Journal of Industrial and Systems Engineering
Vol. 13, No. 4, pp. 39-61
Autumn (November) 2021

Identifying criteria which improve efficiency in an Iranian development bank using artificial neural networks

Negar Mohammadi[1] and Ali Bozorgi-Amiri[2*]

[1]*Department of Industrial Engineering, College of Engineering , Alborz Campus, University of Tehran, Tehran, Iran*
[2]*School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran,*

ngr.mohammadi@ut.ac.ir, alibozorgi@ut.ac.ir

Abstract

Banks, in general, have a direct impact on the macro-economy of all countries. Recognizing the criteria which have momentous influence on bank branches' efficiency is the main purpose of this research. An artificial neural network approach, one of the most applicable data mining techniques, is adopted to identify the criteria that influence the branches' efficiency the most (according to the result of efficiency evaluation base on MCDM). Then, the optimal group of input criteria is determined in order to achieve the most efficient performance. Branches that enjoy more appropriate inputs would have better conditions to increase their efficiency, possess more acceptable position and gain more adequate results. In this paper, utilizing data mining science, we have endeavored to suggest a suitable method in recognizing the most significant inputs with positive impact on enhancing efficiency of branches by the incorporation of relatively neglected indicators which fit the particular conditions of Iranian banks. The strength of this article compared to other related researches is that it provides a mechanism according to which senior managers in the banking sector will be able to identify the most important indicators and implement the best conditions to achieve the highest level of efficiency in the collection.

Keywords: Bank, efficiency, criteria, data mining, artificial neural network

1-Introduction

Efficiency has been evaluated through a wide variety of methods in banking system (Idris & Ahmad 2020). Financial markets are systems which gather economic actors around which require budgets and have considerable resources (Dincer et al. 2011). Given the fiercely competitive nature of the banking industry, methods of performance evaluation and improvement have recently been drawing even more attention. In simple terms, members of the banking industry are in constant competition to provide better services (Shokrollahpour et al. 2016). Evaluating the performance of financial institutes is regarded as a crucial issue for all governments, many of which face one or more banking crises at any given time (Jemric & Vujcic 2002).

*Corresponding author
ISSN: 1735-8272,

This performance evaluation is also important to investors and stakeholders as it helps determine banks' ability to compete effectively within the banking system and broader financial networks and is instrumental to the process of development (Ho & Wu 2009). The banking sector in Iran is the chief provider of cash resources for enterprises and industries and regarded as the main bridge connecting the supply and demand of cash resources; therefore, it has gained widespread economic and political influence over time and maintains a decisive role in preserving the vitality and functionality of the country's financial community.

In real world, many decisions are made at any given moment, usually with uncertainty (Akbari & Osgooei 2020). However, in banking industry all decisions must be made based on scientific research plus previous experiences. Policymakers in every country always seek to prevent banking crises from happening and although they do have some control over their respective country's domestic situation, they face much graver challenges and difficulties when it comes to international crises (Dungey & Gajurel 2015). Factors that affect the performance of banks are highly diverse, to the point that in some cases, the staff's gender may affect the bank's performance level (Owen & Temesvary 2018). Today, in the era of rapid progress and technological breakthroughs, data mining is one of the disciplines drawing keen interest among researchers in various fields and, in particular, among engineering groups. The immeasurable amount of scientific data in today's advanced world can only be accessed and analyzed through modern tools whose efficacy has been consistently proven. Data mining is one such tool which has greatly helped scientists and researchers since its emergence. Data mining techniques have been expanded to such a degree that they can now readily be adapted and added to any software application and make the best of the collected data. Figure (1) illustrates the four levels of information maturity using data mining:

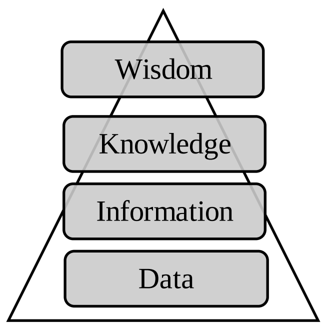

Fig 1. Stages of information maturity using data mining

The growing interdisciplinary field of knowledge discovery and data mining (KDD) integrates various scientific fields and approaches such as statistics, databases, machine learning, etc. and extracts valuable information and knowledge from the data related to each one for use by researchers and practitioners.

In this paper, we aim to achieve two main goals: a) ranking the branches with similar performance levels, and then b) using an artificial neural network (ANN) technique along with the aforementioned ranking to identify the most significant factors affecting the performance and efficiency of the studied bank's branches. In the end, the optimal set of inputs (i.e. the best combination of factors) that leads to the highest output value (performance level) will be determined. The rest of the study is structured into the following sections: literature review, research methodology, case study, and conclusion. The feature that distinguishes the present research from other works is the use of ANN to determine the optimal way of determining input values in order to obtain the highest performance level possible as output. It should be noted that, this paper has used valid data in which all 206 studied branches are evaluated and ranked based on efficiency assessment via MCDM approaches. In fact, first of all required data have been acquired from the results of performance measurement studies which demonstrate the percentage of efficiency of all branches one by one plus the position of each branch compared to other branches and then the authors have used that information in this paper in order to clarify which criteria or conditions could assist branches of the bank to increase their effectiveness. The efficient criteria which we are trying to identify are selected from the same criteria with which the performance evaluation has been performed. The purpose of this paper is to show that in spite of the fact that each organization has specific criteria and scale for measuring the units under its control in accordance with its goals, standards and priorities, it must be considered that some criteria are able to change the results dramatically, so special attention should be paid to them. They are those key factors which are able to enhance the overall efficiency of each set.

Given that banks are one of the most competitive institutions with very high turnover throughout the world, they require to reevaluate their competitive benefits in light of profound changes driven by advances in information technology (Jaksic & Marinic 2019). Therefore, banks are faced with a very large amount of numerical data that must be carefully analyzed in order to assess the bank's position in a competitive market. Correct analysis of this volume of numerical data is very important and needs a tool that can be evaluated with great accuracy. While the numerical equivalent of the data amount is created at categorical rate, the conservation of the information it carries has been a serious matter in terms of gaining efficient results and has been studied a lot (Bacaksız & Esgin 2019). A neuron is a mathematical function that receives inputs and relapses outputs. These outputs use as inputs for the forthcoming layer, and so on until we get to the final, output layer, which is the actual value we return. For each set of inputs, the neural network's purpose is to make each of its outputs as adjacent as possible to the actual awaited values. A Neural Network is a Machine Learning model that, given specified input and output vectors, would make effort to "fit" the outputs to the inputs (Adamu 2019). Data of the banking industry are mostly in numerical form; thus, due to what explained previously and owing to the high power of artificial neural networks in the analysis of numerical data, which is also the same type of banking data, this approach has an acceptable degree of reliability compared to its counterparts in the analysis of banking data. In fact, in this research which is based on performance evaluation, numerical data with mathematical function play critical role; thereby, requiring a technique that its application is in accordance with this kind of inputs. Overall, according to the researches and published articles in this field, some of which were mentioned, the artificial neural network approach is very compatible with these types of studies and can be used as a practical and lucrative tool.

Development Banks are banks that are established with the aim of financing a specific sector of industry, trade, production or other units. Thus, such banks have a much higher value and importance than their counterparts, because not only do they perform the main tasks of a bank, but also the growth and development of one or more other organizations depends on their performance. It is necessary to explain that on a global scale, Development Bank or Regional Financial Institution designed to provide medium- and long-term capital for productive investment, often accompanied by technical assistance, in poor countries.

2-Review of literature

Given the widespread use of data mining techniques in financial institutions, especially in banks, this section is allocated to a brief review of some of the major research works published thus far on the subject. In this part we introduce some studies related to the role of data mining in banking industry (in terms of performance evaluation or some other goals which are relatively close to efficiency measurement).

Santin et al. (2004) used an ANN approach to develop a simulated nonlinear production function and compared its performance with traditional techniques such as SFA (stochastic frontier analysis) and DEA (data envelopment analysis) as well as in various scenarios. The findings suggested that ANN-based approaches could be promising alternatives to conventional techniques of evaluating production performance indicators and measuring efficiency in nonlinear settings. Wu et al. (2006) utilized an integrated ANN and DEA approach to assess the efficiency of the branches of a large bank in Canada. The researchers proposed a number of solutions for improving the studied branches' performance level based on the results. M.Mostafa (2009) considered the efficiency of the most popular and best Arab banks via utilizing two quantitative methodologies: data envelopment analysis and neural networks. On that study he depicted those tools have an excellent for the classification of banks' relative efficiency. Bakar and Tahir (2009) by studying the 7-year performance of thirteen banks, using seven variables including liquidity, credit risk, cost to income ratio, size, concentration ratio, inflation and GDP and via applying several linear regression and neural network proved that the neural network technique powerful method in predicting bank performance. Azadeh et al. (2012) combined two DEA and ANN approaches to measure the efficiency of DMUs (decision making units) in an intelligent system and proposed an integrated approach to optimize the attributes of a large bank's personnel. According to the significance of banking performance assessment, Wang et al. (2014) evaluated the performance of 16 Chinese commercial banks during an extensive study from 2003 to 2011 via two-stage DEA method.

The bulk of the research conducted on data mining in banks revolves around bank customers, e-banking services, and bank fraud detection. Raju et al. (2014) used data mining-based tools and techniques and introduced a model to improve banks' CRM (customer relationship management) by relying on decision trees and SVMs (support vector machines), both of which regarded as data classification and supervised learning tools in data mining. The researchers then applied the model to improve banking services such as credit cards and marketing by grouping the customers based on certain attributes as well as increasing profit and productivity. Mansingh et al. (2015) worked on profiling the e-banking users and the benefits of adopting KDD. The researchers used a set of different models instead of a single one and presented their study on data mining with a clustering-based approach. Wanke et al. (2016) used an integrated ANN and two-stage TOPSIS (technique for order of preference by similarity to ideal solution) approach, including multi-criteria decision-making methods, to forecast the efficiency of Islamic banks in Malaysia. The findings showed that cost structure may be the most decisive factor in the productivity of the studied banks. In that year, Wanke et al. measured the efficiency of 114 Islamic banks from 24 countries using MCDM methods a neural networks and then demonstrated that variables related to both territory sources and cost structure have a noted influence on efficiency.

Farooqi and Iqbal (2017) used data mining and CRM tools to improve risk management practices and fraud detection in the banking industry. The researchers developed an algorithm to classify the transactions made through ATMs and put transactions which appeared somewhat unusual in a separate category. Based on the behavior pattern and common techniques adopted by criminals to discover the actual frauds. Finally, a number of strategic programs were devised to minimize the risk of recurring incidents in the future. Kanmani and Jayapradha (2017) used a deep learning ANN algorithm along with the customers' information in a bank's database to predict default customers. The study indicated that customers are often more loyal to banks with which they had worked previously. Ilie et al. (2017) published papers in which the performance of a Romanian Banking System in achieving customer satisfaction was evaluated. To this end, the paper concentrates on making and training an Artificial Neural Network (ANN) in order to simulate bank customer contentment and detect the indicators that are most substantial for training the ANN. Patil and Dharwadkar (2017) worked on the performance appraisal of the banking system in two areas: fraud detection and customer satisfaction. To achieve this goal, since ANN can process manifold inputs effectively and also it could manage large, complex data easily, supervised artificial neural network algorithm is implemented for classification purpose.

Tavana et al. (2018) analyzed the liquidity risk in a banking system. The researchers utilized a two-phase ANN and developed a Bayesian network model to determine liquidity-related indicators and calculated the liquidity risk involved in the banking system under study. In that year, Kassani et al. presented a hybrid pattern of DEA and methods of data mining to appraise the proficiency of bank branches. Kataria and Nafis (2019) adopted an integrated decision-making tree and a multi-layer perceptron neural network to search for and detect fraud in an e-banking system. The researchers emphasize the significance of fraud detection methods and compare the hidden Markov model (HMM), deep learning, and ANN which are all used to detect fraud in online transactions. In that year, Anouze et al. mixed DEA and data mining tools in order to estimate the effectiveness of banks in MENA countries. Daliri (2020) proposed a model based on ANN and the Harmony search algorithm (HSA) to search for and detect hidden patterns in the transactions of ordinary and fraudulent customers. The model enables organizations and financial institutions to detect suspicious behavior and thereby prevent fraudulent transactions from occurring. Appiahene et al. (2020) illustrated a practical way for forecasting bank operational efficiency by enjoying comparative study of decision tree, random forest, neural networks, and also DEA in order to measure and predict the performance on 444 bank branches in Ghana. In that year, Niknafs et al. utilized dynamic DEA and ANN to evaluate the effectiveness of bank branches and then acquire efficiency trend during the time. Roy (2020) published a paper in which proposes hybridization of DEA and ANN techniques for executive performance evaluation and prediction for Indian banks enjoying the five-year (2015 to 2019) dataset. It clarifies the usage of ANN in banking systems over the years, especially the current extensive studies which is a proof of the efficiency of this method. This research executes prediction duty for gained efficiency marks. Finding of will be advantageous for decision-makers, experts and managers of banking industry for forecasting future operational performance of banks until they are able to make required changes for its development. Ristyawan (2021) published a research in purpose of to present an integrated intelligence algorithm for estimating the strategy of a bank in Indonesia. The algorithm has two basic modules, called ANN and AHP. This framework can be implemented to assist bankers to resolve on bank operations.

3-Research methodology

In this section, the structure of the present paper is detailed. Initially, the methods and approaches selected for the research are introduced, and then the manner in which they were used during the research process are explained.

There are two broad groups of data mining methods: predictive and descriptive. In predictive data mining, a number of variables or fields in the database are utilized in order to identify and predict future (or unknown) behaviors, while descriptive data mining refers to the practice of finding data patterns which are interpretable for humans. Figure 2 illustrates the different methods classified as predictive and descriptive in data mining.

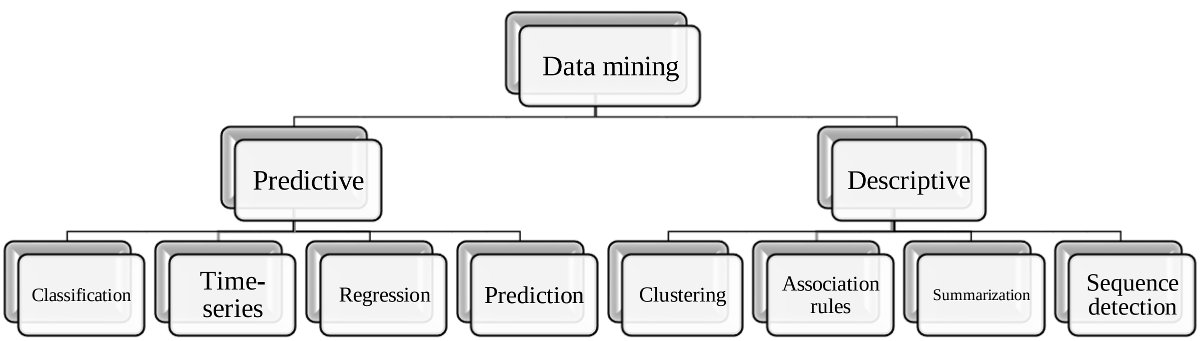

Fig 2. Data mining methods

KDD has found widespread use in various fields, some of which include:

- Fraud detection
- Network management
- Control and scheduling
- Employment market analysis
- Test result analysis (technical, medical, etc.)
- Financial market data analysis
- Analysis and inventory in aerospace industry research

The basis of data aggregation and analysis in the present study is the CRISP-DM (cross-industry standard process for data mining) which is considered a comprehensive model for data mining projects (Wirth & Hipp 2000). This methodology consists of six major phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. Figure 3 is a schematic illustration of the phases of CRISP-DM.

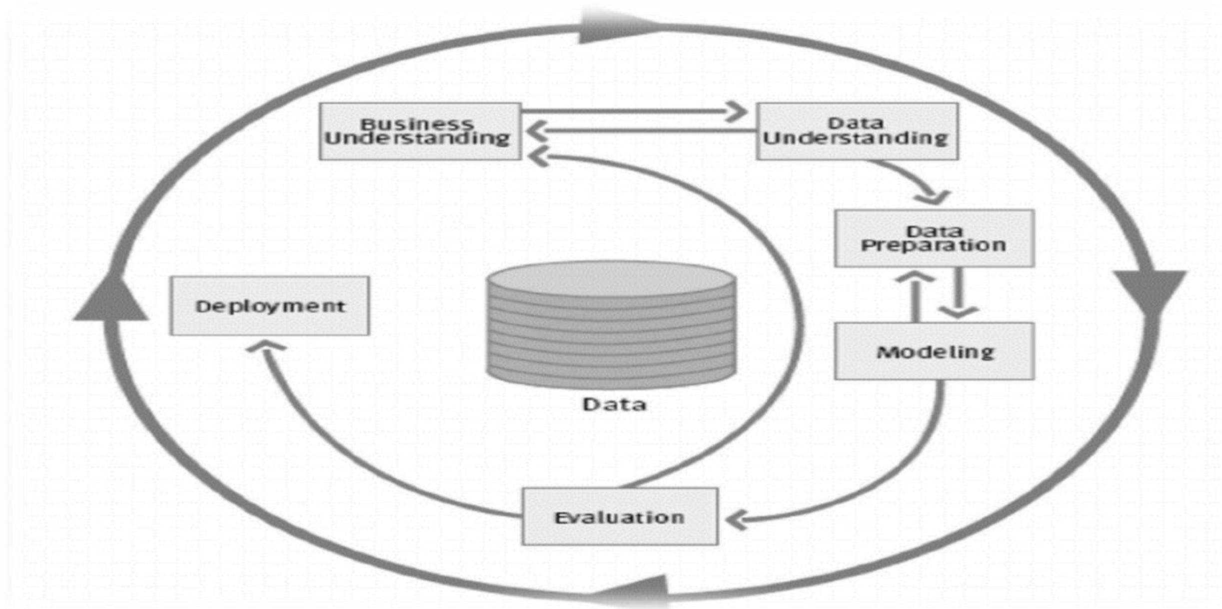

Fig 3. Phases of CRISP-DM

Steps of CRISP-DM:

- Business Understanding
- Data Understanding
- Data Preparation
- Modeling
- Evaluation
- Deployment

From among a number of credible techniques and methods in the field of data mining, the ANN, one of the most prevalent of its kind, was selected as the method of this study (Beale et al. 1996) (Hassoun 1995). ANNs are computing systems inspired by bionetworks which learn about various practices by inspecting numerous examples (training data). ANNs have a number of advantages compared to other methods with regards to operations such as prediction and classification and are usually the preferred choice in executive projects. Although ANNs cannot be seen as the ultimate answer to all computational problems faced by humanity, they are perhaps the best choice in dealing with complex data. In brief, ANNs are composed of the following components:

- A set of nodes
- Inputs and outputs for each node
- Performing computations based on specific functions
- Weighted links connecting the nodes
- Determining the links based on network architecture
- Displaying a highly complex function of the weighted links as the output.

$$I_j = \sum_{i=1}^{n} w_{ij} x_i \qquad Sum\ Of\ Weighted\ Inpyts \tag{1}$$

$$y_j = \oint (I_j) \qquad Function$$

A neural network is a collection of connected units which basically model the vast network of neurons in the human brain. It shapes up when the outputs of some neurons are connected to the inputs of other ones and, together, they form a directed and weighted graph. In the following sections, the measures taken by the authors to obtain the final research results will be explicated step-by-step.

3-1-Data classification and neural network formation

In data mining, classification technique (which comes in different types itself) is used to analyze data and extract patterns in order to describe significant datasets as well as to understand and predict the future behavior of data. Put simply, data classification refers to the practice of assigning labels to datasets which have not yet been classified. Next, the data are assigned – according to their attributes – to different classes whose names have already been determined. Data classification is a two-stage process:

a) First stage - building the model: a number of predefined classes are described based on the model's training datasets (which help with learning). In this process, the target is to build a model using the existing samples based on which the unlabeled data can then be placed in relevant classes.

b) Second stage - using the model: This stage consists of two steps: 1) testing the accuracy of the model built in the first stage, and 2) classifying the data based on the model. In order to measure the model's prediction accuracy, a set of sample data are randomly selected from the available data, on which the model is executed. The main target of this stage is to improve the model's accuracy i.e. enabling the model to assign each datum to a suitable class.

Considering the description provided above, initially about 30% to 40% of the customer record data received from the studied bank's branches are used to train the model. After monitoring the model's performance and approving its accuracy, the rest of the customer record data are added to the model for classification.

As mentioned previously, there are several predictive classification methods. In this study, the ANN, a popular and applicable technique, was selected due to its efficiency and accuracy. The ANN method provides good accessibility, has appeared to be more accurate than similar methods, is flexible and gives the user numerous options, and its fitness can be completely tested. The artificial neural networks are widely used to prop the human decision susceptibilities, eschewing contradiction in practice and errors based on shortage of experiment (Ahamed & Akhtar 2016).

In terms of the practical use of handling data, perhaps the most important part of the data mining operation is selecting the data's properties. In theoretical discussions, the problem's properties are often handed over to users. However, when it comes to practical application, it is the specialists' (in the case of this study, the bank's top executives and experts) job to extract the required properties (that is, influential and decisive properties) from among countless available properties. Data properties (or aspects) form the basis of the majority of data mining and machine learning operations.

In this study, after evaluating the performance of the bank's branches and assigning normalized performance scores to each branch, a network is formed which is able to produce an accurate prediction of the correct placement of each branch in a suitable class. Neural networks are used to model the connections between inputs and outputs with the goal of finding an utilizable pattern. The main issue here is to develop a model that produces the lowest error rate in the prediction operation. The multi-layer perceptron neural network's architecture is represented by ANN(A,N) with the following variables:

N A set of nodes

NI Input layer

NH Hidden layer

NO Output layer

A A set of arcs

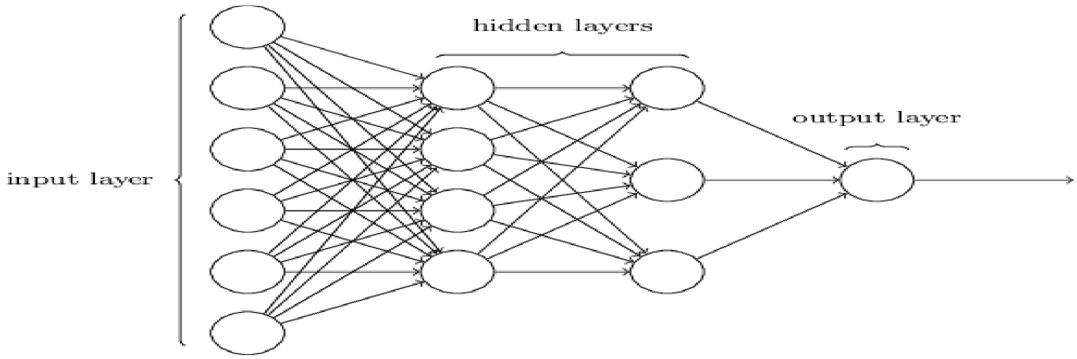

Fig 4. The multi-layer perceptron neural network's architecture

It must be noted that in this study, since the sizes of the investigated levels (i.e. number of branches in each class) are not equal, we deal with imbalanced data, which occurs when the number of samples in one or more classes in the dataset is lower than other classes. This leads to what is known as Skewness in the distribution of datasets and creates problems for most machine learning algorithms.

In this research, classifying the data into groups with exactly equal number of members is logically wrong because the branches are classified based on efficiency; therefore, a well-performing branch cannot be placed in the same group as a low-performing one simply to ensure numerical equality. As a result, the aforementioned imbalance is also evident in this study. There are various approaches to solving this issue, one of the most applicable of which is known as resampling which, itself, comes in different types (e.g. oversampling). It is worth noting that only executable data are balanced and not all the data; this is due to the fact that if every piece of data was balanced, the model would be unrealistic, only superficially accurate, and thereby unreliable.

Feature selection is another widely-discussed subject in data mining, machine learning, and pattern identification. It is of utmost importance in some operations (e.g. data classification) since there are often numerous features to deal with, most of which being either useless or not adequately informative. Failing to eliminate redundant features does not create much difficulty in terms of the available information, yet it needlessly increases the user's already heavy computational load and stores plenty of worthless data along with useful ones.

Feature selection methods are categorized based on the type of search. In some methods, the whole space may be searched, while in others e.g. exploratory or stochastic search, a smaller space is searched for the sake of maximum efficiency and speed. In order to carry out an accurate categorization of feature selection methods, the selection process of all methods is broken down into the following components:

- Generating function: Finds candidate subsets for the selected feature selection method.
- Evaluating function: Evaluates the subset based on the feature selection method and returns a number as the method's accuracy score.
- Termination condition: Determines the point at which sufficient computation has been done and the algorithm should be terminated.
- Validation function: Determines the validity score of the selected subset.

Methods of determining the generating and evaluating functions are listed in the table 1:

Table 1. Methods of determining generating and evaluating functions

Generating function	Evaluating function
Full search	Based on distance
Stochastic search	Based on information
Exploratory search	Based on dependency
-	Based on adaptivity
-	Based on classification error

In this study, using the two statistical tests Chi-square and ANOVA (analysis of variance), a code was written in the programming software Python to select influential features. The neural network is developed based on multi-layer perceptron networks with the following features:

- Input layer: number of indicators
- Hidden layer: 50 layers (in accordance with the number of training stages)
- Output layer: number of levels (A, B, C, D)
- Learning rate: 0.005 (speed of weight modification)
- Max iteration: 1000

Correct determination of the number of hidden layers has a significant impact on the model's efficiency. The most important criterion for determining the efficiency of an evaluation algorithm, on the other hand, is the 'accuracy score' it generates, which also indicates what portion of all evaluated records have been classified correctly.

In addition, we further evaluate the developed model's accuracy by plotting a confusion matrix, an n × n matrix where n represents the number of classes (A, B, C, D). In order to better understand the confusion matrix, let us examine the following example:

		Predicted	
		Positive	Negative
Actual	Positive	True Positives (TP)	False Negatives (FN)
	Negative	False Positives (FP)	True Negatives (TN)

Fig 5. Definition of confusion matrix

- True negative (TN): represents the number of records with a negative number of classes and the classification algorithm has correctly identified the number of classes as negative.
- True positive (TP): represents the number of records with a positive number of classes and the classification algorithm has correctly identified the number of classes as positive.
- False positive (FP): represents the number of records with a negative number of classes, but the classification algorithm has incorrectly identified the number of classes as positive.
- False negative (FN): represents the number of records with a positive number of classes, but the classification algorithm has incorrectly identified the number of classes as negative.

In confusion matrices, the rows represent a real class or a set of data being tested, while the columns represent the tested data's class (Parker 2001).

3-2- Developing information systems

Information systems (a.k.a. data charts, systems of attribute-value, condition-action charts, etc.) are means of representing knowledge in a systematic way. An information system refers to a dataset or table whose rows each represent an event or simply an object, while its columns contain an attribute such as a variable, observation, or characteristic which is measurable for each object or which is computed by information system experts. A decision table could be looked upon as the decision-maker's view on a subject in order that the common decision rule is considered as group knowledge or a general agreement (Skowron & Rauszer 1992) (Inuiguchi & Miyajima 2007). The theory of rough sets, introduced by Pawlak (Pawlak 1982) has found widespread use in today's information systems and rule-based decision-making (Beynon & Peel 2001). An information system can be displayed as follows:

U A non-empty finite set of objects

B A non-empty finite set of attributes

$$A = (U, B) \tag{2}$$

$$b : U \to Vb \quad , \quad b \in B$$

In many information systems, in addition to the inputs' attributes (conditional attributes), the outputs' attributes (decision attributes) are presented as well, in which case the information system turns into a decision system which is displayed as follows:

$$A = (U, B \cup \{d\}) \tag{3}$$

$$d \notin B \qquad \text{decision attribute}$$

Non-discrimination: In many information systems, there are objects with identical attributes. According to the definition:

$$A = (U, B) \tag{4}$$

$$C \subseteq B$$

$$IND_A(C) = \{(x, x') \in U^2 \,|\forall a \in C \;\; a(x) = a(x')$$

Upper and lower bounds of decision attributes: The upper bound set includes objects which are probably in a specific decision attribute class, while the lower bound set includes objects which are definitely in a specific decision attribute class.

$$A = (U, B \cup \{d\}) \tag{5}$$

$$d\,X = \{x \,|[x]_d \cap X \neq \emptyset\} \qquad \textit{Upper Bound}$$

$$d\,X = \{x \,|[x]_d \in X\} \qquad \textit{Lower Bound}$$

Adaptability in information systems: One of the considerable properties of information systems is their adaptability. The adaptability of an information system is computed by calculating the system's positive region or consistency:

$$\text{Information system's adaptable region } = \frac{\text{Union of lower bound sets}}{\text{Value of system's decision attribute}} \tag{6}$$

The concept of information systems: One information system covers all kinds of data about the system. There are two ways in order to lessen data which is unnecessarily large:

- Diminution the number of conditional attributes
- Diminution the number of possible amounts for each conditional attribute

System indicators which do not have a significant impact on the classification of objects based on the values of decision attributes are called redundant indicators. In order to identify effective attributes, we should first discern information systems' redundant indicators. If redundant indicators are eliminated from the system's existing conditional indicators, what is left in the end will be the effective indicators. The total number of effective conditional attributes in an information system with “m attributes” is calculated as follows:

$$\binom{m}{[\frac{m}{2}]}$$

Since the size of an information system increases exponentially, the extent to which the details matter should be specified when identifying the cut-off points. The appropriate and accurate classification of a problem's attributes is crucial and rather complex; thus, heuristics and Meta heuristics have often been proposed as the most suitable solutions. The ANN forms based on the information system and its key purpose is to predict the impact of various factors (with regards to inputs) on the efficiency of each branch. The model should be designed in a way that the weights can be found with minimum margin for error. In this model, learning set is called E and the calculated error is, in fact, the difference between the predicted output p(x,w) and the real value of the output for all elements is x in set E. The network's architecture is, again, as follows:

ANN(A,N)

N A set of nodes

NI Input layer

NH Hidden layer

NO Output layer

A A set of arcs

We assume there are r variables in the network, based on which the decision attributes are predicted and in which there are j hidden layers. Hence, we have:

NI = {1, 2, … , r}

NH = {r+1, r+2, … , r+j}

NO = {s}

Every node i receives a signal from the value of xi in the input layer which is then sent to the hidden layers through the arcs. Then, every node r+j in the hidden layer, receives the input signal:

$$Input(r+j) = w_{r+j} + \sum_{i=1}^{n} x_i w_{i,r+j} \tag{7}$$

Where we have:

W_{i+j} Initial value of node r+j

$W_{i,r+j}$ value of the arc that comes from node i in the input layer to node r+j in the hidden layer.

Every hidden node r+j sends the value of output signal r+j through arc (r+j , s) and, finally, node s in the output layer receives the sum of weights coming from the hidden layer, based on which we will have:

$$NN(x,w) = w_s + \sum_{j=1}^{m} Output(r+j)w_{r+j,s} \tag{8}$$

The main target in this data processing method is to minimize error in set E. Since having a large number of inputs is not conducive to ANN modeling, the best way to achieve optimal results would be to integrate ANN with the rough set theory (Azadeh et al. 2011). The highly applicable theory of rough sets is useful in such practices as knowledge extraction from databases, data mining, fault detection, machine learning, knowledge acquisition, expert systems, and decision support systems (DSS). Developing a DSS can help overcome complications and contradictions and lead to enhanced efficiency (Azadeh et al 2012).

The purpose of this research is to develop a model that accurately predicts the optimal outputs and, to this end, it is essential that the model should be tested properly. Considering the involvement of imbalanced data and the criterion used to classify the data, the best way of assessing the first neural network (developed to measure the accuracy of data classification based on efficiency) may be to perform the algorithm numerous times, test the output, and calculate the mean output value.

Cross-validation approaches have been widely studied due to their ability to guide variable-selection algorithms (Baumann 2003). On this basis, the best validation approach for the second neural network is the k-fold cross-validation. In this approach, the data with n decision-making units (DMUs) are divided into k parts; one part of which used as test data and k-1 parts as the training data. In most cases, 10% of the data are used as test data, hence we will have:

K = [n/10]

Thereby, for each output of the information system (i.e. each set of conditional attributes extracted), i = 1, … , m. The values of k neural networks composed of j neurons are generated in the hidden layer. Next the mean error value when forming the network k times is registered in the hidden layer as the ith conditional network's error in the neural network with j neurons. The number of neurons in the hidden layer is determined considering the targeted error margin in the problem and can be stated as j = 1, … , q. In the end, k × j neural networks are built for each set. In order to assign an efficiency score to this network in the final evaluation, we consider an average of k networks per each number of neurons in the hidden layer as the efficiency score of a neural network composed of j neurons. The variables defined for the problem's error calculations are as follows: MAPEijk: Mean error when executing the validation k times for the ith conditional set with a neural network composed of j neurons in the hidden layer in kth execution

ERRij: Mean error when executing the validation k times for the ith conditional set with a neural network composed of j neurons in the hidden layer ERRij = Average(MAPijk)

AEERi: Mean error of the ith conditional set AEERi = Average(ERRij)

VarERRi: Variance of the ith conditional set VarERRi = Variance(ERRij)

MaxERRi: Maximum error of the ith conditional set MaxERRi = Max(ERRij)

MinERRi: Minimum error of the ith conditional set MinERRi = Min(ERRij)

4-Case study

As it was explained in Introduction, the basis of the data and information used in this article is in accordance with the results of the performance appraisal, and we do not intend to re-evaluate the performance in this article because it has already been done accurately via using effective MCDM tools. In order to describe the process and methods for evaluating the bank's performance in mentioned study the steps was in this format: First of all, the indicators which have direct impact on the performance level of the branches of the studied bank were identified through, field studies as well as holding numerous meetings with chief managers and policy-makers (brainstorming). The next step involved defining the bank's priorities, which was achieved by weighting the resulted criteria using an analytic hierarchical process (AHP) model and a pairwise comparison matrix (PCM). Finally, the three multi-criteria decision-making approaches including data envelopment analysis (DEA), additive ratio assessment (ARAS), and weighted aggregated sum product assessment (WASPAS) were employed to compute the efficiency level of each branch, based on which the branches were ranked. Rather, the purpose of this article is to identify the criteria and conditions that have led to the success of some branches in the performance measurement test. It is imperative for banking experts and market analysts to have sufficient information about the state of a financial institution as it helps identify current or potential problems in order to find timely solutions (Lin 2018). Numerous numbers were also tested to determine the weight modification rate and maximum iteration value; however, the best scenario in order to have the highest possible accuracy and efficiency levels is to have a learning (i.e. weight modification) rate of 0.005 and a maximum iteration value of 1000. The reason behind these figures is that, if the maximum iteration value is lower than 1000, model will have poor accuracy and if it is higher than 1000, the model will generate an extremely accurate but unrealistic output known and is said to "over fit" the data. Table 2 details the results of evaluating the model which contains a different value for each hidden layer determined in the process:

Table 2. Results of model's evaluation

70	65	60	55	50	45	40	35	30	
54.74%	55.23%	64.29%	70.21%	68.32%	62.88%	58.67%	54.32%	55.39%	**1**
52.62%	55.49%	63.56%	67.53%	69.73%	58.32%	62.54%	51.24%	52.67%	**2**
54.29%	53.46%	65.23%	64.23%	73.21%	65.21%	60.69%	53.22%	48.99%	**3**
52.83%	53.32%	61.39%	64.78%	68.78%	64.97%	56.34%	57.43%	51.23%	**4**
50.28%	54.73%	59.48%	63.78%	74.59%	59.89%	59.45%	51.87%	47.76%	**5**
49.39%	52.16%	59.64%	68.28%	76.54%	67.12%	61.03%	56.29%	50.84%	**6**
51.65%	52.29%	60.21%	66.39%	69.91%	66.99%	54.89%	53.29%	52.24%	**7**
48.86%	49.56%	63.83%	69.73%	67.84%	67.23%	59.24%	49.98%	52.03%	**8**
49.64%	48.71%	60.34%	66.32%	73.65%	65.29%	59.09%	49.43%	49.58%	**9**
52.22%	53.63%	58.64%	66.53%	74.21%	58.13%	55.66%	50.63%	50.08%	**10**
53.16%	55.79%	57.43%	70.44%	73.43%	64.23%	58.34%	54.24%	47.87%	**11**
53.63%	55.06%	60.71%	63.23%	70.86%	62.48%	61.42%	55.87%	49.62%	**12**
50.27%	53.17%	60.59%	63.44%	69.74%	61.53%	60.25%	53.43%	53.15%	**13**
51.48%	54.24%	62.60%	66.54%	69.56%	61.07%	59.37%	52.68%	51.08%	**14**
49.53%	51.70%	59.63%	63.86%	73.17%	59.24%	56.48%	53.47%	54.30%	**15**
48.97%	49.98%	59.74%	65.39%	72.91%	59.43%	61.08%	50.29%	51.18%	**16**
51.07%	53.84%	62.09%	63.49%	68.83%	63.76%	58.69%	49.66%	48.99%	**17**
51.39%	53.26%	60.06%	64.51%	69.43%	64.53%	55.26%	51.64%	49.67%	**18**
53.50%	54.68%	61.23%	62.81%	67.57%	62.65%	57.41%	56.36%	53.19%	**19**
50.66%	54.29%	58.88%	64.62%	67.80%	60.39%	55.44%	56.13%	50.90%	**20**
54.74%	55.79%	65.23%	70.44%	76.54%	67.23%	62.54%	57.43%	55.39%	**Max**
48.86%	48.71%	57.43%	62.81%	67.57%	58.13%	54.89%	49.43%	47.76%	**Min**
51.51%	53.23%	60.98%	65.81%	71.00%	62.77%	58.57%	53.07%	51.04%	**Mean**

As can be seen, 50 hidden layers appears to be the optimal number in designing a multi-layer perceptron network due to its ability to classify the execution data accurately. When computing the model's accuracy in different layers, the confusion matrix also seems to have a value among 30 to 70 per each number of hidden layers. However, considering the large amount of information and since there is no need to insert all the confusion matrices in the present text, it was decided to only include the results of 50 hidden layers as the most significant ones, as may be seen below.

$$\begin{pmatrix}2&0&0&1\\0&6&2&4\\1&1&6&4\\1&3&4&22\end{pmatrix}\begin{pmatrix}2&0&0&0\\0&6&3&3\\0&1&7&5\\1&2&3&23\end{pmatrix}\begin{pmatrix}3&0&0&1\\0&6&1&2\\0&1&10&4\\0&2&2&25\end{pmatrix}\begin{pmatrix}2&0&1&0\\0&6&2&3\\0&2&6&4\\1&2&5&23\end{pmatrix}$$

$$\begin{pmatrix}4&0&0&0\\0&7&1&3\\0&1&11&3\\0&1&3&26\end{pmatrix}\begin{pmatrix}4&0&0&0\\0&8&0&2\\0&1&12&2\\0&1&1&28\end{pmatrix}\begin{pmatrix}2&0&0&0\\0&5&2&3\\0&2&8&5\\1&2&5&23\end{pmatrix}\begin{pmatrix}2&0&0&1\\0&5&2&3\\1&2&6&5\\1&3&6&21\end{pmatrix}$$

$$\begin{pmatrix}3&0&0&1\\0&7&1&3\\0&1&10&3\\0&0&3&26\end{pmatrix}\begin{pmatrix}4&0&0&1\\0&7&1&3\\0&0&11&4\\0&1&3&26\end{pmatrix}\begin{pmatrix}3&0&0&1\\0&6&1&3\\0&2&10&5\\0&0&3&25\end{pmatrix}\begin{pmatrix}3&0&0&0\\0&6&2&4\\0&2&8&5\\1&2&3&23\end{pmatrix}$$

$$\begin{pmatrix}2&0&0&0\\0&6&2&3\\0&2&7&4\\1&2&4&23\end{pmatrix}\begin{pmatrix}3&0&1&0\\0&5&2&3\\0&1&7&5\\1&2&2&23\end{pmatrix}\begin{pmatrix}3&0&0&0\\0&5&1&3\\0&1&9&5\\1&2&3&24\end{pmatrix}\begin{pmatrix}3&0&0&1\\0&5&2&4\\0&1&9&5\\1&2&2&24\end{pmatrix}$$

$$\begin{pmatrix}2&0&0&0\\0&6&2&2\\0&2&6&5\\1&2&5&23\end{pmatrix}\begin{pmatrix}2&0&0&1\\0&5&2&2\\0&2&7&5\\1&2&4&23\end{pmatrix}\begin{pmatrix}2&0&0&1\\0&5&2&4\\1&2&6&5\\1&2&6&21\end{pmatrix}\begin{pmatrix}2&0&0&1\\0&5&1&4\\1&1&6&6\\1&3&6&21\end{pmatrix}$$

We will now identify the conditional attributes that affect information systems. Since the system has a large number of conditional attributes and there is no need to consider every attribute (due to the aforementioned over fitting situation), a heuristic algorithm based on a Boolean function is adopted to identify the conditional attributes with the biggest impact. A symmetrical 206 × 206 input-based discriminant matrix is plotted for this information system – with the input consisting of staff and locational attributes – whose elements come in the form of Cij as follows:

$$C_{ij} = \{b \in B \mid b(x_i) \neq b(x_j)\} \qquad for\ i,j = 1,\dots,206 \tag{9}$$

$$f_A(a_1^*,\dots,a_{15}^*) = \wedge \{\vee C_{ij} \mid 1 \leq j \leq i \leq n\,, C_{ij} \neq 0\} \quad \text{n=1,...15}$$

C_{ij}	Elements of the decision matrix
x_i, x_j	Discriminant objects
f_A	Discriminant function
a_n^*	Boolean variable

Note: The Boolean variable corresponds with the system's conditional attributes, of which there are 15 in this problem.

The studied bank has 206 branches and its decision-information system consists of 15 conditional attributes and one decision attribute (efficiency score). Four job categories were identified in the bank: branch manager and deputies (in charge of executive and controlling practices), central office managers and deputies (supervisors in charge of controlling their subordinate branches and conducting other specific operations), clerks (in charge of carrying out all banking transactions and interacting with customers), and other staff (archivists, service employees, security guards, etc.). Table 3 lists the decision-information system's attributes:

Table 3. decision-information system's attributes

Location	Staff
Commercial value of the area	Total number of staff
	Number of branch manager and deputies
	Number of supervisors (central office manager and deputies)
	Number of clerks
	Number of other employees
Income level of area residents	Number of employees with associate or lower degrees
	Number of employees with bachelor's degrees
	Number of employees with master's or higher degrees
	Average age of clerks
	Average work experience of branch manager and deputies
	Average work experience of supervisors
	Average work experience of clerks
	Average work experience of other employees

According to the software's output, effective combinations of conditional attributes which have the biggest impact on the efficiency of the decision-information system, are as follows:

- {Commercial value of the area, number of clerks, average work experience of supervisors, average work experience of branch manager and deputies, number of employees with bachelor's degrees}
- {Commercial value of the area, total number of staff, number of clerks, of employees with bachelor's degrees}
- {Commercial value of the area, income level of area residents, number of clerks, number of employees with master's or higher degrees}
- {Commercial value of the area, average age of staff, number of supervisors, average work experience of branch manager and deputies, number of employees with bachelor's degrees}
- {Income level of area residents, number of supervisors, average work experience of branch manager and deputies, number of employees with master's or higher degrees}
- {Commercial value of the area, income level of area residents, number of clerks, average work experience of supervisors, number of employees with bachelor's degrees}
- {Income level of area residents, number of clerks, average work experience of branch manager and deputies, number of employees with bachelor's degrees}
- {Income level of area residents, total number of staff, average age of staff, average work experience of supervisors, number of employees with master's or higher degrees}
- {Commercial value of the area, income level of area residents, number of supervisors, average work experience of branch manager and deputies, number of employees with bachelor's degrees}
- {Commercial value of the area, income level of area residents, total number of staff, average work experience of supervisors, number of employees with bachelor's degrees}

After determining the best combination of conditional attributes, it is time to find the most influential attributes i.e. factors from among the combinations. The best combination is the one which describes the bank branches' efficiency in the most comprehensive and accurate way, which we find by forming the perceptron neural network again. Next, as with the previous network, a number of neurons are tested for each of the 10 conditional sets we have obtained in order to develop a network which has the lowest error margin and highest accuracy. This model must now be tested for accuracy. As previously mentioned, neural networks are able to estimate the functions with acceptable accuracy. For each of the obtained conditional sets obtained, neural networks with 2, 4, 6, 8, 10, 12, 14, 16, 18, and 20 neurons in the hidden layer were built. Therefore, the number of neural networks built for each set is 40 which add up to a total of 400 networks for the 10 groups. The second neural network is validated using the k-fold cross-validation technique where k = 4. This means that the information system is divided into 4 folds and in each instance, one fold is treated as test data and the other three as training data. In the end, the average neural network error rate – obtained after repeating the test four times on a specific number of neurons in the hidden layer – is considered as the final result.

Fig 6. Sample classification

The results of 4-fold validation tests on the neural network model are detailed in the table below. The network's input consisted of the aforementioned conditional attributes, while the output is the decision attribute i.e. efficiency. Tables 4 to 13 represents the results of conditional sets:

Table 4. Performance of the set

20	18	16	14	12	10	8	6	4	2	**Result**
0.169	0.198	0.159	0.221	0.175	0.084	0.114	0.095	0.126	0.073	**1**
0.248	0.176	0.236	0.149	0.141	0.149	0.176	0.101	0.143	0.154	**2**
0.161	0.229	0.151	0.167	0.226	0.171	0.129	0.096	0.089	0.131	**3**
0.272	0.199	0.287	0.198	0.192	0.143	0.109	0.105	0.191	0.116	**4**
0.213	0.201	0.2084	0.184	0.184	0.137	0.1329	0.099	0.1379	0.119	**AERR$_1$**

Table 5. Performance of the set

20	18	16	14	12	10	8	6	4	2	**RESULT**
0.223	0.204	0.171	0.158	0.143	0.137	0.145	0.124	0.113	0.112	**1**
0.219	0.195	0.186	0.173	0.164	0.148	0.134	0.129	0.127	0.109	**2**
0.215	0.212	0.182	0.177	0.159	0.153	0.142	0.132	0.118	0.121	**3**
0.228	0.214	0.169	0.168	0.172	0.155	0.139	0.136	0.125	0.114	**4**
0.221	0.201	0.177	0.169	0.160	0.148	0.140	0.130	0.121	0.114	**AERR$_2$**

Table 6. Performance of the set

20	18	16	14	12	10	8	6	4	2	RESULT
0.225	0.209	0.193	0.171	0.173	0.158	0.142	0.146	0.141	0.135	**1**
0.229	0.219	0.188	0.165	0.178	0.165	0.151	0.139	0.129	0.127	**2**
0.215	0.198	0.184	0.167	0.169	0.167	0.148	0.142	0.139	0.144	**3**
0.219	0.216	0.205	0.174	0.166	0.159	0.153	0.149	0.135	0.141	**4**
0.222	0.211	0.193	0.170	0.172	0.162	0.149	0.144	0.136	0.137	**$AERR_3$**

Table 7. Performance of the set

20	18	16	14	12	10	8	6	4	2	RESULT
0.195	0.209	0.161	0.175	0.129	0.118	0.124	0.102	0.099	0.097	**1**
0.213	0.186	0.158	0.173	0.115	0.136	0.106	0.116	0.101	0.116	**2**
0.228	0.223	0.179	0.169	0.123	0.133	0.133	0.127	0.105	0.113	**3**
0.231	0.217	0.173	0.157	0.134	0.145	0.127	0.119	0.098	0.125	**4**
0.217	0.209	0.167	0.169	0.125	0.133	0.123	0.116	0.101	0.113	**$AERR_4$**

Table 8. Performance of the set

20	**18**	**16**	**14**	**12**	**10**	**8**	**6**	**4**	**2**	**Result**
0.242	0.215	0.217	0.187	0.163	0.152	0.147	0.121	0.132	0.143	**1**
0.233	0.229	0.209	0.192	0.176	0.157	0.136	0.125	0.137	0.135	**2**
0.236	0.234	0.226	0.201	0.173	0.149	0.149	0.131	0.144	0.131	**3**
0.231	0.238	0.213	0.203	0.167	0.158	0.145	0.113	0.127	0.146	**4**
0.236	0.229	0.216	0.196	0.170	0.154	0.144	0.123	0.135	0.139	**$AERR_5$**

Table 9. Performance of the set

20	**18**	**16**	**14**	**12**	**10**	**8**	**6**	**4**	**2**	**Result**
0.164	0.155	0.148	0.137	0.145	0.131	0.093	0.093	0.112	0.088	**1**
0.176	0.167	0.156	0.146	0.152	0.121	0.102	0.115	0.129	0.126	**2**
0.179	0.173	0.168	0.157	0.138	0.114	0.097	0.133	0.106	0.107	**3**
0.161	0.178	0.164	0.163	0.159	0.119	0.101	0.131	0.118	0.119	**4**
0.170	0.169	0.159	0.151	0.149	0.121	0.098	0.118	0.116	0.110	**$AERR_6$**

Table 10. Performance of the set

20	18	16	14	12	10	8	6	4	2	Result
0.148	0.139	0.143	0.122	0.113	0.101	0.112	0.084	0.078	0.069	**1**
0.139	0.131	0.123	0.135	0.124	0.132	0.095	0.081	0.124	0.084	**2**
0.144	0.147	0.145	0.127	0.138	0.126	0.107	0.079	0.109	0.115	**3**
0.136	0.129	0.129	0.141	0.133	0.108	0.128	0.083	0.093	0.121	**4**
0.142	0.137	0.135	0.131	0.127	0.117	0.111	0.082	0.101	0.097	**$AERR_7$**

Table 11. Performance of the set

20	18	16	14	12	10	8	6	4	2	Result
0.217	0.215	0.194	0.176	0.167	0.164	0.151	0.143	0.129	0.134	**1**
0.227	0.208	0.206	0.182	0.175	0.149	0.148	0.134	0.122	0.138	**2**
0.231	0.223	0.211	0.184	0.166	0.157	0.157	0.137	0.145	0.135	**3**
0.218	0.216	0.199	0.179	0.179	0.161	0.162	0.147	0.132	0.142	**4**
0.223	0.216	0.201	0.180	0.172	0.158	0.155	0.140	0.132	0.137	**$AERR_8$**

Table 12. Performance of the set

20	18	16	14	12	10	8	6	4	2	Result
0.228	0.217	0.177	0.151	0.155	0.134	0.126	0.116	0.104	0.109	**1**
0.235	0.186	0.166	0.176	0.167	0.141	0.132	0.108	0.115	0.104	**2**
0.195	0.209	0.184	0.160	0.153	0.148	0.119	0.103	0.122	0.113	**3**
0.231	0.193	0.171	0.178	0.162	0.152	0.127	0.123	0.117	0.106	**4**
0.222	0.201	0.175	0.166	0.160	0.144	0.126	0.113	0.115	0.108	**$AERR_9$**

Table 13. Performance of the set

20	18	16	14	12	10	8	6	4	2	Result
0.216	0.189	0.163	0.141	0.119	0.134	0.126	0.108	0.123	0.105	**1**
0.194	0.213	0.178	0.158	0.137	0.145	0.132	0.119	0.113	0.124	**2**
0.232	0.227	0.168	0.165	0.128	0.138	0.139	0.111	0.103	0.107	**3**
0.207	0.221	0.201	0.159	0.139	0.146	0.125	0.108	0.121	0.118	**4**
0.212	0.213	0.178	0.156	0.132	0.141	0.131	0.111	0.115	0.116	**$AERR_{10}$**

4-1-Ranking the conditional sets

At this stage, we rank the conditional sets obtained in the previous section. To do so, the best possible combination (a network with a specific number of neurons which has the lowest error rate) from each conditional set is selected. Next, a maximum, minimum, and deviation value is determined for each combination, which are then ranked using a DEA approach by weighting the values. The goal is to find the

neural network with the highest efficiency score. DEA is a suitable method for ranking conditional sets in information systems as it helps measure the efficiency of each mode. Based on the results, it appears that the following conditional set has the highest efficiency score in predicting the performance level of branches studied in this research:
Best type: {Income level of area residents, number of clerks, average work experience of branch manager and deputies, number of employees with bachelor's degrees}
Table 14 represents the best performance of each set and then it demonstrates the best performance among all sets.

Table 14. Identifying Best Set

Rank	Efficiency	Standard Deviation	$AERR_i$	$MaxERR_i$	$MinERR_i$	**Conditional set (Best)**
3	1.028	0.004	0.099	0.105	0.095	**Table 4**
7	0.923	0.004	0.114	0.121	0.109	**Table 5**
10	0.876	0.005	0.136	0.141	0.135	**Table 6**
4	1.009	0.003	0.101	0.105	0.098	**Table 7**
8	0.905	0.006	0.123	0.131	0.113	**Table 8**
2	1.044	0.003	0.098	0.102	0.093	**Table 9**
1	1.073	0.002	0.082	0.084	0.079	**Table 10**
9	0.882	0.008	0.132	0.145	0.122	**Table 11**
5	0.981	0.003	0.108	0.113	0.104	**Table 12**
6	0.954	0.005	0.111	0.119	0.108	**Table 13**

5-Conclusion

Banking managers, experts, and policymakers are constantly on the lookout for reliable, science-based methods based which to devise strategic and fundamental plans in order to maximize effectiveness, efficiency, and productivity. In complex financial institutes such as banks where branches, as the main line units, have a significant role in drawing resources, profit, and other qualitative and quantitative contributions, it is essential to develop precise plans and to take decisive practical steps toward improving the overall efficiency of the branches and thereby the institution as a whole. The first and foremost goal of development banks is achieving maximum profitability for investment in specialized sectors. It is important to conduct various studies and researches in order to achieve the best conditions that can lead to increased efficiency. This study shows how accurately identifying the strengths of premium branches and extending them to other branches can develop the efficiency and profitability of the whole bank and greatly improve the position of the studied bank in today's competitive market. In this study, the initial data were collected in accordance with the efficiency scores and ranking of an Iranian development bank's branches using MCDM techniques. After identifying the highest-scoring branches based on the aforementioned ranking and using data mining techniques, the input variables with the most impact on the branch's performance were determined. This gives bank executives valuable insights about how to (re)locate their branches and recruit the required employees based on the most effective combination of factors and selection criteria in order to improve their efficiency level. If banking experts are given reliable information about the factors responsible for productivity and success of the best bank branches around the world, they will have the necessary tools to localize the factors, adapt them to their respective bank's rules and standards, and ultimately maximize the bank's overall productivity. Using the outcomes of various investigations and computations allows banking decision-makers to define a specific period of time and, taking the general rules and top-down regulations of their respective banks into account, begin to adjust and restructure their branches in order to be maximally compliant with the latest models and thereby take full advantage of their benefits at every organizational level. In the public sector, various guidelines and instructions are routinely issued by government officials, complying with which is mandatory. The bank investigated in this study is a state-controlled financial institution as well.

It should be noted that identifying the factors that affect organizational efficiency is always of utmost importance to policymakers and executives of the banking industry. Therefore, in this study, numerous such factors were identified, tested, and ranked using ANN approaches. Based on the rankings, several combinations of the top influential factors were tested in order to find the most effective combination. The findings of this research could help decision-makers determine the optimal locations and staff attributes for their branches and, ultimately, take considerable steps toward improving the efficiency and profitability of their respective banks' branches.

The approach of this research is relate to banking industry which is considered the most significant organization among all financial and economic bodies. Yet, in order to enhance this research for possible usage in other financial and economic formations the general structure would remain constantly which means the steps of the appraisal process is similar to the measures described in detail in this article, both in the research methodology section and in the case study part. Despite the efficiency of the overall structure of the article for numerous similar studies, it should be noted that since banks have several unique features such as having a wide variety of branches with different services all over the country, very high variety of services and facilities to offer to customers and etc. Therefore, the performance criteria of this type of organization would differ others with limited responsibilities and services. If other financial bodies like Investment Companies, Insurance Companies, Brokerage Companies and etc. want to evaluate their performance utilizing this research, they could determine their impressive criteria via methods mentioned in this paper (the type of criteria for each institution varies depending on its nature and a specific set of criteria cannot be considered for all financial institutions), then enjoy the described structure for performance measurement in order to finding criteria which have direct impact on increasing the level of efficiency.

References

Adamu, J. A. (2019). Advanced Stochastic Optimization Algorithm for Deep Learning Artificial Neural Networks in Banking and Finance Industries. Risk and Financial Management, 1(1), p8-p8.

Ahamed, K., & Akthar, S. (2016). Survey on artificial neural network learning technique algorithms. International Research Journal of Engineering and Technology, 3(02), 36-39.

Akbari, F., & Osgooei, E. (2020). Solving linear equation system based on Z-numbers using big-M method. Journal of Industrial and Systems Engineering.

Anouze, A. L. M., & Bou-Hamad, I. (2019). Data envelopment analysis and data mining to efficiency estimation and evaluation. International Journal of Islamic and Middle Eastern Finance and Management.

Appiahene, P., Missah, Y. M., & Najim, U. (2020). Predicting Bank Operational Efficiency Using Machine Learning Algorithm: Comparative Study of Decision Tree, Random Forest, and Neural Networks. Advances in Fuzzy Systems, 2020.

Azadeh, A., Saberi, M., Moghaddam, R. T., & Javanmardi, L. (2011). An integrated data envelopment analysis–artificial neural network–rough set algorithm for assessment of personnel efficiency. Expert Systems with Applications, 38(3), 1364-1373.

Azadeh, A., Saberi, M., & Jiryaei, Z. (2012). An intelligent decision support system for forecasting and optimization of complex personnel attributes in a large bank. Expert Systems with Applications, 39(16), 12358-12370.

Beale, H. D., Demuth, H. B., & Hagan, M. T. (1996). Neural network design. Pws, Boston.

Bacaksız, A. H., & Esgin, E. (2019, October). Extraction of Numerical data from Categorical Data Set and Artificial Neural Networks. In 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) (pp. 1-4). IEEE.

Bakar, N. M. A., & Tahir, I. M. (2009). Applying multiple linear regression and neural network to predict bank performance. International Business Research, 2(4), 176-183.

Baumann, K. (2003). Cross-validation as the objective function for variable-selection techniques. TrAC Trends in Analytical Chemistry, 22(6), 395-406.

Beynon, M. J., & Peel, M. J. (2001). Variable precision rough set theory and data discretisation: an application to corporate failure prediction. Omega, 29(6), 561-576.

Daliri, S. (2020). Using Harmony Search Algorithm in Neural Networks to Improve Fraud Detection in Banking System. Computational Intelligence and Neuroscience, 2020.

Dincer, H., Gencer, G., Orhan, N., & Sahinbas, K. (2011). A performance evaluation of the Turkish banking sector after the global crisis via CAMELS ratios. Procedia-Social and Behavioral Sciences, 24, 1530-1545.

Dungey, M., & Gajurel, D. (2015). Contagion and banking crisis–International evidence for 2007–2009. Journal of Banking & Finance, 60,

Farooqi, M. R., & Iqbal, N. (2017). Effectiveness of Data mining in Banking Industry: An empirical study. International Journal, 8(5).

Hassoun, M. H. (1995). Fundamentals of artificial neural networks. MIT press.

Ho, C. T. B., & Wu, D. D. (2009). Online banking performance evaluation using data envelopment analysis and principal component analysis. Computers & Operations Research, 36(6), 1835-1842.

Idris, S. M., & Ahmad, F. A. (2020). Islamic Banking Efficiency: A Systematic Litrature Review. Journal of Islamic, 5(29), 41-50.

Ilie, M., Moraru, A. D., & Ghita-Mitrescu, S. (2017). THE HIERARCHICAL DETERMINATION OF CUSTOMER SATISFACTION WITH BANKING SERVICES USING AN ARTIFICIAL NEURAL NETWORK. Transformations in Business & Economics, 16.

Inuiguchi, M., & Miyajima, T. (2007). Rough set based rule induction from two decision tables. European journal of operational research, 181(3), 1540-1553.

Jakšič, M., & Marinč, M. (2019). Relationship banking and information technology: The role of artificial intelligence and FinTech. Risk Management, 21(1), 1-18.

Jemric, I., & Vujcic, B. (2002). Efficiency of banks in Croatia: A DEA approach. Comparative Economic Studies, 44(2-3), 169-193.

Kanmani, W. S., & Jayapradha, B. (2017). Prediction of Default Customer in Banking Sector using Artificial Neural Network. International Journal on Recent and Innovation Trends in Computing and Communication, 5(7), 293-296.

Kassani, S. H., Kassani, P. H., & Najafi, S. E. (2018). Introducing a hybrid model of DEA and data mining in evaluating efficiency. Case study: Bank Branches. arXiv preprint arXiv:1810.05524.

Kataria, S., & Nafis, M. T. (2019, March). Internet Banking Fraud Detection Using Deep Learning Based on Decision Tree and Multilayer Perceptron. In 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 1298-1302). IEEE

Lin, T. Y. (2018). Two-Stage Performance Evaluation of Domestic and Foreign Banks in Taiwan. Asian Journal of Economic Modelling, 6(2), 191-202.

Mansingh, G., Rao, L., Osei-Bryson, K. M., & Mills, A. (2015). Profiling internet banking users: A knowledge discovery in data mining process model based approach. Information Systems Frontiers, 17(1), 193-215.

Mostafa, M. M. (2009). Modeling the efficiency of top Arab banks: A DEA–neural network approach. Expert systems with applications, 36(1), 309-320.

Niknafs, J., Keramati, M. A., & Monfared, J. H. Estimating Efficiency of Bank Branches by Dynamic Network Data Envelopment Analysis and Artificial Neural Network. Advances in Mathematical Finance and Applications, 5(3), 1-15.

Patil, P. S., & Dharwadkar, N. V. (2017, February). Analysis of banking data using machine learning. In 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 876-881). IEEE.

Parker, J. R. (2001). Rank and response combination from confusion matrix data. Information fusion, 2(2), 113-120.

Pawlak, Z. (1982). Rough sets. International journal of computer & information sciences, 11(5), 341-356.

Owen, A. L., & Temesvary, J. (2018). The performance effects of gender diversity on bank boards. Journal of Banking & Finance, 90, 50-63.

Raju, P. S., Bai, D. V. R., & Chaitanya, G. K. (2014). Data mining: Techniques for enhancing customer relationship management in banking and retail industries. International Journal of Innovative Research in Computer and Communication Engineering, 2(1), 2650-2657.

Ristyawan, M. R. (2021). Artificial Neural Network and Analytical Hierarchy Process Integration: A Tool to Estimate Business Strategy of Bank (No. jfbr179). Global Academy of Training and Research (GATR) Enterprise.

Roy, S. (2020, December). Performance Modelling on Banking System: A Data Envelopment Analysis-Artificial Neural Network Approach. In International Working Conference on Transfer and Diffusion of IT (pp. 581-597). Springer, Cham.

Santin, D., Delgado, F. J., & Valino, A. (2004). The measurement of technical efficiency: a neural network approach. Applied Economics, 36(6), 627-635.

Shokrollahpour, E., Lotfi, F. H., & Zandieh, M. (2016). An integrated data envelopment analysis–artificial neural network approach for benchmarking of bank branches. Journal of Industrial Engineering International, 12(2), 137-143.

Skowron, A., & Rauszer, C. (1992). The discernibility matrices and functions in information systems. In Intelligent decision support (pp. 331-362). Springer, Dordrecht.

Tavana, M., Abtahi, A. R., Di Caprio, D., & Poortarigh, M. (2018). An Artificial Neural Network and Bayesian Network model for liquidity risk assessment in banking. Neurocomputing, 275, 2525-2554.

Wanke, P., Azad, M. A. K., & Barros, C. P. (2016). Predicting efficiency in Malaysian Islamic banks: A two-stage TOPSIS and neural networks approach. Research in International Business and Finance, 36, 485-498.

Wanke, P., Azad, M. A. K., Barros, C. P., & Hassan, M. K. (2016). Predicting efficiency in Islamic banks: An integrated multicriteria decision making (MCDM) approach. Journal of International Financial Markets, Institutions and Money, 45, 126-141.

Wang, K., Huang, W., Wu, J., & Liu, Y. N. (2014). Efficiency measures of the Chinese commercial banking system using an additive two-stage DEA. Omega, 44, 5-20.

Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining (pp. 29-39). London, UK: Springer-Verlag.

Wu, D. D., Yang, Z., & Liang, L. (2006). Using DEA-neural network approach to evaluate branch efficiency of a large Canadian bank. Expert systems with applications, 31(1), 108-115.