Journal of Economic Cooperation and Development, **44**, 1 (2023), 111-138

Analysis of Turkish Deposit Banking Sector with Market Basket Analysis and Genetic Algorithm

Assoc.Prof.Dr. Mehmet ÖZÇALICI[1] and Assoc.Prof.Dr. Onur AKKAYA[2]

ABSTRACT

The paper aims to develop an integrated system with Genetic Algorithm and Market Basket Analysis to analyze the financial performance of deposit banks. Genetic Algorithm is used to discretize the banking dataset while Market Basket Analysis is used to extract rules. Sixty-seven ratios of thirteen deposit banks in the Turkish banking system currently traded on BİST are used with the proposed method. Data and ranking of 47 banks by asset sizes is acquired from The Banks Association of Turkey (BAT). The dataset covers the period between 1990-2018. Financial ratios, stock market performance and bank type are used as the dataset. A perfect positive correlation has been obtained between pairs of variables expressed below: total loans and receivables and total assets, total deposits and total assets, total deposits and total loans and receivables, total loans and total assets. A perfect negative correlation was obtained between the following variables: FC Liabilities/total liabilities, and TC liabilities/total liabilities, FC assets/total assets and TC assets/total assets. Using association rule mining to analyze Turkish banking sector, according to the case database, it is concluded that there are discernible [Net on Balance-sheet Position / Total Shareholders' Equity] characteristics of the Turkish banking sector, and a superposition relationship between [Net on Balance-sheet Position / Total Shareholders' Equity] and [Total Loans / Total Assets].

ملخص

يهدف المقال إلى تطوير نظام متكامل مع الخوارزمية الجينية وتحليل سلة السوق لتحليل الأداء المالي لبنوك الإيداع. ويتم استخدام الخوارزمية الجينية لحجب مجموعة البيانات المصرفية بينما يتم استخدام تحليل سلة السوق لاستخراج القواعد. ويتم استخدام 67 نسبة من ثلاثة عشر مصرفًا للودائع في النظام المصرفي التركي المتداول حاليا على بورصة إسطنبول مع الطريقة المقترحة. ويتم الحصول على بيانات وتصنيف 47 بنكا حسب أحجام الأصول من جمعية البنوك التركية .(BAT) وتغطي مجموعة البيانات الفترة بين 1990–2018. وتُستخدم النسب

[1] Department of International Trade and Logistics, Kilis 7 Aralık University, Kilis, Türkiye. E-mail: mozcalici@kilis.edu.tr
[2] Department of International Trade and Logistics, Kilis 7 Aralık University, Kilis, Türkiye. E-mail: onr.akkaya@gmail.com

المالية وأداء سوق الأسهم ونوع البنك كمجموعة بيانات.وجرى التوصل إلى علاقة إيجابية مثالية بين أزواج المتغيرات المبينة أدناه: إجمالي القروض والمبالغ المستحقة القبض وإجمالي الأصول، وإجمالي الودائع وإجمالي الأصول، وإجمالي الودائع وإجمالي القروض والمبالغ المستحقة القبض، وإجمالي القروض وإجمالي الأصول. كما سُجلت علاقة سلبية مثالية بين المتغيرات التالية: خصوم/ إجمالي خصوم التكاليف الثابتة(FC) ، وخصوم/ إجمالي خصوم التكاليف الإجمالية(TC) ، وأصول/ إجمالي أصول التكاليف الثابتة(FC) ، وأصول/ إجماي أصول التكاليف الإجمالية .(TC) وباستخدام أسلوب التنقيب عن قواعد الارتباط لتحليل القطاع المصرفي التركي، وفقا لقاعدة بيانات الحالة، تم استنتاج أن هناك خصائص [الصافي على وضع الميزانية العمومية/ إجمالي حقوق المساهمين] للقطاع المصرفي التركي، وعلاقة تراكب بين [صافي وضع الميزانية العمومية/ إجمالي حقوق المساهمين] و [إجمالي القروض/ إجمالي الأصول].

ABSTRAITE

Ce document vise à développer un système intégré avec un algorithme génétique et une analyse du panier de consommation pour analyser la performance financière des banques de dépôt. L'algorithme génétique est utilisé pour discrétiser l'ensemble des données bancaires, tandis que l'analyse du panier de biens et services est utilisée pour extraire des règles. Soixante-sept ratios de treize banques de dépôt du système bancaire turc actuellement cotées à la BİST sont utilisés avec la méthode proposée. Les données et le classement de 47 banques en fonction de la taille de leurs actifs proviennent de l'Association des banques de Türkiye (BAT). L'ensemble des données couvre la période 1990-2018. Les ratios financiers, la performance du marché boursier et le type de banque sont utilisés comme ensemble de données. Une corrélation positive parfaite a été obtenue entre les paires de variables exprimées ci-dessous : total des prêts et créances et total des actifs, total des dépôts et total des actifs, total des dépôts et total des prêts et créances, total des prêts et total des actifs. Une corrélation négative parfaite a été obtenue entre les variables suivantes : Passif du CF/total du passif, et Passif du CT/total du passif, Actif du CF/total de l'actif et Actif du CT/total de l'actif. En utilisant l'extraction de règles d'association pour analyser le secteur bancaire turc, selon la base de données de cas, il est conclu qu'il existe des caractéristiques discernables [Position nette au bilan / Total des capitaux propres] du secteur bancaire turc, et une relation de superposition entre [Position nette au bilan / Total des capitaux propres] et [Total des prêts / Total des actifs].

Keywords: Genetic Algorithm, Association Rule Mining, Banking Sector, Data Mining

JEL Classification: G21, E51, D40

1. Introduction

In 2001, the Turkish economy was lived a severe financial crisis. The economy management had implemented many sanctions which are called “GEGP ” program to remove the harmful effects of financial crisis and reorganized the economy. The program aims to protect the domestic economy for external shocks, decreasing inflation rate, complete financial reforms and strengthen the sector. During the reorganized period, the Turkish banking sector started substantial capital increase for recovery. Some of the banks were not attending capital increase, and then they merged or took over by Saving Deposit Insurance Fund of Turkey. The result of all, the total number of banks declined than 61 to 48.

After 2002 to 2008, the Turkish banking sector has seen stable growth. However, the last quarter of 2008 was that Gross Domestic Product (after their GDP) of Turkey was suddenly and sharply decreasing, saving deficit was rising, and the inflation rate was starting to fluctuate. At the beginning of 2009, the Turkish economy was again facing the economic crisis. Unlike in 2001, this crisis source was external facts like developed countries. The global financial system was profoundly affected by the Mortgage Crisis. Many developed and developing countries (including Turkey) to deal with the crisis took quick decisions to recover its economy. Even so, a lot of banks went bankrupt after the mortgage crisis like another financial crisis [1].

In 2011, uncertainties and risks in the financial sector had started to be concerned. Economic problems of institutions and the high public sector borrowing ratio in some European Union (after their EU) countries have caused uncertainty and risk. The use of the debt crisis statement instead of the global crisis has become widespread. Also, they have reduced the adverse effects (such as narrow the domestic savings deficit, credit growth measures have been taken to slow down) of uncertainties in the international markets.

Association rule mining successfully applied for various fields such as designing facility layout [2], hotel revenue management [3]; traffic data analysis [4–6]; determining operational problems of building heating [7]. Moreover, in the past few years, evolutionary computation-based rule mining systems have emerged [8]. Applying association rule mining to a dataset obtained from the banking sector can reveal insights about the sector. Banking dataset (such as financial ratios, stock market performance etc) are available in continuous type. To apply association rule mining, the dataset must be in categorical or Boolean type. In this study to discretize the banking dataset which is continuous in nature, the Genetic Algorithm is applied.

In this context, the paper aims that 67 ratios of thirteen deposit banks in the Turkish banking system currently traded on Borsa İstanbul (BİST) are used to analyse with Market Basket Analysis and Genetic Algorithm for 1990-2018 period. The second part of the paper has a literature review of the ratio of the banking sector as well as methodologies. The third and fourth parts are the methodology and results of the analysis, respectively. The final section is dedicated to conclusions.

2. Literature Review

Studies using financial ratios for analyze the banking sector can be summarized as follow.

Pehlivan analysed Turkish banking sector by ratio analysis between 2005 and 2014 [9]. According to the rates used in the study, the rate of return on equity of public-owned banks is higher than the rates of return of development and investment banks. On the other hand, in-branch efficiency and personnel productivity ratios, public capital with development and investment banks are in good condition. Altınay, examined the 26-year period covering the period 1990-2015 by using the ratios of the banking sector [10]. General equity ratio and asset quality ratio averages have been examined. It has been observed that the loans extended have an effect on decreasing asset quality rates. In the 2001 crisis period all ratios, except the equity ratio, were negatively affected.

Demirel and colleagues researched the operating ratios and the profitability trends of the banks in Turkey [11]. It was that return on assets, net profit margin, return on equity and the information related to the other operational expenses / total assets are analyzed in the research between the years 2002 March – 2012 June. Result of that (personnel expenses / total assets) and (operational expenses/total assests) ratios of public banks were worked significantly more efficiently. On the other hand, it had been found that banks with foreign capital have high values in (personnel expenses / total assets) and (operational expenses/total assests) ratios.

Some of the studies employed market basket analysis with another method can be summarized as follow:

Kuo and colleagues employed Particle Swarm Optimization and Genetic Algorithm to improve the computational efficiency of association rule mining [13]. They verified the proposed algorithm with some other datasets and with a case study on the stock market. Results indicate that their proposed strategy can be used for the formulation of marketing strategy. Kaya and Alhajj, proposed a genetic algorithm based framework for mining fuzzy association rules [14]. They emphasized that users cannot always intervene in the mining process. They proposed an automated GA based rule mining model. They report that the number of interesting rules obtained with the GA-based approach is greater than other models.

Ghosh and Nath, emphasized the multi-objective nature of association rule finding attempts and developed a GA based system to extract interesting and useful rules from market-basket type database [16]. It is reported that the proposed algorithm is suitable for large databases. Wang and colleagues addressed the importance of determining intervals when using numerical dataset in association rule mining [17]. They proposed an interestingness-based criterion for determining intervals to be merged.

Miller and Yang, considered the problem of mining association rules over interval data for which the separation between data points is essential [18]. They implemented the proposed algorithm on large real-life datasets. Djenouri and Comuzzi, argue that exact approaches to frequent itemset mining are suffered from poor runtime performance when dealing with large databases [19]. They proposed a bio-inspired approach based on GA and PSO.

Kumar and Singh, focused on the problem of generating association rules from numeric data [20]. They evaluated their proposed model on four different datasets and compared the results with four different modified rule extraction method and report that their proposed model has superior performance on other techniques.

3. Data and Methodology

3.1. Genetic Algorithm Optimization Process

Genetic algorithm which is introduced by John Holland [21] in 1970s and are biologically inspired search approaches that are suitable to a wide range of optimization problems [22].Genetic Algorithm (hereafter GA) is a repetitive investigation and optimization technique that is created by natural evolution and belongs to the stochastic class of techniques. GA works on a population of intended solutions for the problem in hand. Each member of the population is known as a chromosome.

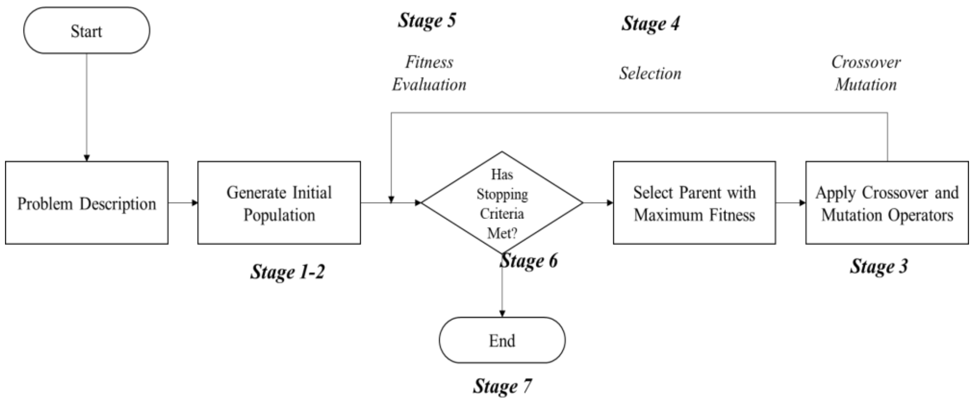

Figure 1: GA Procedure

GA's working stages can be summarized as follow [Figure.1]:

Stage.1: A solution Group is created where possible solutions are encoded. Solution due to its similarity in biology to the group population, codes of solutions see also it is called a chromosome. The number of individuals in the population starts by identifying. There is no standard for this number. Generally recommended It is a magnitude in the range of 100 - 300. Operations on size selection, the complexity and depth of the search are essential.

Stage.2: After this process, the population is randomly generated. It is found how good each chromosome is. The function that finds that it is as good is called the fitness function. This function is the only part that works specifically for the problem in GA. Most of the time, the success of GA is connected with the efficiency and sensitivity of this function.

Stage.3: Copying and modifying these chromosomes by mapping operators are implemented. In this way, a new population is created. The chromosomes are mapped according to the fitness values of chromosomes. There are selection methods such as Roulette wheel selection, tournament selection to make this selection.

Stage.4: Eliminate old chromosomes to make room for new chromosomes. The old chromosomes are removed.

Stage.5: All chromosomes N are re-evaluated with fitness function. The success of the new population can be found by recalculating chromosomes.

Stage.6: GA is run many times to create a large number of populations is calculated. If time is not up or another stopping criterion has not been met, it will be returned to stage 3.

Stage. 7: The individual with has the best fitness value is determined as an optimal solution.

Genetic Algorithm has advantages and disadvantages [25]. Some of the advantages are: parallelism, liability, using only function evaluations, can be easily modified for different problems, can handle large or poorly understood search spaces easily. Limitations are: the problem of identifying fitness function, the definition and representation of the problem, cannot use gradients, no effective terminator and require a large number of response (fitness) function evaluations.

3.2. Market Basket Analysis

Association Rules

Association rules are developed in the computer science field. They are often used in important applications such as measuring the association between products purchased by a particular consumer or for measuring the associations between pages viewed sequentially by a visitor [26]. The objective of association rules is to uncover hidden patterns in large data sets by detecting relationships or associations between specific values of categorical variables [27]. An association rule is a statement between two item sets and can be written as $x \rightarrow y$ where both x and y are item sets. Rule indicates that x and y occur together, that is if x occurs, then y also occurs.

The strength and accuracy of the association rule is measured using following three metrics:

- **Support:** Support of x and y together is represented as the number of times they appear together divided by total number of transactions in data set n.

$$S(x \rightarrow y) = \frac{Number\ of\ times\ x\ and\ y\ appear\ together}{Number\ of\ transactions} = \frac{n_{x\rightarrow y}}{n} \tag{1}$$

Support values will be always in the range [0,1]. Support value of zero indicates that rule is never occurred in transaction dataset. Support value of one indicates that rule is observed in each and every record.

- **Confidence:** Conditional probability of y when x already occurred plays a significant role when evaluating the strength of the rule. Confidence for $x \rightarrow y$ is defined as the ratio of the support for x and y together to the support for x [28]. In other words it is the frequency of occurrence of y, conditionally on x being true [26].

$$C(x \rightarrow y) = \frac{support(xy)}{support(x)} = \frac{n_{x\rightarrow y}}{n_x} \tag{2}$$

Confidence values will be in the range of [0,1]. A confidence value equals to one indicates that all of the transactions where x is occurred, y is also occurred. Similarly a confidence value equals zero indicates that none of the transactions include x, also includes y.

- **Lift:** Lift is the ratio between the probability that both elements occur together and the relative probability of the same event. The removal value indicating the positive relationship is greater than 1, and the removal value indicating the negative relationship is less than 1.

$$L(a \rightarrow b) = \frac{C(x \rightarrow y)}{S(y)} \tag{3}$$

A lift ratio greater than 1.0 suggests that the level of association between the antecedent and consequent item sets is higher than would be expected if they were independent [29]. In other words, if the result is greater than one, then x and y are positively correlated (when one is present, the other is likely to be present) [30]. If the lift value is less than one, then the appearance of x and y are negatively correlated (when one is present the other is likely to be absent). The larger the lift ratio, the greater the strength of the association. If the result of the formula is equal to one, then x and y are independent, and there is no correlation between the two [30].

The task in association rule discovery is to find all rules fulfilling given pre-specified frequency and accuracy criteria [31]. Apriori algorithm developed by the studies of Agrawal can be used for extraction rules from large dataset [32].

It is possible to create enormous number of rules. In order to limit the number of rules some minimal support (minsup) and minimal confidence (minconf) thresholds can be considered. Thus only the rules that satisfy the following inequalities are considered as valid.

$$S(x \to y) \geq minsupp$$

$$C(x \to y) \geq minconf.$$

4. Empirical Results

Proposed Method

Types of the financial ratio dataset of banks are continuous in nature. On the other hand, the apriori algorithm requires Boolean data type. It is possible to use the quantitative dataset in the apriori algorithm by discretizing dataset [34,35].

In discretization process, the values are mapped to consecutive integers such that the orders of the values are preserved. In order to map continuous dataset to Boolean type, users have to determine the number of groups (intervals). The number of intervals should be determined carefully. If intervals are too large, important rules at smaller resolution may be missed; and if they are too small, there may not be enough data to mine rules [35]. For example, consider 'total deposit sector share' ratio which is a financial ratio used in this study. It is distributing in the range of [0.9555 100]. It is possible to discretize this feature into two discrete groups where the banks having total deposit share up to 50% will be assigned to the first group while other banks will be assigned to the second group. Moreover, it is also possible to discrete this feature into 5 equal-width groups with intervals [0, 20), [20,40), [40,60), [60,80), [80, 100]. It is unknown that in which case the apriori algorithm will produce rules with higher lift values. Moreover, there are 64 features that must be considered in this study (Actually, there are 65 features however 'bank characteristics' feature has only three values, which indicates that it has already discretized). In other words, the user must determine the number of groups that will be used in the discretization process for each feature. If there are 9 groups (2, 3, …, 10) that will be considered and there are 64 features, there will be 9^{64} possibilities in the search space. Searching for all possibilities will not be terminated in a reasonable time. An intelligent approach must be developed.

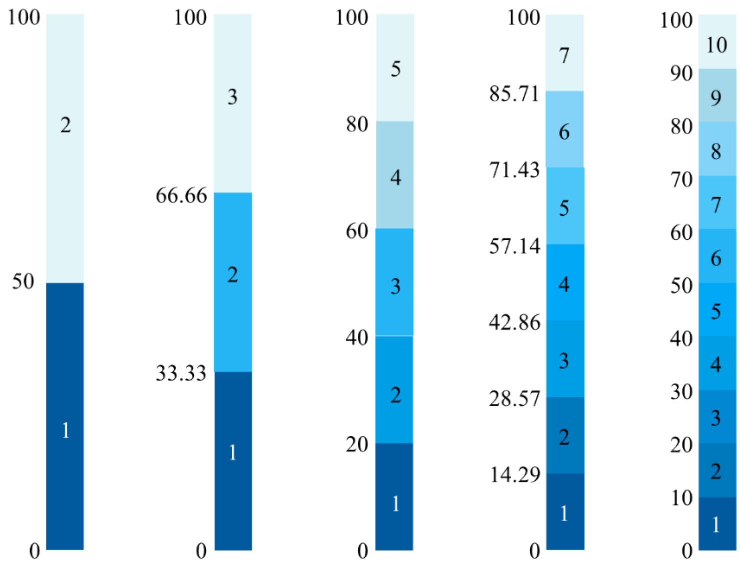

Figure 2: Discretization of dataset into different number of intervals.

In this study GA optimization process is employed to determine the number of groups for each feature [Figure.2, 3 and 4].

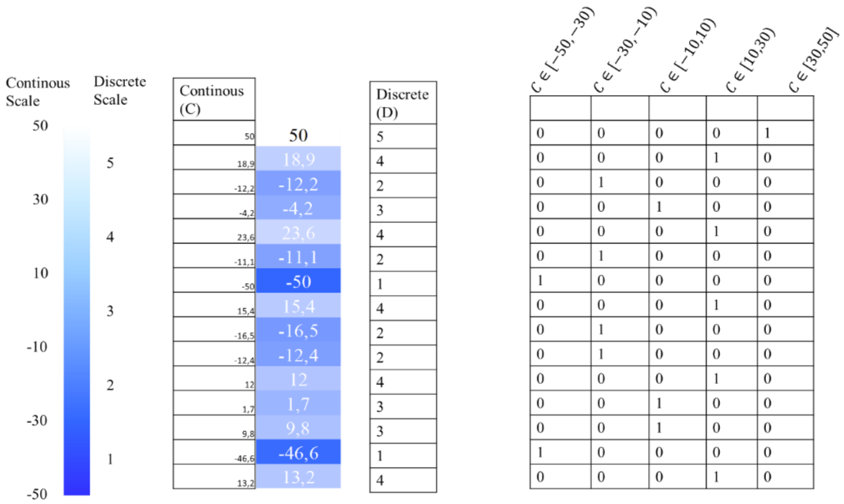

Figure 3: Mapping continuous feature to a Boolean type.

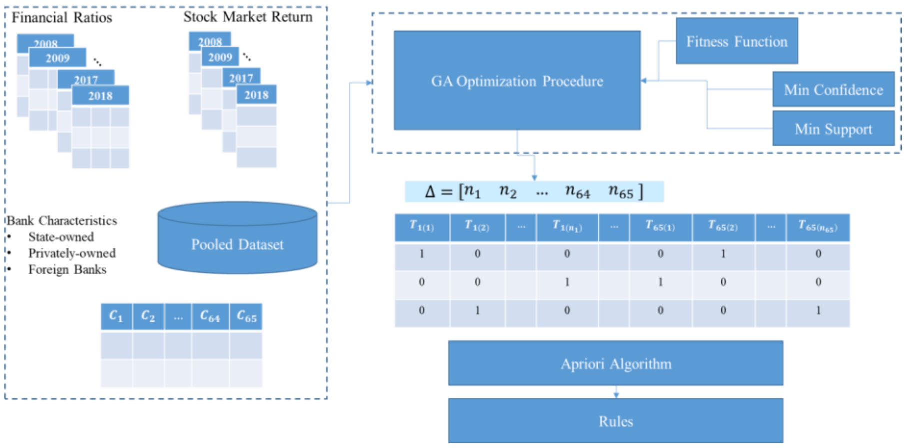

Figure.4 Model of the Proposed System.

Dataset Description

Bank Selection

Financial ratios belong to 47 banks are published by The Banks Association of Turkey (BAT)[3]. These include deposit, development and investment banks (consisting of 14 development and investment banks and 33 deposit banks). The fact that the banks considered in the study

[3] Retrieved from “tbb.org.tr” in 30.06.2019.

show similar characteristics to each other is important for the rules to be meaningful. The rates of banks with only one branch and the rates of banks operating widely throughout the country are different. To obtain a homogeneous data set, only deposit banks were considered in the study. However, it was preferred to select deposit banks with significant weight in the sector. 13 banks with a total deposit rate of more than 1% were taken into account in this study.

Banks and their characteristics are as shown in the [Table.1].

Table 1: Selected Banks

Number	Bank Name	Stock Quote	Total Deposits (TL)	Total Deposits Share (%)	Bank Characteristic	Establishment Year
1	Türkiye Cumhuriyeti Ziraat Bankası A.Ş.	-	392606	17.59	State-owned	1863
2	Türkiye İş Bankası A.Ş.	ISCTR	258852	11.60	Privately-owned	1924
3	Türkiye Halk Bankası A.Ş.	HALKB	263580	11.81	State-owned	1938
4	Yapı ve Kredi Bankası A.Ş.	YKBNK	217578	9.75	Privately-owned	1944
5	Türkiye Garanti Bankası A.Ş.	GARAN	231802	10.39	Foreign	1946
6	Türkiye Vakıflar Bankası T.A.O.	VAKBN	214295	9.60	State-owned	1954
7	Akbank T.A.Ş.	AKBNK	206235	9.24	Privately-owned	1948
8	QNB Finansbank A.Ş.	QNBFB	95274	4.27	Foreign	1987
9	Denizbank A.Ş.	DENIZ	89513	4.01	Foreign	1997
10	Türk Ekonomi Bankası A.Ş.	-	65859	2.95	Privately-owned	1927
11	ING Bank A.Ş.	-	36195	1.62	Foreign	1984
12	HSBC Bank A.Ş.	-	25032	1.12	Foreign	1990
13	Şekerbank T.A.Ş.	SKBNK	24115	1.08	Privately-owned	1953
	Sector Total		2231634			

Features or Variables

There are three types of features in the dataset

- Financial ratios of banks
- Stock market returns
- Bank characteristics

Dataset compiled from two different data sources. The first data source includes the financial ratios of deposit banks operating in Turkey. The Banks Association of Turkey (tbb.org.tr) issues selected ratios of banks which are based on financial statements of banks. Association reports 63 ratios under 10 main categories. In the study, we used annual financial ratios on the institution's website between 2008 and 2018. In addition, the information about whether banks are state-owned, privately-owned or foreign banks is given in the table.

The stock market dataset is retrieved from investing.com which provides historical price and volume information of stocks in numerous stock markets. The shares of only nine of the 13 banks listed in the table are traded on the Borsa İstanbul (BİST). The historical closing prices of the nine banks whose shares are traded on the stock exchange in Borsa Istanbul are shown in the figure. The earliest data on 4 May 2010 were available. Older data were not available on the site. Stock closing prices start from 6 November 2012 for Şekerbank and 24 March 2014 for QNB Finansbank.

Figure 5: Historical Closing Price of Banks.

When we look at the “closing price” change of 9 different banks, it is seen that volatility is high in terms of periods. In the series discussed in general, it can be said that while “closing prices” peaked in 2013, 2015 and 2018 periods, they reached the bottom position in the 2014 period. HALKBANK is at the highest price in terms of “closing price” values. The lowest price belongs to Şekerbank [Figure.5].

In order to calculate the stock market performance, stock returns are calculated based on the following equation:

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} \tag{4}$$

[Eq.1] where, p_t and p_{t-1} represents the price of the day t and day $t-1$ respectively. For each year, sum of the daily returns is calculated and results are presented in [Table.2].

Table 2: Sum of Daily Returns of Stocks for different years

Banks	**2018**	**2017**	**2016**	**2015**	**2014**	**2013**	**2012**	**2011**	**2010**
AKBNK	0.4267	-0.2303	0.0639	0.3409	-0.2395	0.3767	-0.3159	0.4255	-0.1195
DENIZ	-1.3444	0.2249	-0.9065	0.0317	0.7338	2.3576	0.2001	-0.0442	0.0502
GARAN	0.4188	-0.3415	-0.0617	0.3703	-0.277	0.4179	-0.4038	0.3548	-0.0815
HALKB	0.5489	-0.0846	0.1232	0.389	-0.1289	0.5376	-0.5421	0.0867	-0.0716
ISCTR	0.5279	-0.3114	-0.1353	0.4369	-0.3638	0.3562	-0.5501	0.4601	-0.0998
QNBFB	-0.312	-0.102	0.1493	-0.3468	-0.1824	0	0	0	0
SKBNK	0.5937	-0.3747	0.3978	0.0319	-0.1968	0.3826	-0.0733	0	0
VAKBN	0.6702	-0.4084	-0.1042	0.3423	-0.2114	0.3195	-0.5863	0.5724	-0.0471
YKBNK	0.6575	-0.1968	0.5321	0.4225	-0.2748	0.4505	-0.5706	0.7039	-0.1371

Thus, 63 financial ratios, stock returns and bank characteristics one year, a total of 65 within the scope of this study variable are combined. The data set is a typical example of panel data. Observations of 13 banks between 2008 and 2018 (11 years) have been gathered together in the form of pooled data. Thus a pooled data set consisting of 143 (=11×13) rows and 65 columns was used. Initial feature pool and its descriptive statistics are presented in [Table.3].

Table 3: Initial Feature Pool and Descriptive Statistics

	Feature Name	Min	Max	Mean	Std
f01	Capital Adequacy Ratio	12.57	23.22	16	2.08
f02	Shareholders Equity / Total Assets	7.05	18.01	11	1.76
f03	(Shareholders Equity-Permanent Assets) / Total Assets	1.76	13.86	7.64	2.2
f04	Shareholders Equity / (Deposits + Non-Deposit Funds)	8.48	23.95	13.81	2.6
f05	On Balance-sheet FC Position / Shareholders Equity	-23.53	235.12	78.28	64.35
f06	Net on Balance-sheet Position / Total Shareholders Equity	-185.93	39.69	-51.88	52.46
f07	N(on+off) Balance-sheet Position / Total Shareholders Equity	-17.48	12.25	0.15	5.23
f08	TC Assets / Total Assets	41.7	87.63	69	8.59
f09	FC Assets / Total Assets	12.37	58.3	31	8.59
f10	TC Liabilities / Total Liabilities	38.84	84.7	60.51	8.73
f11	FC Liabilities / Total Liabilities	15.3	61.16	39.49	8.73
f12	FC Assets / FC Liabilities	40.17	108	79.13	16.28
f13	TC Deposits / Total Deposits	29.34	82.17	61.49	9.55
f14	TC Loans and Receivables / Total Loans and Receivables	55.99	95.96	73.84	9.23
f15	Total Deposits / Total Assets	48.02	83.22	61.51	6.53
f16	Funds Borrowed / Total Assets	0.02	30.65	11.48	6.3
f17	Financial Assets (Net) / Total Assets	4.56	57.01	21.04	10.16
f18	Total Loans / Total Assets	29.49	75.02	61.87	8.06
f19	Total Loans / Total Deposits	36.76	148.86	102.18	18.76
f20	Loans under follow-up (gross) / Total Loans	1.21	12	4.04	1.87
f21	Permanent Assets / Total Assets	0.88	10.52	5.02	2.05
f22	Consumer Loans / Total Loans	4.77	66.71	31.72	10.52
f23	Liquid Assets / Total Assets	8.61	45.18	26.42	7.83
f24	Liquid Assets / Short-term Liabilities	14	105.49	46.79	14.29
f25	TC Liquid Assets / Total Assets	0.89	36.52	14.07	7.45
f26	Liquid Assets / (Deposits + Non-Deposit Funds)	10.16	56.02	32.64	9.63
f27	FC Liquid Assets / FC Liabilities	10.86	62.45	31.84	10.3
f28	Average Return on Assets	-1.27	3.07	1.56	0.71
f29	Average Return on Shareholders Equity	-14.45	39.64	14.11	6.79
f30	Income Before Taxes / Total Assets	-1.45	3.58	1.78	0.8
f31	Net Profit (Losses) / Paid-in Capital	-54.67	298.04	67.91	55.63
f32	Net Interest Income After Specific Provisions / Total Assets	1.31	6.78	3.35	0.98
f33	Net Interest Income After Specific Provisions / Total Operating Income	24.6	100.71	59.9	12.39
f34	Non-Interest Income (Net) / Total Assets	-0.5	3.42	1.48	0.67
f35	Non-Interest Income (Net) / Other Operating Expenses	-11.78	137.72	59.77	27.43
f36	Other Operating Expenses / Total Operating Income	20.23	76.9	47.31	12.53
f37	Provision For Loan or Other Receivables Losses / Total Assets	0	3.43	1.18	0.59
f38	Interest Income / Interest Expense	128.26	297.78	194.34	27.28
f39	Total Income / Total Expense	114.2	185.58	142.2	14.17
f40	Interest Income / Total Assets	6.2	16.44	8.98	2.08
f41	Interest Expense / Total Assets	2.67	9.13	4.73	1.4
f42	Interest Income / Total Expenses	66.81	104.51	85.72	5.89
f43	Interest Expense / Total Expenses	36.55	90.31	63.69	10.3
f44	Share in Sector - Total Assets	0.8	100	14.35	25.21
f45	Share in Sector - Total Loans and Receivables	0.65	100	14.27	25.18
f46	Share in Sector - Total deposits	0.96	100	14.69	25.18
f47	Share in Group - Total Assets	0.84	16.24	7.29	4.76

	Feature Name	Min	Max	Mean	Std
f48	Share in Group - Total Loans	0.71	17.23	7.3	4.44
f49	Share in Group - Total Deposits	0.96	20.47	7.31	4.93
f50	Total Assets / No. of Branches	32.17	419.52	152.99	91.49
f51	Total Deposits / No. of Branches	23.73	303.5	92.78	54.75
f52	TRY Deposits / No. of Branches	13.19	140.86	54.08	26.06
f53	FX Deposits / No. of Branches	6.29	211.59	38.7	32.39
f54	Total Loans and Receivables* / No. of Branches	19.16	252.14	95.26	58.82
f55	Total Employees / No. of Branches (person)	12.98	35.42	18.16	2.85
f56	Net Income / No. of Branches	-3.96	7.54	2.17	1.62
f57	(Personnel Expenses + Reserve for Employee Termination Benefit) / Total Assets	0.55	2.88	1.24	0.5
f58	(Personnel Expenses + Reserve for Employee Termination Benefit) / Number of Personnel (Thousand TRY)	44.39	198.78	87.16	31.24
f59	Reserve for Employee Termination Benefit / Number of Personnel (Thousand TRY)	0	9.88	1.69	1.62
f60	Personnel Expenses / Other Operating Expenses	35.83	55.21	42.97	4.34
f61	Other Operating Expenses / Total Asset	1.43	6.47	2.8	1.04
f62	Total Operating Income / Total Assets	2.36	12.08	5.65	1.43
f63	Net Operating Income(Loss) / Total Assets	-1.45	3.58	1.74	0.81
f64	Stock Market Return	-1.34	2.36	0.08	0.49
f65	Bank Characteristic	1	3	2.15	0.77

Correlation among Features

In order to investigate the dataset, correlation coefficients among all features are calculated. Since the number of correlation coefficients is 2080 $(= 65!/(2! \times (65-2)!))$, it would be impractical to report all of the correlations coefficients. So heatmap of correlation coefficients is presented as a heatmap format [Figure.6].

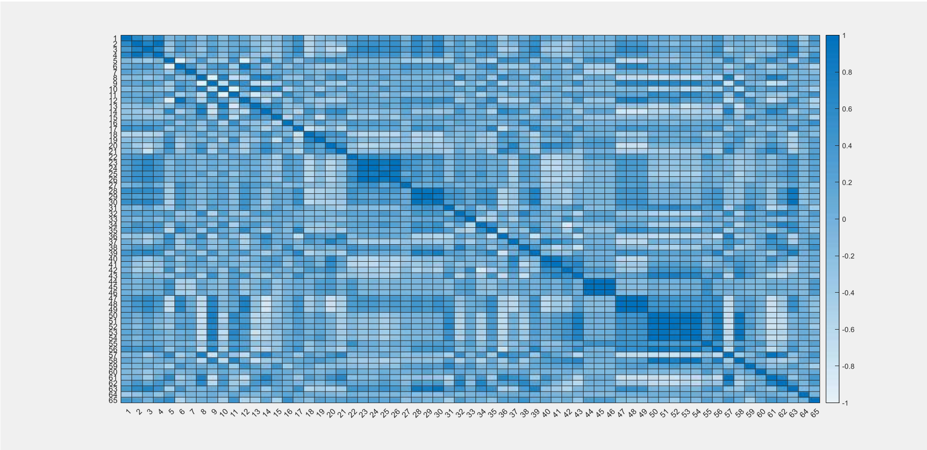

Figure 6: Heat map of correlation coefficients.

Although all of the correlation coefficients are shown as heat graphs, some of the prominent correlation coefficients will be summarized here. A perfect positive correlation has been obtained between pairs of variables expressed below: total loans and receivables and total assets, total deposits and total assets, total deposits and total loans and receivables, total loans and total assets. A perfect negative correlation was obtained between pairs of variables. These; FC Liabilities/total liabilities and TC liabilities/total liabilities, FC assets/total assets, and TC assets/total assets. This is the correlation coefficient between pairs of variables. Total employees/no of branches (person) and net interest income after specific provisions/total assets, total assets/number of branches and provision for a loan or other receivables losses / total assets, other operating expenses/total assets, and funds borrowed/total assets.

Determining Minimum Support And Minimum Confidence Threshold Values

The minimum support and minimum confidence threshold values are required for the Apriori algorithm to reveal the rules and to work with another expression. In other words, minimum support and minimum confidence values must be determined by the user. To determine the minimum support and minimum confidence values a series of experiments are carried out. For each discretization number n_i $(i = 2, 3, \dots, 10)$, 41 values ([0.1 0.11 0.12 0.13 … 0.48 0.49 0.50] in step 0.01) in the range of 0.1 to 0.5 were tried as both minimum support and minimum confidence and the results are presented in Figure. In each chart, all the variables in the data set were subjected to a number of discretization processes determined in the individual chart. For example, in the subfigure located in the upper left of the figure X, the dataset is discretized to two categories (n=2). Minimum support and minimum confidence values are depicted in horizontal axes while the mean of lift values of rules whose lift value is greater than 2 is depicted in the vertical axes.

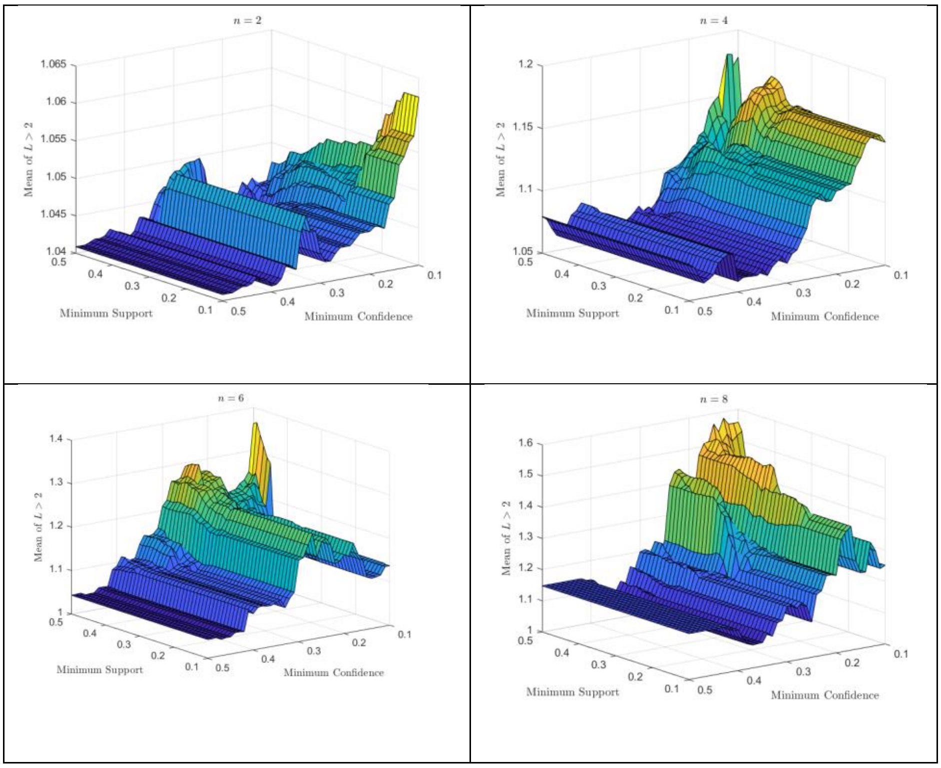

Figure 7: Change of Mean of lift values greater than 2 for different support, confidence and discretization number.

From [Figure.7], it is possible to assert that the minimum confidence threshold value has more impact on the lift values (mean of lift values greater than 2) than minimum support threshold. The increase or decrease of the minimum confidence value allows the average value to be moved to different levels. A strong rule must have a large confidence and support value and a lift value greater than one. In some of the previous studies, minimum support and confidence threshold values are determined by researchers [5,36–38]. In this study, minimum support and minimum confidence thresholds are fixed at 0.2.

Optimizing number of categories for each variable with Genetic Algorithm

In this context, after the dataset is prepared for the analysis, the GA optimization procedure is started. In this study lift value is accepted as more important for determining the strength of an association rule [5].

The aim of GA employed in this study is to maximize the number of rules that has lift values greater than two by changing the number of categories for each variable. To transform the problem to minimization type, the number of rules that has lift values greater than two is multiplied by (-1). During iterations, an individual that has the highest fitness value is survived to the next generation. While individuals who have the lowest fitness value is eliminated. In other words, strong individuals are candidates for the next generation. In this experimental setting, the elitist strategy is adapted. In this strategy, the best solutions are copied to the next generation. In this study, three individuals who have the best fitness values are passed to the next generation without any modification such as crossover or mutation.

Function: Fitness function
input: dataset and candidate solution **output:** fitness value **dataset:** a matrix including continuous features with 143 rows and 65 columns. **candidate solution:** $N = [n_1 \quad n_2 \quad \dots \quad n_{64} \quad n_{65}]$ - Create matrix D by discretizing each variable with corresponding n_i Matrix D has 143 rows and 65 columns consisting from discrete values. - Create matrix T by transforming discrete values to binary values Matrix T has 143 rows and $\sum n_i$ columns - Generate rules by applying apriori algorithm with minimum support and minimum confidence thresholds - **output:** calculate number of rules which have lift values greater than 2.

Figure 8: Fitness function procedure.

Lower bound and upper bound for the number of categories are determined as 2 and 10 respectively. The elite count is fixed at 3. The number of generations is and stall generation limit is fixed at 100. There are 100 individuals in populations. Crossover probability is 0.8. Genetic algorithm procedure stopped running after 100 generations [Figure.8].

Results of the genetic algorithm process are presented in [Table.4]. This table includes the optimal number of categories for each variable as well as intervals in these optimal categories. For example, the optimal number of categories for variable "f05 - On Balance-sheet FC Position / Shareholders Equity" ratio is determined as 4. Interval for this feature is [-60, 20), [20, 100), [100, 180) and [180, 260). Mean of the categories is calculated as 6.125

Table 4: Results of GA optimization procedure.

Label	Optimal n_i^*	Intervals
f01	9	[12, 13.3, 14.6, 15.9, 17.2, 18.5, 19.8, 21.1, 22.4, 23.7]
f02	9	[7, 8.3, 9.6, 10.9, 12.2, 13.5, 14.8, 16.1, 17.4, 18.7]
f03	10	[1, 2.3, 3.6, 4.9, 6.2, 7.5, 8.8, 10.1, 11.4, 12.7, 14]
f04	9	[8, 9.8, 11.6, 13.4, 15.2, 17, 18.8, 20.6, 22.4, 24.2]
f05	4	[-60, 20, 100, 180, 260]
f06	8	[-200, -170, -140, -110, -80, -50, -20, 10, 40]
f07	5	[-20, -13, -6, 1, 8, 15]
f08	4	[40, 52, 64, 76, 88]
f09	7	[12, 19, 26, 33, 40, 47, 54, 61]
f10	7	[36, 43, 50, 57, 64, 71, 78, 85]
f11	7	[12, 20, 28, 36, 44, 52, 60, 68]
f12	6	[40, 52, 64, 76, 88, 100, 112]
f13	8	[24, 32, 40, 48, 56, 64, 72, 80, 88]
f14	9	[52, 56.9, 61.8, 66.7, 71.6, 76.5, 81.4, 86.3, 91.2, 96.1]
f15	8	[48, 52.5, 57, 61.5, 66, 70.5, 75, 79.5, 84]
f16	3	[0, 11, 22, 33]
f17	8	[0, 8, 16, 24, 32, 40, 48, 56, 64]
f18	3	[20, 39, 58, 77]
f19	4	[20, 60, 100, 140, 180]
f20	6	[1, 2.9, 4.8, 6.7, 8.6, 10.5, 12.4]
f21	4	[0, 2.7, 5.4, 8.1, 10.8]
f22	5	[0, 14, 28, 42, 56, 70]
f23	8	[8, 12.7, 17.4, 22.1, 26.8, 31.5, 36.2, 40.9, 45.6]
f24	7	[10, 24, 38, 52, 66, 80, 94, 108]
f25	7	[0, 5.3, 10.6, 15.9, 21.2, 26.5, 31.8, 37.1]
f26	6	[7, 16, 25, 34, 43, 52, 61]
f27	6	[8, 18, 28, 38, 48, 58, 68]
f28	8	[-1.5, -0.92, -0.34, 0.24, 0.82, 1.4, 1.98, 2.56, 3.14]
f29	8	[-18, -10, -2, 6, 14, 22, 30, 38, 46]
f30	7	[-2.1, -1.2, -0.3, 0.6, 1.5, 2.4, 3.3, 4.2]
f31	8	[-80, -32, 16, 64, 112, 160, 208, 256, 304]
f32	5	[1, 2.2, 3.4, 4.6, 5.8, 7]
f33	6	[20, 34, 48, 62, 76, 90, 104]
f34	4	[-0.9, 0.2, 1.3, 2.4, 3.5]
f35	4	[-30, 20, 70, 120, 170]
f36	6	[18, 28, 38, 48, 58, 68, 78]
f37	6	[0, 0.6, 1.2, 1.8, 2.4, 3, 3.6]
f38	4	[120, 170, 220, 270, 320]
f39	10	[112, 119.4, 126.8, 134.2, 141.6, 149, 156.4, 163.8, 171.2, 178.6, 186]
f40	5	[6, 8.1, 10.2, 12.3, 14.4, 16.5]

Label	Optimal n_i^*	Intervals
f41	7	[1.8, 2.9, 4, 5.1, 6.2, 7.3, 8.4, 9.5]
f42	5	[63, 72, 81, 90, 99, 108]
f43	4	[30, 46, 62, 78, 94]
f44	7	[0, 15, 30, 45, 60, 75, 90, 105]
f45	5	[0, 20, 40, 60, 80, 100]
f46	10	[0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100]
f47	2	[0, 9, 18]
f48	5	[0, 3.5, 7, 10.500, 14, 17.5]
f49	9	[0, 2.3, 4.6, 6.9, 9.2, 11.5, 13.8, 16.1, 18.4, 20.7]
f50	4	[0, 110, 220, 330, 440]
f51	7	[0, 44, 88, 132, 176, 220, 264, 308]
f52	5	[0, 29, 58, 87, 116, 145]
f53	5	[0, 50, 100, 150, 200, 250]
f54	3	[0, 90, 180, 270]
f55	4	[10, 17, 24, 31, 38]
f56	8	[-4, -2.5, -1, 0.5, 2, 3.5, 5, 6.5, 8]
f57	5	[0.4, 0.9, 1.4, 1.9, 2.4, 2.9]
f58	4	[30, 80, 130, 180, 230]
f59	8	[0, 1.3, 2.6, 3.9, 5.2, 6.5, 7.8, 9.1, 10.4]
f60	6	[33, 36.8, 40.6, 44.4, 48.2, 52, 55.8]
f61	7	[1.4, 2.2000, 3, 3.8, 4.6, 5.4, 6.2, 7]
f62	5	[2, 4.1, 6.2, 8.3, 10.4, 12.5]
f63	7	[-2.1, -1.2, -0.3, 0.6, 1.5, 2.4, 3.3, 4.2]
f64	8	[-1.6, -1.1, -0.6, -0.1, 0.4, 0.9, 1.4, 1.9, 2.4]
f65	3	[0.6, 1.4, 2.2, 3]
Total	401	
Mean	6.125	

Application of Apriori Algorithm and Rules

After determining the optimal n_i values with Genetic Algorithms, features in the dataset is discretized with using the optimal n_i values. Discretized features are transformed to the Boolean data type. As a result, the final transformed matrix consists of 143 rows and 401 columns. Minimum support and minimum confidence threshold values are fixed as 0.2 as used in the fitness function.

When the final Apriori algorithm is executed at intervals determined by the genetic algorithm, 150 rules are created. However, due to the small volume of the data set, predecessors and successors are included in the rules are in pairs. In other words, in all 150 rules, predecessor – successor and successor – predecessor rules are listed consecutively. Rules with a low confidence value are eliminated from these pairs of rules, and rules with high confidence values are left in the dataset. Thus, 75 rules have been extracted.

Table 5: Top Twenty Rules with the Highest Lift Value

	Antecedent	Consequent	Lift	Support	Confidence	# of transactions	Interval of Antecedent	Interval of Consequent	Correlation Coefficient
Rule 1	Shareholders Equity / (Deposits + Non-Deposit Funds) (3)	Shareholders Equity / Total Assets (3)	**3.1544**	20.979	88.2353	30	[11.60,13.40]	[9.600,10.90]	0.96
Rule 2	On Balance-sheet FC Position / Shareholders Equity (1)	Net on Balance-sheet Position / Total Shareholders Equity (7)	**2.7427**	20.2797	80.5556	29	[-60,20]	[-20,10]	-0.98
Rule 3	Shareholders Equity / Total Assets (4)	Shareholders Equity / (Deposits + Non-Deposit Funds) (4)	**2.6074**	22.3776	82.0513	32	[10.90,12.20]	[13.40,15.20]	0.96
Rule 4	Net on Balance-sheet Position / Total Shareholders Equity (6)	On Balance-sheet FC Position / Shareholders Equity (2)	**2.4655**	21.6783	100	31	[-50,-20]	[20,100]	-0.98
Rule 5	On Balance-sheet FC Position / Shareholders Equity (1)	FC Assets / FC Liabilities (5)	**2.432**	20.979	83.3333	30	[-60,20]	[88,100]	-0.94
Rule 6	FC Assets / FC Liabilities (4)	On Balance-sheet FC Position / Shareholders Equity (2)	**2.2344**	20.2797	90.625	29	[76,88]	[20,100]	-0.94
Rule 7	On Balance-sheet FC Position / Shareholders Equity (1)	Interest Expense / Total Expenses (3)	**2.0187**	21.6783	86.1111	31	[-60,20]	[62,78]	-0.30
Rule 8	TC Assets / Total Assets (2)	On Balance-sheet FC Position / Shareholders Equity (2)	**1.6787**	22.3776	68.0851	32	[52,64]	[20,100]	0.37
Rule 9	Average Return on Assets (6)	On Balance-sheet FC Position / Shareholders Equity (2)	**1.375**	20.2797	55.7692	29	[1.400,1.980]	[20,100]	-0.42
Rule 10	Shareholders Equity / (Deposits + Non-Deposit Funds) (4)	Income Before Taxes / Total Assets (5)	**1.3427**	20.979	66.6667	30	[13.40,15.20]	[1.500,2.400]	0.56
Rule 11	On Balance-sheet FC Position / Shareholders Equity (2)	Interest Income / Total Assets (1)	**1.3149**	22.3776	55.1724	32	[20,100]	[6,8.100]	0.33
Rule 12	Capital Adequacy Ratio (2)	Share in Group - Total Assets (1)	**1.2456**	20.979	73.1707	30	[13.30,14.60]	[0,9]	0.23
Rule 13	Capital Adequacy Ratio (2)	Share in Sector - Total deposits (1)	**1.2335**	20.2797	70.7317	29	[13.30,14.60]	[0,10]	0.01
Rule 14	On Balance-sheet FC Position / Shareholders Equity (2)	Average Return on Shareholders Equity (4)	**1.2328**	21.6783	53.4483	31	[20,100]	[6,14]	-0.41
Rule 15	Capital Adequacy Ratio (2)	Total Loans / Total Assets (3)	**1.2254**	27.2727	95.122	39	[13.30,14.60]	[58,77]	-0.57
Rule 16	On Balance-sheet FC Position / Shareholders Equity (2)	Income Before Taxes / Total Assets (5)	**1.2154**	24.4755	60.3448	35	[20,100]	[1.500,2.400]	-0.50
Rule 17	On Balance-sheet FC Position / Shareholders Equity (2)	Provision For Loan or Other Receivables Losses / Total Assets (2)	**1.2154**	24.4755	60.3448	35	[20,100]	[0.6000,1.200]	0.44
Rule 18	Shareholders Equity / Total Assets (4)	Total Loans / Total Deposits (3)	**1.2083**	20.2797	74.359	29	[10.90,12.20]	[100,140]	-0.20
Rule 19	Capital Adequacy Ratio (2)	Total Operating Income / Total Assets (2)	**1.2001**	22.3776	78.0488	32	[13.30,14.60]	[4.100,6.200]	0.12
Rule 20	Shareholders Equity / Total Assets (3)	Total Operating Income / Total Assets (2)	**1.1917**	21.6783	77.5	31	[9.600,10.90]	[4.100,6.200]	0.38

[Table.5] summarized the extracted rules. Rules are ranked according to their lift measure. The antecedent, consequence, support, confidence, lift values and number of transactions values are reported for each rule. In addition, the ranges of the predecessor and successor are also listed in [Table.5].

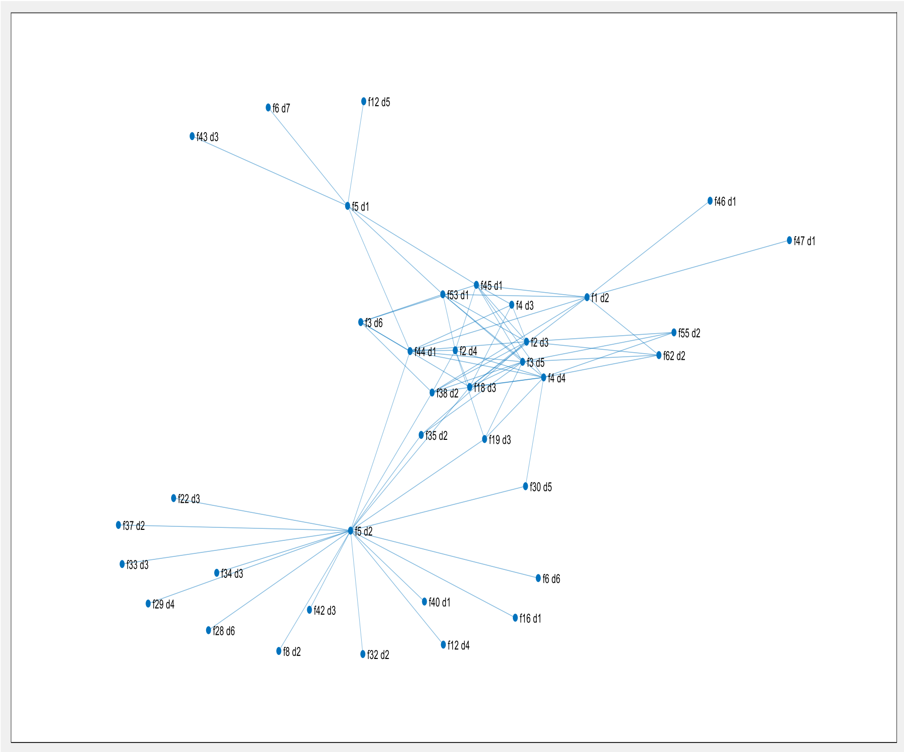

Figure 9: Network Plot of Rules

Network plot of rules is presented in figure. The [Figure.9] shows interconnections between 75 rules. Each interval of the rules is represented by a node. Connections between nodes are represented by links. The features that are frequently observed in rules are placed in the centre of the figure. This figure visually represents the relations between rules.

Performance Comparison with Single Discrete Values

Performance of the proposed method is compared with single models where each feature is discretized with fixed numbers 2, 3… 10. In other words, in this model where optimization is not used, each variable is categorized with fixed numbers. With each new data set, the apriori algorithm has been run once again, and the average confidence, average support, average leverage values and leverage values are greater than two. According to the table, the average support and average lift values in the optimal model are superior to other models. The average confidence value appears to be lower at optimal value compared to other models. Moreover, the number of rules with a lift value greater than two is highest in the optimal model [Table.6].

Table 6: Performance Comparison with Single Discrete Values

	Average Support	Average Confidence	Average Lift	Number of Rules ($L > 2$)
$n_i = 2$	0.7242	0.5140	1.0184	0
$n_i = 3$	0.6003	0.3424	1.0522	0
$n_i = 4$	0.5379	0.2632	1.0538	0
$n_i = 5$	0.5365	0.2585	1.0904	0
$n_i = 6$	0.5466	0.2519	1.1293	4
$n_i = 7$	0.5774	0.2507	1.1849	6
$n_i = 8$	0.5820	0.2418	1.1280	8
$n_i = 9$	0.5635	0.2372	1.1711	8
$n_i = 10$	0.6277	0.2478	1.5551	10
Optimal n^*	0.7440	0.2118	2.5222	14

.

5. Conclusion

The banking sector (especially commercial banks) is one of the major sectors of the country's economies. When we consider the economic structure as the financial sector and the real sector in two different ways, banks are the institutions that provide transitivity within these two sectors. It is sufficient to explain the importance of the financial sector and the structures that support production in the real sector. In this context, the situation of the Turkish banking sector was examined with the help of the apriori algorithm, one of the data mining techniques. The dataset collected from the Turkish banking sector is continuous. On the other hand, to implement the Apriori algorithm, the data set must be discretized. In this study, the complexity of determining group number in mapping continuous feature to a Boolean data type in applying association rules is addressed. A genetic algorithm-based approach is developed to solve the problem.

The proposed system optimizes the number of categories by applying a genetic algorithm optimization method. This optimization procedure tries to maximize the number of rules that have lift values greater than 2. The effect of the different fitness function design is the subject of other studies. The system we propose maps continuous variables into a number of categories that are optimized by genetic algorithm. It is in harmony with the quality of the rules produced by the system and the correlation coefficients calculated in the cases of variables before they are divided into categories. The performance of the proposed system was compared with the performance of models divided into fixed numbers of groups. The highest performance is achieved in the model optimized by the genetic algorithm. The best results in the optimal model in terms of lift value may be due to the consideration of lift value in the fitness function when applying the genetic algorithm. In future studies, besides the lift value, support and confidence values can be used in the fitness function.

The results provided reveal information about the general state of the banking sector. Using association rule mining to analyze the Turkish banking sector, according to the case database, it is concluded that there are discernible [Net on Balance-sheet Position / Total Shareholders' Equity] characteristics of Turkish banking sector, and a superposition relationship between [Net on Balance-sheet Position / Total Shareholders' Equity] and [Total Loans / Total Assets]. Through the relationship between mechanism and results, the optimal results for the Turkish banking sector are put forward as follows:

First under the highly strong association rules, in view of the [Net on Balance-sheet Position / Total Shareholders' Equity] in the Turkish banking sector, the strategy of total shareholders' equity should be adopted. The financial structure owned by commercial banks could become more robust with this ratio rising. Secondly, [Shareholders' Equity / (Deposits + Non-Deposit Funds)] ratio has a dominant influence in the Turkish banking system. As can be seen from this ratio, it can be said that the concept of "equity" is important in the sector..

References

F. Kaya, Bankacılık Giriş ve İlkeleri, Beta Yayınlları, İstanbul, 2012.

S. Halim, T. Octavia, C. Alianto, Designing facility layout of an amusement arcade using market basket analysis, in: Procedia Comput. Sci., 2019. doi:10.1016/j.procs.2019.11.165.

D. Solnet, Y. Boztug, S. Dolnicar, An untapped gold mine? Exploring the potential of market basket analysis to grow hotel revenue, Int. J. Hosp. Manag. (2016). doi:10.1016/j.ijhm.2016.04.013.

W.J. Saunders, S.R. Mousa, J. Codjoe, Market basket analysis of safety at active highway-railroad grade crossings, J. Safety Res. (2019). doi:10.1016/j.jsr.2019.09.002.

A. Pande, M. Abdel-Aty, Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool, Saf. Sci. (2009). doi:10.1016/j.ssci.2007.12.001.

J. Hong, R. Tamakloe, D. Park, Application of association rules mining algorithm for hazardous materials transportation crashes on expressway, Accid. Anal. Prev. (2020). doi:10.1016/j.aap.2020.105497.

C. Zhang, X. Xue, Y. Zhao, X. Zhang, T. Li, An improved association rule mining-based method for revealing operational problems of building heating, ventilation and air conditioning (HVAC) systems, Appl. Energy. (2019). doi:10.1016/j.apenergy.2019.113492.

A. Telikani, A.H. Gandomi, A. Shahbahrami, A survey of evolutionary computation for association rule mining, Inf. Sci. (Ny). (2020). doi:10.1016/j.ins.2020.02.073.

P. Pehlivan, Turk Bankacilik Sektoru Verimliliginin Rasyo Yontemi ile Incelenmesi: 2005:2014, Dayanisma Derg. (2015).

A. Karakcıoglu Altinay, Turk Bankacilik Sektorunde Kirilganlik, Finansal Krizlerin Kamu ve Ozel Sermayeli Bankalara Etkisinin Sermaye Yeterliligi ve Aktif Kalitesi Rasyosu ile Tespiti ve Karsilastirilmasi (1990-2015 Donemi), Finans Ekon. ve Sos. Araştırmalar Derg. 3 (2018) 544–563. doi:DOI: 10.29106/fesa.358675.

E. Demirel, A. Atakişi, S. Abacıoğlu, Bankacılık Faaliyet Oranlarının Panel Veri Analizi: Türkiye'deki Kamu, Özel Ve Yabancı Sermayeli Bankaların Durumu, Muhasebe ve Finans. Derg. (2013).

K. Alper, T. Capacioglu, Banking Sector's Liquidity Position and its Relation to Deposit Rates, 2016. https://www.tcmb.gov.tr/wps/wcm/connect/8ff2daf4-0bbb-405d-bcd0-1b7b8b5838e8/en1631eng.pdf?MOD=AJPERES&CACHEID=ROOTWORKSPACE-8ff2daf4-0bbb-405d-bcd0-1b7b8b5838e8-m3fw5t1.

R.J. Kuo, C.M. Chao, Y.T. Chiu, Application of particle swarm optimization to association rule mining, Appl. Soft Comput. J. (2011). doi:10.1016/j.asoc.2009.11.023.

M. Kaya, R. Alhajj, Genetic algorithm based framework for mining fuzzy association rules, Fuzzy Sets Syst. (2005). doi:10.1016/j.fss.2004.09.014.

A. Salleb-Aouissi, C. Vrain, C. Nortet, QuantMiner: A genetic algorithm for mining quantitative association rules, in: IJCAI Int. Jt. Conf. Artif. Intell., 2007.

A. Ghosh, B. Nath, Multi-objective rule mining using genetic algorithms, Inf. Sci. (Ny). (2004). doi:10.1016/j.ins.2003.03.021.

K. Wang, B. Liu, S. Tay, Interestingness-based interval merger for numeric association rules, in: Proc. 4th Int. Conf. Knowl. Discov. Data Mining, KDD, 1998.

R.J. Miller, Y. Yang, Association Rules over Interval Data, SIGMOD Rec. (ACM Spec. Interes. Gr. Manag. Data). (1997). doi:10.1145/253262.253361.

Y. Djenouri, M. Comuzzi, Combining Apriori heuristic and bio-inspired algorithms for solving the frequent itemsets mining problem, Inf. Sci. (Ny). (2017). doi:10.1016/j.ins.2017.08.043.

P. Kumar, A.K. Singh, Efficient Generation of Association Rules from Numeric Data Using Genetic Algorithm for Smart Cities, in: 2019. doi:10.1007/978-3-030-01560-2_14.

J.H. Holland, Adaptation in Natural and Artificial Systems, 2019. doi:10.7551/mitpress/1090.001.0001.

O. Kramer, Genetic Algorithm Essentials, 2017. doi:10.1007/978-3-319-52156-5.

S. Sivanandam, S.N. Deepa, Introduction to Genetic Algorithms, Springer Berlin, 2008. doi:10.1007/978-3-540-73190-0_2.

P. Giudici, S. Figini, Applied Data Mining for Business and Industry, 2009. doi:10.1002/9780470745830.

R. Nisbet, G. Miner, K. Yale, Handbook of statistical analysis and data mining applications, 2017. doi:10.1016/c2012-0-06451-4.

P. Bhatia, Data Mining and Data Warehousing principles and practical techniques, 2019.

G. Shmueli, P.C. Bruce, N.R. Patel, Data Mining for Business Analytics, 2016.

Q. Li, Y. Zhang, H. Kang, Y. Xin, C. Shi, Mining association rules between stroke risk factors based on the Apriori algorithm, in: Technol. Heal. Care, 2017. doi:10.3233/THC-171322.

D. Hand, H. Mannila, P. Smyth, Principles of Data Mining Cambridge, 2001. doi:10.1007/978-1-4471-4884-5.

R. Agrawal, R. Srikant, Fast Algorithms for Mining Association Rules in Large Databases, in: Proc. 20th Int. Conf. Very Large Data Bases, 1994.

R. Srikant, R. Agrawal, Mining Quantitative Association Rules in Large Relational Tables, SIGMOD Rec. (ACM Spec. Interes. Gr. Manag. Data). (1996). doi:10.1145/235968.233311.

B. Alatas, E. Akin, A. Karci, MODENAR: Multi-objective differential evolution algorithm for mining numeric association rules, Appl. Soft Comput. J. (2008). doi:10.1016/j.asoc.2007.05.003.

Y. Shi, B. Wu, N. Chen, A. Chen, J. Li, H. Li, Determination of effective management strategies for scenic area emergencies using association rule mining, Int. J. Disaster Risk Reduct. (2019). doi:10.1016/j.ijdrr.2019.101208.

Y. Zhou, C. Li, L. Ding, P. Sekula, P.E.D. Love, C. Zhou, Combining association rules mining with complex networks to monitor coupled risks, Reliab. Eng. Syst. Saf. (2019). doi:10.1016/j.ress.2019.02.013.

D.B. Calçada, S.O. Rezende, M.S. Teodoro, Analysis of green manure decomposition parameters in northeast Brazil using association rule networks, Comput. Electron. Agric. (2019). doi:10.1016/j. compag. 2019.02.013.