \sum_{k=1}^{c}\sum_{i=1}^{n}\left(\mu_{k,i}^{(t)}\right)^{m} d^{2}\left(\mathbf{x}_i,\mathbf{v}_k^{(t)}\right) > 0 \qquad (9)$$

FCM algorithm is ended at the end of a particular iteration or according to a termination rule defined as $\left|v_k^{(t)} - v_k^{(t-1)}\right| \le \varepsilon$ (Celikyilmaz and Turksen, 2009).

2. The data and empirical implementation

2.1. Data

166 active firms listed and traded on National Market of Istanbul Stock Exchange (BIST-Borsa Istanbul) were selected as a sample for empirical implementation. Selected sample does not include financial service firms and the companies with lack of data. Annual accounting numbers of 2013 and 2014 are obtained from Public Disclosure Platform of BIST (KAP). BIST100 index is used for '$\mathrm{r_M}$' calculation. Historical stock prices, trading volumes and index values are obtained from 'Matriks' which is the one of formal data distributor of BIST. The dates for post and pre n^{th} trading days are shown in Table 1 according to announcement dates of financial statements for each firm.

Table 1. 't', pre nth and post nth days of each selected firms

Firm Codes *(Date notation: day/month/year)*	Post 10^{th} Trading Day	Announcement ← Date (t) →	Pre 10^{th} Trading Day
ARCLK	13.02.2015	30.01.2015	16.01.2015
TOASO	16.02.2015	02.02.2015	19.01.2015
ARENA	17.02.2015	03.02.2015	20.01.2015
AFYON, TTKOM	19.02.2015	05.02.2015	22.01.2015
HEKTS	20.02.2015	06.02.2015	23.01.2015
MAALT	23.02.2015	09.02.2015	26.01.2015
IZOCM	24.02.2015	10.02.2015	27.01.2015
CEMTS, EREGL	25.02.2015	11.02.2015	28.01.2015

Firm Codes (Date notation: day/month/year)	Post 10th Trading Day	Announcement ← Date (t) →	Pre 10th Trading Day
EGGUB, ERBOS, TCELL, VESBE	26.02.2015	12.02.2015	29.01.2015
OTKAR, PKART	27.02.2015	13.02.2015	30.01.2015
USAK	27.02.2015	14.02.2015	02.02.2015
LOGO, TKNSA, VERUS	02.03.2015	16.02.2015	02.02.2015
BOLUC, MRDIN	03.03.2015	17.02.2015	03.02.2015
FROTO	04.03.2015	18.02.2015	04.02.2015
TAVHL	05.03.2015	19.02.2015	05.02.2015
AKENR, AKSA, TATGD	06.03.2015	20.02.2015	06.02.2015
CRFSA, KARTN, KONYA	09.03.2015	23.02.2015	09.02.2015
AKCNS, COMDO, CIMSA, NETAS, PIMAS, VESTL	10.03.2015	24.02.2015	10.02.2015
ASUZU, BAGFS, BOYNR, THYAO, TTRAK	11.03.2015	25.02.2015	11.02.2015
BIZIM, BRISA, BUCIM, KOZAL	12.03.2015	26.02.2015	12.02.2015
ALCAR, ALKA, DITAS, DOAS, EGSER, GOODY, INTEM, KORDS, KRSTL, OLMIP, SANKO, SASA, SODA, TRKCM, TUKAS, YUNSA	13.03.2015	27.02.2015	13.02.2015
ADEL, AKSUE, AKPAZ, ALKIM, ANACM, AYGAZ, BAKAB, BSOKE, BOSSA, BURVA, DMSAS, DENCM, DERIM, DYOBY, ENKAI, IHEVA, IHGZT, KAREL, KENT, KLMSN, KNFRT, KUTPO, LINK, OZBAL, PRKME, PETUN, PINSU, PNSUT, TEKTU, TUPRS, UYUM, VKING, YATAS	16.03.2015	02.03.2015	16.02.2015
BIMAS, KCHOL, PARSN, SELEC	17.03.2015	03.03.2015	17.02.2015
AKFEN, BRSAN, IHYAY, IZMDC, MNDRS, PGSUS, ULKER	18.03.2015	04.03.2015	18.02.2015
AEFES, CCOLA, KERVT, NUHCM, TKFEN	19.03.2015	05.03.2015	19.02.2015
AKSEN, ASELS, BMEKS, CMBTN, CMENT, ZOREN	20.03.2015	06.03.2015	20.02.2015
HURGZ, PETKM	20.03.2015	07.03.2015	23.02.2015
ALARK, ALCTL, ALYAG, AYEN, BANVT, EDIP, IHLAS, INDES, ULUSE	23.03.2015	09.03.2015	23.02.2015
CLEBI, DEVA, DOHOL, GENTS, POLHO, ROYAL, SARKY, TRCAS, VAKKO	24.03.2015	10.03.2015	24.02.2015
ANELE, ARSAN, BTCIM, BURCE, CEMAS, DGKLB, ECILC, EMKEL, ESCOM, GEREL, GLYHO, GOLTS, GUBRF, IPEKE, ITTFH, KARSN, KILER, KOZAA, MRSHL, MGROS, TIRE, NTHOL, PENGD, SKTAS, TMSN, TBORG, YAZIC	25.03.2015	11.03.2015	25.02.2015
NTTUR	26.03.2015	12.03.2015	26.02.2015
ODAS	08.06.2015	25.05.2015	11.05.2015
MARTI	23.06.2015	09.06.2015	26.05.2015

Source: Developed by authors.

2.2. Empirical Implementation: Producing Fuzzy Financial Health Scores

In order to calculate F-FHSs for selected firms, the steps outlined below are performed.

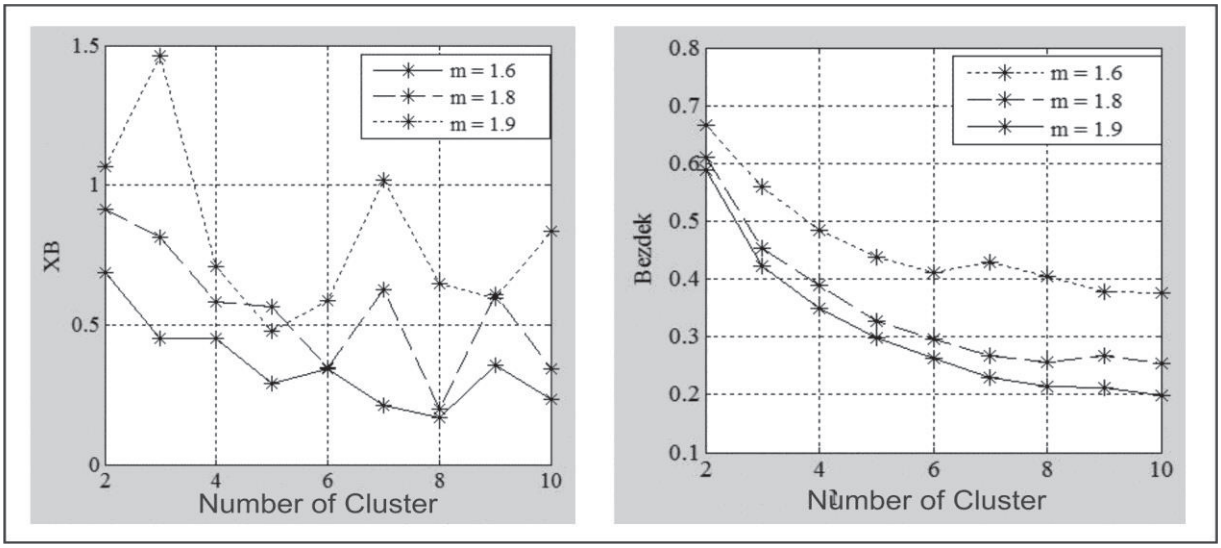
Step 1. Due to the existing heterogeneity in measurement units of variables, it is necessary to perform a homogenization process. By utilizing the normalization of the variables, weighting variables more or less is prevented. The normalization process is performed with the following relation:

$$X_{new} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (10)$$

where X_{min} is the minimum value and X_{max} is the maximum value of corresponding variable. All variables with normalization is scaled to the range $[0, 1]$.

Step 2, Optimum value of the number of cluster (c) and degree of fuzziness (m) are determined by utilizing CVI analysis.

Figure 1. The change in cluster validity indices according to the number of cluster, (left) XB index, (right) Bezdek's partition coefficient



Source: Developed by authors

In **Figure 1**, the results of two validity indexes, Xie – Beni (XB) Index (Xie and Beni, 1991) and Bezdek's partition coefficient (Pal and Bezdek, 1995), are displayed. The proper value of the number of cluster and degree of fuzziness that satisfies the minimization of XB index and maximization of Bezdek's partition coefficient are determined as $c = 5$ and $m = 1.6$, respectively.

Step 3. Cluster center vectors and partition matrix are determined by applying FCM clustering algorithm with the prior information, c and m, obtained at previous step.

For $c = 5$ and $m = 1.6$ by applying FCM clustering method, cluster center vectors, $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c\} \in \mathcal{R}^{c \times d}$ are determined as,

$$V = \begin{bmatrix} 0.348 & 0.271 & 0.212 & 0.409 & 0.508 & 0.539 \\ 0.573 & 0.285 & 0.219 & 0.493 & 0.716 & 0.593 \\ 0.508 & 0.276 & 0.217 & 0.470 & 0.621 & 0.637 \\ 0.684 & 0.279 & 0.223 & 0.493 & 0.720 & 0.531 \\ 0.499 & 0.281 & 0.213 & 0.478 & 0.626 & 0.526 \end{bmatrix} \quad (11)$$

Step 4. Euclidean norm is calculated for each cluster center vector.

In this implementation, it is claimed that the norm values allow an assessment of the general level of financial health for each cluster. Thus, while the value of the calculated norm for each cluster increases, the level of financial health rises in accordance with defined determinants, and while the norm value becomes smaller, the level of financial health of cluster will be reduced similarly. As a result, calculated Euclidean norms for center vectors of five clusters are given in **Table 2**.

Table 2. Euclidean norms calculated for the cluster center vectors

Cluster Number	Norm (h_i)
1	0.978
2	1.251
3	1.181
4	1.281
5	1.127

Source: Developed by authors.

Step 5. The advantage of FCM clustering algorithm is to produce the degree of membership of each country to c cluster. Let the degree of memberships of i. firm to c number of cluster is denoted as $\boldsymbol{\mu}_i = [\mu_{1,1}, \mu_{2,1}, \dots, \mu_{c,1}]$ and the vector consisting of the norms of cluster center vectors is represented by $\boldsymbol{h}$. Accordingly, the F-FHS for each firm is determined with the following formula,

$$\lambda_i = \boldsymbol{\mu}_i \boldsymbol{h} \quad (12)$$

Step 6. F-FHSs of each firm are presented in Table 3 and F-FHS, $i = 1, 2, \dots, n$ is calculated with the following relation,

$$\mathrm{F-FHS}_i = \frac{\lambda_i - \lambda_{min}}{\lambda_{max} - \lambda_{min}}, i = 1, 2, \dots, n \quad (13)$$

Table 3. F-FHS of each firm

Firm Codes	F-FHS	Firm Codes	F-FHS	Firm Codes	F-FHS	Firm Codes	F-FHS	Firm Codes	F-FHS
ADEL	0.34	BRSAN	0.61	FROTO	0.57	LINK	0.30	SKTAS	0.73
AFYON	0.88	BOSSA	0.68	GENTS	0.49	LOGO	0.87	TATGD	0.84
AKCNS	0.92	BOYNR	0.66	GEREL	1.00	MRDIN	0.91	TAVHL	0.89
AKENR	0.45	BRISA	0.80	GLYHO	0.29	MAALT	0.74	TEKTU	0.78
AKFEN	0.42	BURCE	0.83	GOODY	0.67	MRSHL	0.59	TKFEN	0.86
AKSA	0.67	BURVA	0.83	GOLTS	0.63	MARTI	0.57	TKNSA	0.59
AKSEN	0.77	BUCIM	0.92	GUBRF	0.77	MNDRS	0.60	TOASO	0.63
AKSUE	0.49	CRFSA	0.98	HEKTS	0.59	MGROS	0.81	TRKCM	0.74
AKPAZ	0.61	CCOLA	0.66	HURGZ	0.33	TIRE	0.90	TUKAS	0.40
ALCAR	0.65	COMDO	0.68	IHEVA	0.77	NTHOL	0.52	TRCAS	0.60
ALARK	0.39	CLEBI	0.72	IHGZT	0.94	NTTUR	0.64	TCELL	0.53
ALCTL	0.87	CEMAS	0.25	IHLAS	0.78	NETAS	0.70	TMSN	0.05
ALKIM	0.83	CEMTS	0.62	IHYAY	0.92	NUHCM	0.92	TUPRS	0.57
ALKA	0.64	CMBTN	0.54	INDES	0.58	ODAS	0.81	THYAO	0.70
ALYAG	0.00	CMENT	0.93	INTEM	0.71	OLMIP	0.68	TTKOM	0.72
ANACM	0.68	CIMSA	0.63	IPEKE	0.50	OTKAR	0.69	TTRAK	0.12
AEFES	0.41	DMSAS	0.89	ITTFH	0.52	OZBAL	0.66	TBORG	0.54
ASUZU	0.37	DENCM	0.90	IZMDC	0.77	PRKME	0.31	ULUSE	0.50
ANELE	0.68	DERIM	0.43	IZOCM	0.92	PARSN	0.91	USAK	0.70
ARCLK	0.68	DEVA	0.56	KAREL	0.60	PGSUS	0.68	UYUM	0.70
ARENA	0.53	DITAS	0.86	KARSN	0.44	PENGD	0.90	ULKER	0.67
ARSAN	0.93	DOHOL	0.56	KARTN	0.39	PETKM	0.51	VAKKO	0.68
ASELS	0.57	DGKLB	0.61	KENT	0.87	PETUN	0.68	VERUS	0.69
AYEN	0.91	DOAS	0.65	KERVT	0.67	PINSU	0.85	VESBE	0.93
AYGAZ	0.68	DYOBY	0.66	KILER	0.65	PNSUT	0.62	VESTL	0.82
BAGFS	0.80	EDIP	0.93	KLMSN	0.58	PIMAS	0.39	VKING	0.94
BAKAB	0.69	EGGUB	0.95	KCHOL	0.65	PKART	0.48	YATAS	0.81
BANVT	0.76	EGSER	0.92	KNFRT	0.75	POLHO	0.52	YAZIC	0.35
BTCIM	0.92	ECILC	0.53	KONYA	0.91	ROYAL	0.57	YUNSA	0.82
BSOKE	0.87	EMKEL	0.82	KORDS	0.92	SANKO	0.69	ZOREN	0.85
BIMAS	0.65	ENKAI	0.65	KOZAL	0.68	SARKY	0.69		
BMEKS	0.68	ERBOS	0.93	KOZAA	0.51	SASA	0.91		
BIZIM	0.54	EREGL	0.66	KRSTL	0.36	SELEC	0.69		
BOLUC	0.82	ESCOM	0.15	KUTPO	0.99	SODA	0.85		

Source: Developed by authors.

Step 7. Correlation analysis is executed in order to see if F-FHSs work or not. The results of significance test are given in Table 4 and there is a statistically significant relationship between F-FHSs and r_B, r_C respectively. That means F-FHSs are coherent with short-term price formations and so, the scores could be used as summarized single indicators to sort firms according to changes in their financial health in a more sensitive way.

Table 4. The relation between F-FHS and r_A, r_B, r_C

Return Inputs	F-FHS	
	Corr. Coeff.	Sig. Level (p)
r_A	0,138	0,794
r_B	0.652	0,000*
r_C	0,529	0,000*

$^* p < 0{,}05$,

Source: Developed by authors.

Conclusion

The paper suggests a methodological perspective for the first time on how Fuzzy c-means (FCM) clustering algorithm could be used in order to sort firms according to changes in their financial health compared to previous year. Accordingly, to show this methodology, we applied FCM clustering algorithm and produced fuzzy financial health scores (F-FHSs) of selected 166 active firms listed and traded on National Market of Istanbul Stock Exchange by benefit from F-Score's delta determinants calculated via using the accounting numbers of 2013 and 2014. This implementation enables us to classify firms in a more sensitive way based on single numerical indicator.

A correlation analysis was executed between calculated scores and realized returns for a given short term in order to investigate the employability of F-FHSs. The results indicate that FCM clustering algorithm is beneficial tool to sort firms according to their financial health level and can provide a sensitive and single summarized indicator for investment decisions based on recently announced accounting numbers especially for markets with low efficiency.

In this paper, we tried to show this methodological perspective through empirical implementation by using F-Score's delta determinants. On the other hand, this is not the only option. Also, the best-fit mix of determinants to produce most efficient F-FHSs can be investigated which is also closely related with the subject of value or behavioral relevance of accounting numbers.

REFERENCES

1. Aitken, B. (1998), Have institutional investors destabilized emerging markets?, *Contemporary Economic Policy*, vol. 16, no. 2, pp. 173-184, DOI 10.1111/j.1465-7287.1998.tb00510.x.
2. Balaban, E. (1995), Informational efficiency of the Istanbul Securities Exchange and some rationale for public regulation, *Research paper in banking and finance. UK: Institute of European Finance*, available online at https://core.ac.uk/download/pdf/7061411.pdf, accessed on 28.01.2017.
3. Ball, R. and Brown, P. (1968), An empirical examination of accounting income numbers, *Journal of Accounting Research*, vol. 6, no. 2, pp.159-178, DOI 10.2307/2490232.
4. Bezdek, J.C. (1981), Pattern recognition with fuzzy objective function algorithms, *Plenum Press, New York*, DOI 10.1007/978-1-4757-0450-1.
5. Bezdek, J.C. and Pal, S.K. (1992), *Fuzzy models for pattern recognition: Methods that search for structure in data*, IEEE Press, New York.
6. Brimble, M. and Hodgson, A. (2007), On the intertemporal value relevance of conventional financial accounting in Australia, *Accounting and Finance*, vol. 47, no. 4, pp. 599-662, DOI 10.1111/j.1467-629x.2007.00241.x.
7. Buguk, C. and Brorsen, B.W. (2003), Testing weak-form market efficiency: Evidence from the Istanbul Stock Exchange, *International Review of Financial Analysis*, vol. 12, no. 5, pp. 579-590, DOI 10.1016/s1057-5219(03)00065-6.
8. Cajueiro, D.O. and Tabak, B.M. (2004), Ranking efficiency for emerging markets, *Chaos, Solitons & Fractals*, vol. 22, no. 2, pp. 349-352, DOI 10.1016/j.chaos.2004.02.005.

9. Celikyilmaz, A. and Turksen, I.B. (2008), Enhanced fuzzy system models with improved fuzzy clustering algorithm, *IEEE Transitions on Fuzzy Systems*, vol. 16, no. 3, pp. 779–794, DOI 10.1109/tfuzz.2007.905919.
10. Celikyilmaz, A. and Turksen, I.B. (2009), *Modeling uncertainty with fuzzy logic: with recent theory and applications,* Springer-Verlag, Berlin Heidelberg.
11. Core, J.E., Guay, W.R. and Buskirk, A.V. (2003), Market valuations in the new economy: An investigation of what has changed, *Journal of Accounting and Economics*, vol. 34, no. 1-3, pp. 43–67, DOI 10.1016/s0165-4101(02)00087-3.
12. D'Ambrosio, C. (1980), Random walk and the stock exchange of Singapore, *Financial Review*, vol. 15, no. 2, pp. 1–12, DOI 10.1111/j.1540-6288.1980.tb00475.x.
13. Dunn, J.C. (1973), A fuzzy relative ISODATA process and its use in detecting compact well-separated clusters, *Journal of Cybernetics*, vol. 3, no. 3, pp. 32-57, DOI 10.1080/01969727308546046.
14. Fama, E.F. (1998), Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics*, vol. 49, no. 3, pp. 283-306, DOI 10.1016/s0304-405x(98)00026-9.
15. Fama, E.F., Fisher, L., Jensen, M.C. and Roll, R. (1969), The adjustment of stock prices to new information, *International Economic Review*, vol. 10, no. 1, pp. 1-21, DOI 10.2307/2525569.
16. Grieb, T.A. and Reyes, M.G. (1999), Random walk tests for Latin American equity indexes and individual firms, *Journal of Financial Research*, vol. 22, no. 4, pp. 371–383, DOI 10.1111/j.1475-6803.1999.tb00701.x.
17. Hammah, R.E. and Curran, J.H. (1998), Optimal delineation of joint sets using a fuzzy clustering algorithm, *International Journal of Rock Mechanics and Mining Sciences,* vol. 35, no. 4-5, pp. 495-496, DOI 10.1016/s0148-9062(98)00151-x.
18. Harrison, J.M. and Kreps, D.M. (1978), Speculative investor behavior in a stock market with heterogeneous expectations, *The Quarterly Journal of Economics*, vol. 92, no. 2, pp. 323-336, DOI 10.2307/1884166.
19. Harvey, C. (1993), Portfolio investment using emerging markets and conditioning information, *Washington, DC: World Bank. Working paper.*
20. Kawakatsu, H. and Morey, M.R. (1999), An empirical examination of financial liberalization and the efficiency of emerging market stock prices, *Journal of Financial Research*, vol. 22, no. 4, pp. 385–411, DOI 10.1111/j.1475-6803.1999.tb00702.x.
21. Kim, M. and Ramakrishna, R.S. (2005), New indices for cluster validity assessment, *Pattern Recognition Letters,* vol. 26, no. 15, pp. 2353–2363, DOI 10.1016/j.patrec.2005.04.007.
22. Malkiel, B.G. and Fama, E.F. (1970), Efficient capital markets: A review of theory and empirical work, *The Journal of Finance*, vol. 25, no. 2, pp. 383-417, DOI 10.1111/j.1540-6261.1970.tb00518.x.
23. Morris, S. (1996), Speculative investor behavior and learning, *The Quarterly Journal of Economics*, vol. 111, no. 4, pp. 1111-1133, DOI 10.2307/2946709.
24. Nanda, S.R., Mahanty, B. and Tiwari, M.K. (2010), Clustering Indian stock market data for portfolio management, *Expert System with Application*, vol. 37, no. 12, pp. 8793–8798, DOI 10.1016/j.eswa.2010.06.026.
25. Nefti, S. and Oussalah, M. (2004), Probalilistic-fuzzy clustering algorithm, În *2004 IEEE Intemational Conference on Systems, Man and Cybemetics,* pp. 4786–4791, DOI 10.1109/icsmc.2004.1401288.
26. Pal, N.R. and Bezdek, J.C. (1995), On cluster validity for the fuzzy c-means model, *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 370–379, DOI 10.1109/91.413225.
27. Pattarin, F., Paterlini, S. and Minerva, T. (2004), Clustering financial time series: an application to mutual funds style analysis, *Computational Statistics & Data Analysis*, vol. 47, no. 2, pp. 353-372, DOI 10.1016/j.csda.2003.11.009.
28. Piotroski, J.D. (2000), Value investing: the use of historical financial statement information to separate winners from losers, *Journal of Accounting Research*, vol. 38, no. 3, pp. 1-41, DOI 10.2307/2672906.
29. Tola, V., Lillo, F., Gallegati, M. and Mantegna, R.N. (2008), Cluster analysis for portfolio optimization, *Journal of Economic Dynamics and Control*, vol.

32, no. 1, pp. 235-258, DOI 10.1016/j.jedc.2007.01.034.

30. Urrutia, J.L. (1995), Test of random walk and market efficiency for Latin American emerging equity markets, *Journal of Financial Research*, vol. 18, no. 3, pp. 299-309, DOI 10.1111/j.1475-6803.1995.tb00568.x.
31. Wang, Y.J., and Lee, H.S. (2008), A clustering method to identify representative financial ratios, *Information Sciences*, vol. 178, no. 4, pp. 1087-1097, DOI 10.1016/j.ins.2007.09.016.
32. Xie, X.L. and Beni, G.A. (1991), Validity Measure for Fuzzy Clustering, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 3, no. 8, pp. 841-846, DOI 10.1109/34.85677.
33. Zhang, Y.J. (1996), A survey on evaluation methods for image segmentation, *Pattern Recognition*, vol. 29, no. 8, pp. 1335-1346, DOI 10.1016/0031-3203(95)00169-7.
34. Xie, X.L. şi G.A. Beni (1991), Validity Measure for Fuzzy Clustering, *IEEE Trans. Pattern and Machine Intelligence*, vol. 3, nr. 8, pp. 841-846.
35. Zhang, Y.J. (1996), A Survey on Evaluation Methods for Image Segmentation, *Pattern Recognition*, vol. 29, nr. 8, pp. 1335-1346.