ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

MULTI CRITERIA ABC CLASSIFICATION OF LIGHT RAILWAY SPARE PARTS USING ARTIFICIAL NEURAL NETWORK APPROACH

M.Sc. Thesis by
Jarkyn JUMABAEVA

Department: Management Engineering

Programme: Management Engineering

JUNE 2011

İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

MULTI CRITERIA ABC CLASSIFICATION OF LIGHT RAILWAY SPARE PARTS USING ARTIFICIAL NEURAL NETWORK APPROACH

M.Sc. Thesis by
Jarkyn JUMABAEVA
(507071030)

Date of submission : 06 May 2011
Date of defence examination: 06 June 2011

Supervisor (Chairman) : Prof. Dr. Demet BAYRAKTAR (ITU)
Members of the Examining Committee : Assoc. Prof. Dr.İlker TOPCU (ITU)
Asst. Prof. Dr. Dilay ÇELEBİ (ITU)

JUNE 2011

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

HAFİF RAYLI SİSTEMLERDE YEDEK PARÇA STOKLARININ SINIFLANDIRILMASI İÇİN ÇOK ÖLÇÜTLÜ ABC ANALİZİ İLE YAPAY SİNİR AĞI YAKLAŞIMI

YÜKSEK LİSANS TEZİ

Jarkyn JUMABAEVA
(507071030)

Tezin Enstitüye Verildiği Tarih : 06 Mayıs 2011
Tezin Savunulduğu Tarih : 06 Haziran 2011

Tez Danışmanı : Prof. Dr. Demet BAYRAKTAR (İTÜ)
Diğer Jüri Üyeleri : Doç.Dr. İlker TOPCU (İTÜ)
Yrd. Doç. Dr. Dilay Çelebi (İTÜ)

HAZİRAN 2011

FOREWORD

I am very grateful to Prof. Dr. Demet Bayraktar for her supervision and helpful directions throughout the study. I would also like to express my deep appreciation and thanks forAsst.Prof.Dr. Dilay Çelebi for helping with gathering data for this thesis.

I also thankful to my husband and family for their continuous support and encouragement during the preperation of this thesis.

June 2011

Jarkyn Jumabaeva

Management Engineering

TABLE OF CONTENTS

Page

ABBREVIATIONS

ABC-FC **:** ABC–Fuzzy Classification
AHP **:** Analytical Hierarchy Process
ANN **:** Artificial Neural Network
App **:** Appendix
BPA **:** Back Propagation Algorithm
DEA **:** Data Envelopment Analysis
EFNN **:** Enhanced Fuzzy Neural Network
FAHP **:** Fuzzy Analytical Hierarchy Process
GA **:** Genetic Algorithm
k-NN **:** k-nearest neighbour
MBPN **:** Moving Back-propagation neural network
MC **:** Multi criteria
MCDA **:** Multi criteria decision aid
MCDM **:** Multi Criteria Decision Making
MDA **:** Multiple Discriminant Analysis
METRIC **:** Multi-echelon technique for recoverable item control
MFNN **:** Moving Fuzzy Neuron Network
MLP **:**Multi Layer Perception
MSE **:** Mean Squared Error
PE **:** Processing Element
RCM **:** Reliability Centered Maintenance
SVM **:** Support Vector Machine
VLSI **:**Very-large-scale-integrated

LIST OF TABLES

Page

LIST OF FIGURES

Page

MULTI CRITERIA ABC CLASSIFICATION OF LIGHT RAILWAY SPARE PARTS USING ARTIFICIAL NEURAL NETWORK APPROACH

SUMMARY

Management of spare parts inventories is crutial for successfull execution of the maintenance processes. Main attention is paid to maintenance operations in railway systems management because it directly affects the performance of railway machines and play important role in railway service, in order to provide uninterrupted and high quality service to passengers. The most important characteristics of railway services are the availability and reliability of spare items, which are necessary for maintenance operations. The role of spare parts is to support the maintenance operations in keeping vehicles in a functioning condition. Supplying an adequate and yet efficient amount of spare parts to support the maintenance activities in railway service is a tough management problem.

In this thesis, artificial neural network model was developed to classify spare parts inventory based on multi criteria ABC analysis in light railway systems. After thorough literature review on inventory classification, the relevant classification criteria and control chracteristics of maintenance spare parts are identified and discussed in terms of their effects on maintenance operations, purchasing characteristics, positioning of materials, responsibility of control and control principles. The classification model of railway systems spare parts have been performed based on following criteria: annual unit cost, lead time, value usage, substitutability, commonality and stock-out penalty.

Most widely used neural network structure is chosen, which contains one input, one hidden and one output layer. For the training of the network which will be used for classification of items into A, B and C class items, a multi layer perception network with back propagation algorithm was used. The optimum architechture is determined by estimating number of nodes in hidden layer by five-fold cross validation method. The comparison of classification accuracy was done between neural networks with hyperbolic tangent function and sigmoid transfer function. As a result of training, neural network with hyperbolic tangent function gave more accurate results, by giving the lowest the minimum mean squared error, MSE. The proposed artificial neural network architecture has 6-8-3 structure, which constitutes six nodes in the input layer, eight nodes in the hidden layer and three nodes in the output layer.

The evaluation data of 71 maintenance spare part items were utilized for training and testing processes, specifically, 60 item data for training and 11 items data for testing the developed artificial neural network.

The main contribution of this thesis is that, it applies a different approach, which is ANN, for classification of spare parts inventory in railway systems, which was not widely investigated area in the literature. The previous researches in literature on multi criteria inventory management have been done using analytical hierarchy process and data envelopment analysis). The developed network has shown good classification accuracy and this network can be used for classification of other real-world inventory data in railway systems maintenance.

HAFİF RAYLI SİSTEMLERDE YEDEK PARÇA STOKLARININ SINIFLANDIRILMASI İÇİN ÇOK ÖLÇÜTLÜ ABC ANALİZİ İLE YAPAY SİNİR AĞI YAKLAŞIMI

ÖZET

Bakım işlerinin başarılı bir şekilde gerçekleşmesi için yedek parça stoklarının yönetimi çok önemlidir. Raylı sistemlerin yönetiminde bakım işlemlerine büyük önem verilmesinin sebebi, bakım işlemlerinin raylı sistemlerde kullanılan araçların performasını direk etkilemesi ve yolculara kesintisiz ve yüksek kalitede hizmet sağlanmasında çok önemli rol oynamaktadır. Raylı sistemlerin hizmetinde en önemli iki husus, bakım işleri için gerekli olan yedek parçaların yeterli miktarda elde bulundurulması ve güvenilir olmasıdır. Yedek parçaların amacı araçların çalışır vaziyette tutulması için gerekli bakım işlerinde kullanılması. Raylı sistemlerin hizmetinde yapılan bakım işleri için yeterli ve aynı zamanda etkin bir şekilde yedek parça sağlamak, yönetim için zor bir iştir.

Bu çalışmanın amacı hafif raylı sistemlerde kullanılan yedek parça stoklarının çok kriterili ABC analizi ile sınıflandırılması için yapay sinir ağı modeli geliştirmektir. Envanter sınıflandırma üzerine kapsamlı literatür incelemesi yapıldıktan sonra, bakım için kullanılan yedek parçaların sınıflandırılma ölçütleri ve kontrol parametreleri belirlendi. Belirlenen sınıflandırma ölçütlerin bakım işleri, satın alma özellikleri, malzemenin konulmandırması, kontrol sorulluluğu ve denetim prensipleri açılarından tartışıldı. Hafif raylı sistemlerin yedek parça sınıflandırma modelinde değerlendirme süreci için elde bulundurmama maliyeti, ikame edilebilirlik, kullanım adedi, maliyet, tedarik süresi ve ortak nokta kriterileri belirlenmiştir.

En yaygın kullanılan, tek girdi, tek gizli ve tek çıktı katmanlı, yapay sinir ağı modeli seçilmiştir. Yedek parçaların A,B ve C gruplarına sınıflandırmasın için kullanılan ağın eğitililmesi için çok katmanlı algılama ağ modeli ile geri yayılım algoritmasından yararlanılmıştır. Uygun ağ yapısını blirlenebilmesi için beşli çapraz doğrulama yöntemi ile gizli katmandaki sinir hücre sayısı tespit edilmiştir. Sınıflandırma doğrulaması için, hiperbolik tanjantlı ve sigmoid transfer fonsiyonlu iki farklı yapay sinir ağı modeli karşılaştırılmıştır. Ağın eğitimi sonucunda, hiperbolik tanjant fonksiyonlu yapay sinir ağının daha düşük hata oranı gösterdiği, ortalama karesel hatadan belirlenmiştir. Önerilen YSA modelinde; yedek parça envanteri sınıflandırma kriterilerini temsil eden 6 adet girdi sinir hücresi, tek gizli katmanda yer alan 8 adet sinir hücresi ve sınıflandırma kümelerini gösteren 3 çıktılı ileri beslemeli bir ağ yapısı oluşturulmuştur.

Değerlendirmede eğitim ve test işlemleri için toplam 71 adet yedek parça verisi kullanılmıştır, bunların, 60 adet yedek parça verisi eğitim işlemi ve 11 adet yedek parça verisi test işlemi için değerlendirilmiştir.

Yapılan çalışmanın en önemli katkısı, hafif raylı sistemlerin bakım işlerinde kullanılan yedek parça envanter sınıflandırması çalışmasına farklı yöntemle yaklaşılmış olması ve bu alanın yapılan çalışmalar arasında daha önce incelenmemiş olması. Daha önce yapılan çalışmalar arasında çok kriterili ABC analizi ile hafif raylı sistemlerin yedek parça stokların sınıflandırmasında analitik hiyerarşi prosesi ve veri zarflama analizi yöntemleri kullanılarak incelenmiştir. Bu çalışmada, hafif raylı sistemlerinde kullanılan envanter sınıflandırması için yapay sinir ağı yaklaşımı önerilmiştir. Yedek parça sınıflandırma modeli olarak geliştirilen YSA modeli başarılı sonuçlar göstermiştir. Geliştirilen YSA modeli diğer raylı sistemlerin de bakım işleri için kullanabileceği bir yapıya dönüştürülebilir.

INTRODUCTION

Inventory management is one the main objectives of production and operations management. The role of inventory managers is to maintain sufficient inventoires to meet demand and achieve productivity, while at the same time to incur the lowest possible cost. In the literature there are several models and approaches that have been developed for planning and controlling inventories (Leenders et.al., 1989). ABC analysis is used to classify items into three groups for management and control of inventories with large number of different items. Items are classified according to their annual dollar usage in traditional ABC analysis. Howewer, there are instances where criteria other than unit price and anuual usage are important. Thus, inventory managers have to decide how to take these different criteria into account. It has been suggested by Flores and Whybark (1986) that multiple criteria ABC classification can provide a more comprehensive managerial approach.Withmultiple criteria certain mechanical methods are required to reduce the classification to an ABC grouping. An artificial neural network approach, pioneered by McCulloch and Pitts (1943), can be used to classify large amount of items in terms of multi criteria classification in order to reach inventory management objectives.

1.1 Purpose of the Thesis

For effective functioning of any organization, it is essential to implement efficient performance measurement system. Performance in a company must be controlled throughout every operation and functioning units. There are several performance and operation control methods and almost in every organization main attention is paid to maintenance policies, which are known as age replacement, periodic repair, failure limit, etc. Each of these maintenance policies have many advantages and disadvantages. Several known maintenance techniques are productive maintenance, condition based maintenance and predictive maintenance, which all somehow necessitate the importance of effective inventory management, to provide efficient operating conditions in any organization.

Inventory management is one of the major points in production scheduling and operations management. It is very important to maintain sufficient amount of inventory to meet operational demand and achieve productivity and at the same time maintain possible costs at the minimum level. There are several techniques developed to plan and control inventory levels in the organization to minimize costs of keeping inventory and diminish levels of stock-out of inventory. Methods developed are mostly based on mathematical and classification approaches. Mathematical approaches developed for inventory management are linear programming, dynamic programming, goal programming, simulation and etc. During the last decades, various models have been examined and developed for inventory management based on mathematical optimization approaches. Multi-echelon technique for recoverable item control (METRIC) developed by Sherbrooke (1968) is the early study, which concentrates on optimization of inventory costs and service levels connected with a potential spare parts management policy based on economic order quantity, reorder point, safety stock and so on. There are several drawbacks in these developed models like complexity, abstract and oversimplified structures and they do not include factors like obsolescence, standard characteristics of the item, type and quality of suppliers, etc.

Main attention is given to inventory classification schemes as a spare parts management, because of its need to keep ample sizes of inventories to provide for unexpected breakdowns and at the same time keeping minimum inventory level in terms of stock costs. Generally, classification problems are often encountered in a variety of fields including finance, marketing, environmental and energy management, human resources management, medicine, etc. Many investigations have been done on inventory classification in various industries like production, pharmaceuticals, services industries like hospitals, and many other areas. However, little research has been done on inventory management in railway systems. Specifically, the objective of this study is to develop multi criteria inventory classification model based on ABC classification for maintenance spare parts in railway systems.

The research methodology will include developing an artificial neural network which will be used to classify spare parts inventories in railway industry based on multi criteria ABC classification. The reason why neural network approach is chosen for the classification in this study is that there is learning process that enable the model to catch up the nonlinear relation in the problem pattern.

1.2 Research Scope

In the new millennium, the transportation of people and materials has even greater significance world-wide than it had in the immediate past. The birth and evolution of railways during the last two centuries have had an enormous influence on industrial, social and economic development (Bonnett,2005).

Every railway needs to establish for itself a regime for the regular maintenance of rolling stock. The various activities, involving both inspection and work to components, will necessitate some time when stock is not available to get back to normal operation. It is difficult to get the balance right between over and under maintenance and this must be watched carefully, allowing suitable adjustments to be made in both procedures and frequencies when these are shown to be necessary (Bonnett,2005).

A little investigation has been done on inventory management in the area of railway systems. The importance to provide reliable and effective performance in railway systems is based on continuous or scheduled maintenance and unplanned repair. In other words, maintenance operations directly affect the performance of railway vehicles and play a crucial role in railway services to provide uninterrupted and high quality services to passengers.

Inventory management of spare parts in light railway systems has been a researh topic for studies done by Çelebi et al. (2008) using analytical hierarchy process (AHP) with data envelopment analysis (DEA) and later, by Sönmeztürk et.al (2008) using fuzyy AHP and DEA. In the context of our research knowledge through the literature, spare parts inventory classification with ABC analysis has not been mentioned to be implemented by the usage of artficial neural network for classification of spare parts inventory in light railway systems. So, the contribution of this research will be to propose a multi criteria classification model and develop artificial neural network model based on ABC analysis for classification of spare parts inventory of company performing maintenance operations in light railway systems.

1.3 Research Methodology

In this study artificial neural network model will be proposed for the classification of spare parts inventory in light railway systems. The design of the study will be conducted in three steps. First, a literature review will be done on inventory classification based on multi criteria ABC analysis and artificial neural network topologies will be researched. Within the literature review, findings on inventory classification will be analyzed and evaluation criteria for the study in light railway systems will be selected. Then, artificial neural network model that is suitable for inventory classification will be chosen and its structure and parameters will be analyzed. Two network structures will be trained and the more successful one will be chosen as the proposed model for the study. Finally, the developed neural netwrok will be tested to validate the performance of the network and results will be discussed.

2. INVENTORY MANAGEMENT

The objective of inventory management is to replace a very expensive asset called "inventory" with a less-expensive asset called "information." In order to accomplish this objective, the information must be timely, accurate, reliable and consistent (Viale, 1996). There are five basic types of inventory: raw material, work-in-process, finished goods, distribution inventory and maintenance, repair, and operating supplies which are present in almost every type of organizations. All of the following types of inventory act to buffer fluctuation in supply and demand. They add no value to the process and result in additional cost to carry, store, etc. However, they are necessary in order to ensure high levels of customer service. All types of businesses (retailers, manufacturers, banks) are challenged with the conflicting objectives of minimizing inventory and ensuring high levels of customer service (Viale, 1996).

Inventory management in most companies necessitates maintenance operations of the most inventories because efficient and effective management helps to maintain competitive advantage to keep up in a time of accelerating globalization and innovations in management techniques. The number of stock keeping units that is held by large companies may easily approach ten of thousands and it may not seem economical to design inventory management policy to each of the inventory unit. There is a need to differentiate the inventory according to the business model of the firm, because almost every business company differentiates in a way they manage and operate company. So, several techniques have been developed to classify the inventory according to their importance in the firm to enable meaningful categorization and thus provide to manageable way to control all inventory items in each category.

2.1 Inventory Classification

Inventory classification should be done in management of inventory because there are so many types of materials and spare parts and it is easy to lose sight of managing materials effectively (Çelebi et al., 2008; Parvoti and Anandarajan, 2002). To prevent a company from misallocating its materials management resources, planning and control systems must be implemented. Effective procurement planning and control system maintenance must be developed so that it would keep the balance between the keeping inventory costs to the minimum achievable level and protect the company from critical stock-outs of raw materials, work-in-progress inventories and finished goods.

Multi criteria decision aid (MCDA) has several distinctive and attractive features, involving mainly its decision support orientation which has found its crucial applications in inventory classification. The significant advances in multi criteria decision aid over the last three decades constitute a powerful non-parametric alternative methodological approach to study inventory classification problems. Although the MCDA researches until the late 1970s, have been mainly oriented towards the fundamental aspects of this field, as well as to the development of choice and ranking methodologies, during the 1980s and 1990s significant research has been undertaken on the study of the classification problem within the MCDA framework (Doumpos, 2002).

Classification itself is referred as to assign a finite set of alternatives into predefined groups. In this case the alternatives belonging into different groups have different characteristics, without being possible to establish any kind of preference relation between them (i.e., the groups provide a description of the alternatives without any further information). One of the most well-known problems of this form is the iris classification problem used by Fisher (1936) with a pioneering work on the linear discriminant analysis. This problem involves the distinction between three species of flowers, iris setosa, iris versicolor and iris virginica, given their physical characteristics (length and width of the sepal and petal). Obviously, each group (specie) provide a description of its member flowers, but this description does not incorporate any preferential information. Pattern recognition is also an extensively studied problem of this form with numerous significant applications in letter recognition, speech recognition, recognition of physical objects and human characteristics (Doumpos, 2002).

Doumpos (2002) has also made attention to the difference between classification and clustering of alternatives in decision making problems because they are usually confused. In classification the groups are defined a priori, whereas in clustering the objective is to identify clusters (groups) of alternatives sharing similar characteristics. In other words, in a classification problem the analyst knows in advance what the results of the analysis should look like, while in clustering the analyst tries to organize the knowledge embodied in a data sample in the most appropriate way according to some similarity measure.

The significance of the classification problems extends to a wide variety of practical fields of interest. Some characteristic examples are the following:

Medicine: medical diagnosis to assign the patients into groups (diseases) according to the observed symptoms (Tsumoto, 1998).

Pattern recognition: recognition of human characteristics or physical objects and their classification into properly defined groups (Nieddu and Patrizi, 2000).

Human resources management: personnel evaluation on the basis of its skills and assignment to proper working positions.

Production management: monitoring and control of complex production systems for fault diagnosis purposes (Shen, 2000).

Marketing: selection of proper marketing policies for penetration to new markets, analysis of customer characteristics, customer satisfaction measurement, etc. (Siskos et al., 1998).

Environmental management and energy policy: analysis and in time diagnosis of environmental impacts, examination of the effectiveness of energy policy measures (Diakoulaki et al., 1999).

Financial management and economics: bankruptcy prediction, credit risk assessment, portfolio selection (stock classification), country risk assessment (Zopounidis, 1998).

As Doumpos (2002) outlined about classification problems application field, it is widely used in modeling inventory classification problems. As was pointed out in previous sections, the aim of inventory management is to make right decision regarding the appropriate level of inventory. In practice, all inventories cannot be controlled with equal attention. The most widespread technique used in inventory systems is the ABC classification system.

The utilization of classification techniques for spare part management tool represents a widely used approach in industrial world. ABC inventory classification is a frequently used procurement planning and control method that was designed to achieve balance between two conflicting economic forces: inventory costs and critical stock-outs of spare parts.

2.2 ABC Analysis

ABC analysis was discovered by Pareto, an Italian economist, approximately 100 years ago. He discovered that a small percentage of a population always has the greatest effect. Because of its easy-to implement nature and remarkable effectiveness in many inventory systems, this approach is widely used in practice (Chen et.al. 2006). Pareto's law was further expanded to the ABC classification and is summarized below. Viale (1996) has explained ABC analysis like this: "When considering how to apply this tool to establish inventor y levels, consider the following: From a practical standpoint, ask yourself: "Which products (and which customers) generate 80 percent of the revenue?" Answer: Approximately 20 percent of the products and customers generate 80 percent of the revenue. From the Table 2.1, detailed explanation on the A, B and C grouping can be viewed:

Table 2.1: ABC analysis grouping

20% of Customers, products, or parts=	80% of the company's revenue and inventory investment	These are called "A" customers "A" products "A" parts
30% of customers, products, or parts=	15% of the company's revenue and inventory investment	These are called "B" customers "B" products "B" parts
50% of customers, products, or parts=	5% of the company's revenue and inventory investments	These are called "C" customers "C" products "C" parts

The fluctuation in demand for the B and C products causes most of the product mix problems, the changes on the shop floor (capacity) and the changes in the supplier due dates Viale (1996).

Consider doing the following:

- Reduce the forecast error by improving the forecast model.
- Use the standard deviation formula A customers and A and B parts.

- For C parts, give a predetermined number of days' supply and allow the production floor to build this during the beginning of each quarter.

The standard deviation of forecast error tool can also be used to determine

- ✓ The amount of extra (safety) lead time needed to ensure on-time delivery.
- ✓ The amount of work to release to the shop floor to make sure machines do not run out of work. That is, how much "queue" buffer is needed in back of the bottleneck work center?
- ✓ The number of extra (buffer) pieces to start on the initial machining operation to ensure a particular yield after the final operation (for example, scrap allowance).
- ✓ The amount of machine downtime (safety) to allow for in planning utilization of available capacity (Viale, 1996).

Classical ABC technique differentiates the inventory into three classes: A – very important, the inventory small in volume but large in cost, B – moderately important, inventory intermediate in volume and cost; and last, C – least important, that is inventory is large in volume and however low in cost usage. "A" items have the highest value. These are relatively few items (15– 20 percent) whose value accounts for 75– 80 percent of the total value of the inventory. As a general rule, 20 percent of the items constitute 80 percent of the annual requirements. All "A" items are counted monthly.

"B" items have medium value. These are a larger number in the middle of the list, usually about 30– 40 percent of the items, accounting for about 15 percent of the value. All "B" items are counted quarterly.

"C" items have low value. These are the bulk of the inventory, usually about 40– 50 percent of the items, whose total inventory value is almost negligible, accounting for only 5–10 percent of the value. All "C" items are counted annually. Many times these physical counts are based on estimates (Viale,1996).

The amount of time, effort, money and other resources spent on inventory planning and control needs to be in accordance with importance of each item. So, the purpose of this technique is to provide all inventory items with appropriate levels of control.

Traditional ABC classification is based on only one criterion. Generally, the only criterion is taken as annual cost usage or average unit cost, and sometimes it may be the number of orders and purchasing conditions (Simunovic et al., 2008). However, there may be other important criteria that represent various important considerations for management and it is a disadvantage of the classical ABC classification that it takes into account only one criteria and leaves out many important characteristics according to which inventory may differentiate. Thus multi-criteria techniques have been developed to enable efficient decision making which will include crucial criteria like criticality of an item, stock-out penalty, lead time of supply, part criticality, availability, average unit price, the scarcity, the rate of obsolescence, substitutability, etc. (Simunovic et al., 2008). Therefore, it has been usually recognized that traditional ABC analysis may not be able to provide a good classification of inventory in practice (Guvenir & Erel, 1998; Huiskonen, 2001; Partovi & Anandarajan, 2002). There are many instances when other criteria become important in deciding the importance of an inventory item. This problem becomes a multi-criteria inventory classification that has been studied by some researchers in the past. In general, complex computational tools or procedures are needed for multi-criteria ABC classification (Chu et. al., 2008).

2.3 Multi Criteria ABC Analysis

A typical company holds multitude of items in inventory and only a small fraction deserves the close attention. The decision maker's task is determining the inspection levels on each inventory item; thus, first and foremost, assigning the priorities for each item and realizing a sensible classification . As Sharaf and Helmy (2001) have noted, thousands of items may be potentially held in inventory by a typical organization, however only a small portion of them deserve management's close attention and accurate control.

To overcome the limitations of the traditional classification analysis, many researchers concentrated on incorporating multiple criteria judgments into the inventory classification procedure. Flores et. al. (1985) have transformed the classical ABC analysis for the inventory classification, considering another important criterion to so called bi-criteria inventory classification. They used traditional ABC analysis by classifying the inventory by first criterion and then by second criterion. Here the disadvantage of this method was that they took weights of each criterion to be equal. Flores et al. (1985) suggested using multiple criteria like lead time, criticality, commonality, obsolescence and substitutability.

Since the multi-criteria classification idea was introduced, several authors reported ondifferent approaches to the problem. Cluster analysis method (Cohen and Ernst, 1988), a classification expert system (Petrovic & Petrovic, 1992), joint criteria matrix approach (Flores, Olson, & Dorai, 1992), a heuristic procedure employing genetic algorithms (Guvenir & Erel, 1998), and a weighted linear optimization methodology (Ramanathan, 2006) have been successfully applied in the past years.

Other developments of multi-criteria of ABC analysis has been done by Ernst and Cohen (1990). They proposed technique based on statistical clustering that uses a full range of operationally significant attributes. Their technique can accommodate a large number of combinations of attributes which is mostly important for strategic and operational purposes. Yet, their method required utilization of large amount of data, the use of factor analysis and clustering procedure which may end up to be impractical for typical stockroom environment. Moreover, the clusters should be measured again in order to classify new stock items and there is also chance that each time a new item is added, the clusters previously classified will be re-classified again. This model may be too sophisticated and might disturb the inventory control procedure.

3. AN ARTIFICIAL NEURAL NETWORK

The simplest definition for artificial neural network (ANN) is that, it is a mathematical or a computational model which is inspired by the structure and/or functional aspects of biological neural networks. A neural network is an adaptive system that changes its structure based on external or internal information that flows through its network and generally is learning to generalize or classify or organize data in the learning phase. Modern neural networks are non-linear statistical data modeling tools. They are used to model complex relationships between inputs and outputs or to find correlations in the data. Unlike traditional computational models, ANN model has brain like structure and functionality of the brain cells.

Researches on artificial neural networks have been inspired by the recognition of how human brain computes in an entirely different way from the computer functioning. The structure of the brain is highly complex, nonlinear and parallel, like information processing systems in computer. It has the ability to do certain computations such as pattern recognition, perception and motor control in a very short term. For instance, it usually accomplishes perceptual recognition tasks such as memorizing a familiar face in an unknown scene in 0,1-0,2 milliseconds. On the other hand, computer will spend much more time like several minutes or even hours to perform less difficult tasks. The reason why human brain has been interesting topic for researchers is that its structure and functioning gives thorough insight to how a computer can be developed to perform functions usually completed with human intelligence such as reasoning, learning and self improvement (Haykin, 1999). Haykin (1999) explains that a brain has great structure and ability to build up its own rules through time using that is known as experience.

Human brain undergoes a dramatic development within two years after birth and continues to evolve after that. It has capability to develop its structural elements known as *neurons* that allow the building nervous system to adjust to its surroundings. Neurons play same role, which is information processing unit, in both human brain and neural networks. In other words, it can be explained that the neural network are designed like a machine to perform particular tasks like human brain (Haykin, 1999). To specify the above statements made by Haykin (1999), his definition for artificial neural network is that "it is a massive parallel distributed processor for storing experiential knowledge and making it available for use. It resembles the brain in two respects: knowledge is acquired by the network from its environment through a learning process and secondly interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge".

The model of learning process is known as learning algorithm and its function is to modify the synaptic weights of the network in an orderly fashion to attain a desired design objective (Haykin, 1999).

3.1 Application Areas of Artificial Neural Networks

The application area of artificial neural networks can be extended to practically every kind of problem where regression based models and statistical models can't be applied or do not give desired results. Neural networks give the best results in building nonlinear models.

- *Classification:D*iscrimination of elements based on similarities
- *Pattern recognition:* Recognition of an output according to given input data.
- *Modeling:*Generalization of few examples
- *Clustering:*Grouping the elements according to common attributes or features.
- *Forecasting:*Prediction of future outputs by analysis of current and past data.
- *Optimization:*Satisfying multiple conflicting goals and constraints.
- *Finance:*ANN's can be used for funds investment, credit analysis, risk insurance, options and futures prediction, trend prediction, stock investment optimization , cash flow forecasting, etc.
- Medicine: ANN' can be used medical diagnosis, classification of deceases, genetic mapping, blood mapping, treatment cost estimation, etc.
- Science: ANN's can be used for modeling complex problems, nonlinear problems, physical system modeling, chemical compound identification, botanical classification, odor analysis and identification, etc.
- Sales and Marketing: ANN's can be used for sales forecasting, targeted marketing, service usage forecasting, etc.

- Energy: ANN’s can be used for electrical load forecasting, energy demand forecasting, power control systems, etc.
- Production: ANN’s can be used for quality control, process control, temperature and force predictions, production routing optimization, etc.

3.2 Benefits of Neural Networks

The ability to learn and generalize is the main advantage of neural networks. The ability of generalization makes it possible for neural network to give a reasonable output for inputs that have not been encountered before. Below are the properties and capabilities of neural networks reported by Haykin (1999):

- *Nonlinearity:* the nonlinear structure of neural network is a very important property in cases where a nonlinear relation is required. A regression based methods are hard to control and implement where a cases of linear relations takes place.
- *Input-Output Mapping:* A popular way of training a neural network is to use supervised training. Supervised training teaches a neural network to produce the ideal output with a given inputs. Every training iteration calculates how close the actual output is to the ideal output. The closeness to ideal output is expressed as an error percent. Each iteration modifies the internal weights of the neural network to get the error rate as low as possible. The training is repeated until there no further significant changes in error rate. The neural network learns from training and constructs an input-output mapping for a given problem.
- *Adaptivity:* Neural Networks have a capability to adapt their synaptic weights to changes in surrounding environment. A neural network trained to operate in a specific environment can easily be adapted to minor changes environmental conditions.
- *Evidential Response:* In pattern classification a neural network can be designed to provide information not only about the output but also about the confidence in the decision made. The confidence in decision making can be used to eliminate ambiguous patterns and improve classification performance of the network.

- *Contextual Information:* Knowledge is presented in by the structure and activation of a neural network. Every neuron in the network is potentially affected by global activity of all other neurons in the network. Consequently, contextual information is dealt with naturally by a neural network.
- *Fault Tolerance:* A neural network which is used in hardware form has the ability to be intrinsically fault tolerant or capable of robust computation, in the sense that its efficiency deteriorates gracefully undesirable operating conditions. That is if the neuron or its connection links are damaged, the recall of the saved patters in worsened in quality. On the other hand, because of distributed nature of information saved through the network, the adverse condition has to be extensive before the overall response of the network is damaged. Therefore, a neural network shows a graceful degradation rather than catastrophic failure.
- *VLSI Implementability:* A neural network has the potential to compute certain tasks very quickly because of its massively parallel nature. This characteristic enables neural network well suited for implementation using *very-large-scale-integrated (VLSI)* technology. One of the beneficial virtues of VLSI is that it provides means of recognizing very complex behavior in a highly hierarchical fashion.
- *Uniformity of Analysis and Design:* Neural networks have universality as information processors, because the same notation is used in all domains involving the application of neural networks. This characteristic reveals itself in several ways; at first neuron represents an ingredient common to all neural networks. Then, this commonality makes it possible to share theories and learning algorithms in different applications of neural networks. And finally, modular networks can be designed a seamless integration of modules.
- *Neurobiological Analogy:* the architecture of neural network is motivated by analogy of brain, which is a living proof that fault tolerant parallel processing is not only physically possible but also fast and powerful (Haykin,1999).

3.3 Artificial Neural Network Structure: An Explanation Based on Biologic Nervous System

Haykin (1999) has explained the biologic nervous system in a three staged system as can be seen from diagram in Figure 3.1. The brain is in the center of neural system which represents the neural net. It continuously receives information, perceives it, and gives appropriate decision. There are two arrow nets in the figure. The arrows pointing from left to right indicate the forward transmission of the information-bearing signals through the system. Moreover, there is the feedback process which is transmitted from right to the left by arrows. The receptors convert stimuli from the human body or the external environment into electrical impulses that conveys information to the neural net (brain). The effector convert electrical impulses generated by the neural net into discernible responses as system outputs.

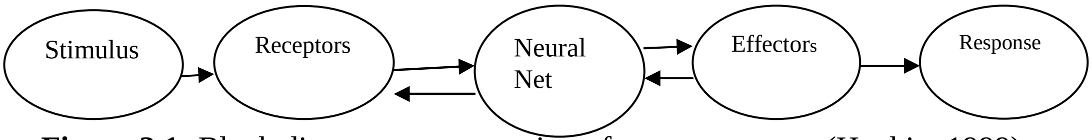
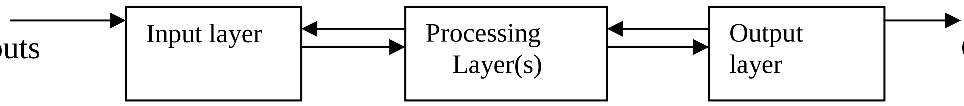


Figure 3.1: Block diagram representation of nervous system (Haykin, 1999).

There are several types of ANN architecture in literature; however a general ANN architecture can be shown as in Figure 3.2. There is at least one input element in the input layer. In this layer input elements generate the same values of input without any processes (Tosun, 2007). There is at least one output element and in spite of input elements in output elements, there is a process that generates output (Tosun, 2007). Processing layers usually called black box because understanding each processing elements behaviors is a very difficult task. This layer(s) and functions used in these layers can change according to the ANN type.

Inputs



Output

Figure 3.2: Block diagram representation of ann architecture.

Neurons are the structural constituents of biological nervous system. The structure of neuron and its parts are demonstrated in Figure 3.4. With dendrites neuron collects the stimulus of previous neurons. In cell body this stimulus is evaluated and an output stimulus is generated. This stimulus is send to following neurons with the axon. At this point axon can be divided into many parts and send the stimulus to all connected neuron dendrites. The axon- dendrite connection areas are called synapse. The most common type of synapse is called chemical synapse, which operates as follows; a presynaptic process liberates a transmitter substance that diffuses across the synaptic junction between neurons and then acts on the postsynaptic process

(Haykin, 1999). In an adult brain, plasticity, which permits the developing nervous system to adapt to its surrounding environment, may be accounted for by two mechanisms: the creation of new synaptic connections between neurons, and the modification of existing synapses (Haykin, 1999).

Partovi and Anandarajan (2002) defines ANN as follows: “ The ANN in general can be characterized in the following components: a set of nodes and connection between nodes. The nodes can be seen as computational units. They receive inputs and process them to obtain an output. This process might be very simple computation like summing or it might also be complex as to contain another network. The interaction of nodes through the connections leads to a global behavior of the network, which cannot be observed in the elements of the network. This global behavior is said to be emergent that is the abilities of the network supersede the ones of its elements, making networks a very powerful tool. In ANNs, the network sees the nodes as ‘artificial neurons’. An artificial neuron is a computational model inspired by the natural neurons. Natural neurons receive signals through synapses located on the dendrites or the membrane of the neuron. When the signals received are strong enough, the neuron is activated and emits a signal through the axon. This signal might be sent to another synapse and might activate other neuron”. A typical neuron structure described by Partovi and Anandarajan (2002) can be seen in Figure 3.3 below:

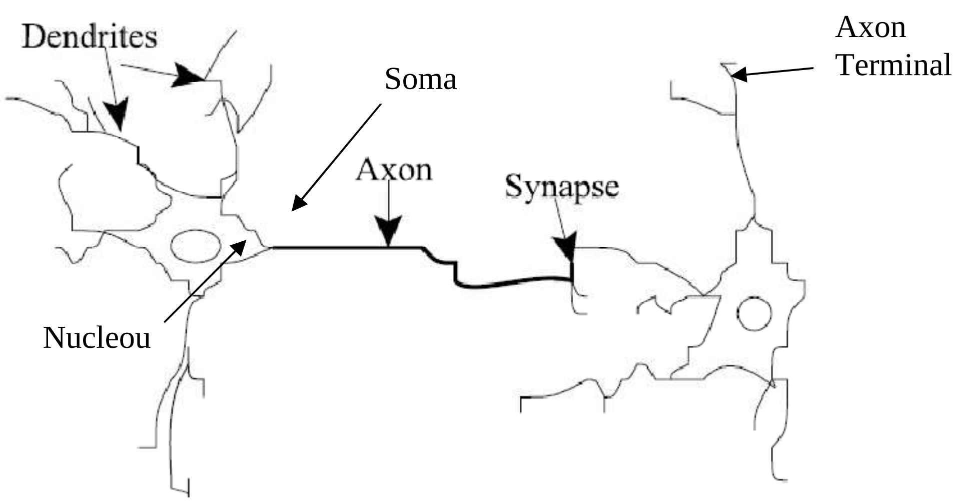


Figure 3.3: Structure of a typical neuron (Partovi and Anandarajan, 2002).

The complexity of the real neuron is highly abstracted when modeling artificial neurons. These basically consist of inputs (like synapses) which are multiplied by weights (strength of the respective signals), and then computed by the mathematical function which determines the activation of the neuron. Another function, which may be the identity, computes the output of the artificial neuron. ANNs combines artificial neurons in order to process information.

A processing element can be shown as in Figure 3.4. We can express the following similarities between processing elements (PEs) or neurons of ANN and neurons of nervous systems. The weights in processing elements are synapses. The summing junction is the dendrites that collect the inputs. Activation function is the cell body that processes the stimulus. And the output element is the axon that transports the output to the other neurons.

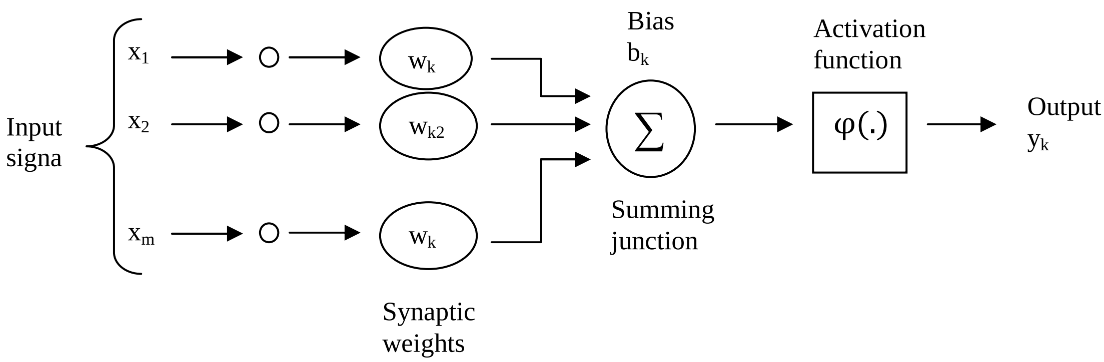


Figure 3.4: Model of a typical processing element (Haykin, 1999).

The neuron model shown in Figure 3.4 also includes an externally applied bias, denoted by b_k.. The bias b_k has the effect of increasing or decreasing the net input of the activation function depending on whether it is positive or negative (Haykin, 1999). Mathematically we can describe a neuron k by the following pair of equations 3.1 and 3.2 :

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \qquad \textbf{(3.1)}$$

$$y_k = \varphi(u_k + b_k) \qquad \textbf{(3.2)}$$

Here, $x_1, x_2, x_3, \dots x_m$ are the input signals; $w_{k1}, w_{k2}, \dots w_{km}$ are synaptic weights of neuron k; u_k is the linear combiner output due to the input signals; b_k is the bias; $\varphi(.)$ is the activation function; and $y_k$is the output signal of the neuron. The use of bias b_k has the effect of applying the affine transformation to the output u_k of the linear combiner in the model (Figure 3.4), as shown by equation 3.3 below:

$$v_k = u_k + b_k \qquad \textbf{(3.3)}$$

The higher the weight of an artificial neuron, the stronger the input will be, which is multiplied by it. Depending on the weights, the computation of the neuron will be different. By adjusting the weights of an artificial neuron we can obtain the output we want for specific inputs. But if we have an ANN of thousand of neurons, their weights of the ANN may be adjusted only by algorithms in order to obtain desired output from the network. This adjusting process is called learning or training of the network. The number of types of ANNs is very high and there are more than hundred of different models considered as ANNs. Since the functions of ANNs is to process information, they are used in many fields like study behavior and control in animals and machines, also as well as engineering purposes like pattern recognition, forecasting, data compression and classification. For example, Partovi and Anandarajan (2002) had used two types of learning methods to examine classification accuracy of ANN as an aid to facilitate the decision making process of classifying inventory items. More specifically, two learning methods used by Partovi and Anandarajan (2002) are back propagation (BP) and genetic algorithms. Also, Wei et al. (1997) had used ANN which takes into account quantitative and qualitative data and nonlinear problem with multiple inputs in supplier evaluation process.

3.4 Learning Process in Artificial Neural Networks

Learning in the conext of the neural network has been defined by Haykin (1999) as; learning is a process by which the free parameters of a neural network are adjusted through a process of stimulation by the environment in which the network is embedded and the type of learning is determined by the manner in which the parameter changes take place. Haykin (1999) also mentioned that this definiton of learing process implies the following sequence of events:

- The neural network is stimulated by an environment.
- The neural network undergoes changes in its free parameters as a result of the changes that have occurred in its internal structure.

Çınar (2007) has defined learing process in ANN as a kind of price-penalty system. If the output of artificial neural network and the desired output are in the same direction, the weights of ANN are strengthened. If the output of ANN and desired output are in opposite direction, the weights are weakened to teach ANN to respond differently (Çınar, 2007).

In practice, artificial neural networks that have only one hidden layer may easily learn the process with few data and continuous functions (Çınar, 2007). Regarding the hidden layer, Çınar (2007) has mentioned that hidden layer is only required if the function is not continuous for some points. For many ANN model in literature, researchers reported that only one hidden layer is enough and that the second hidden layer slows the learning process (Çınar, 2007).

Feedforward network indicates that data flow from input layer to output layer. Output of each layer is the input of the next layer and is the functions of its inputs (Çınar, 2007). Activation fucntion computes the output value of each neuron. In complex models it is important to utilize nonlinear activation functions. The shape of the activation function affects the learning performance of the neural network; however, it does not influence the overall performance of the network (Çınar, 2007).

The learning of network can be in two ways: online or batch learning. First one, in online learning data used one by one for learning. In second way, batch learning, the whole data is used at once for learning. Apaydin (2004) mentioned that in batch learning, learing (changes in free parameters of neural network) accumulated over the all patterns and the change is made once after a complete pass over the whole training set is run. A complete pass over all the patterns is called an epoch (Apaydin, 2004).

There are three kinds of learning process: supervised learning, unsupervised learning and hybrid learing.

3.4.1 Supervised learning

It is know as learning under supervison of teacher, since in conceptual terms a teacher with the knowledge about environment teaches the neural network with that knowledge being presented by a set of input-output examples (Haykin, 1999). Apaydin (2004) gave examples of supervised learning as regression and classification problems.

The following example gives better explanation of the supervised learning: let the teacher and the neural network both be exposed to training vector drawn from the environment. By the virtue of built-in knowledge, the teacher is able to provide the neural network with desired responce for that training vector (indeed, the desired respense represents the optimum action to be performed by the neural network). The parameters of the network are modified according to the traininf vector and the signal error. The error signal is computed by the difference between the desired output and the actual output of the network. This adjustment is carried out iteratively in a step by step fashion with objective to eventually make the neural network emulate the teacher. The emulation process is presumed to be optimum in some statistical sense.

3.4.2 Unsupervised learning

It is also called learning without a teacher, because in spite of supervised learning only inputs of the problem are known. In unsupervised learning the goal is to determine the formation along the inputs (Çınar, 2007). Input space has a pattern and if analyzed it can be deduced which input are more repeated and which are less repeated. This is called density estimation in statistics (Alpaydın, 2004). When the patterns are discovered learning is completed; a new input’s cluster can be determined (Haykin, 1999).

One method for density estimation is clustering where the aim is to find clusters or groupings of input. The following example of clustering is given by Alpaydın (2004): in the case of a company with a data of past customers. The data contains the demographic information as well as the past transactions with the company, and the company may want to see the distribution of the profile of tis customers, to see what type of customers frequently occur. The author noticed that in such a case, a clustering model allocates customer similar in their attributes to the same group, providing the company with natural groupings of its customer. Alpaydın (2004) also added that once such groups are found, the company may decide strategies (for example, specific services and products to different groups).

3.4.3 Reinforcement learning

In some applications, the output of the system is a sequence of actions. In such a case, a single action is not important; the policy, which is the sequence of correct actions to reach the goal, is important. In this case, neural network should be able to assess the goodness of policies and learn from past good action sequences to be able to generate a policy. Such learning methods are called reinforcement learning (Alpaydın, 2004).

In reinforcement learning, like unsupervised learning, certain outputs are not used to train the neural network. But the desired outputs are defined as good output or bad output and then used to train the neural network (Çınar, 2007). But defining good or bad outputs are somehow similar to supervised learning.

Chess game can be an example of this type of learning because the rules of the game are limited but in many situations there is large number of possible moves (Aplaydın, 2004). In such a case one move is not important, the series of moves are important to win the game. For further information about unsupervised and reinforcement learning please look at Haykin (1999) and Schalkoff (1997).

In this study, supervised learning methodology is going to be described in detail because it will be utilized in the classification of the spare parts inventory based on ABC analysis.

3.5 Multi-Layer Perception (MLP) and Back Propagation Algorithms

3.5.1 Multi layer perception

The multi-layer perception is one the the most used neural network topologies. Principle et al. (1999) have indicated that, the MLP is capable of approximating arbitrary functions. MLP plays important role in the study of nonlinear dynamics and other function mapping problems. A common multi layer perception architecture can be seen in Figure 3.5:

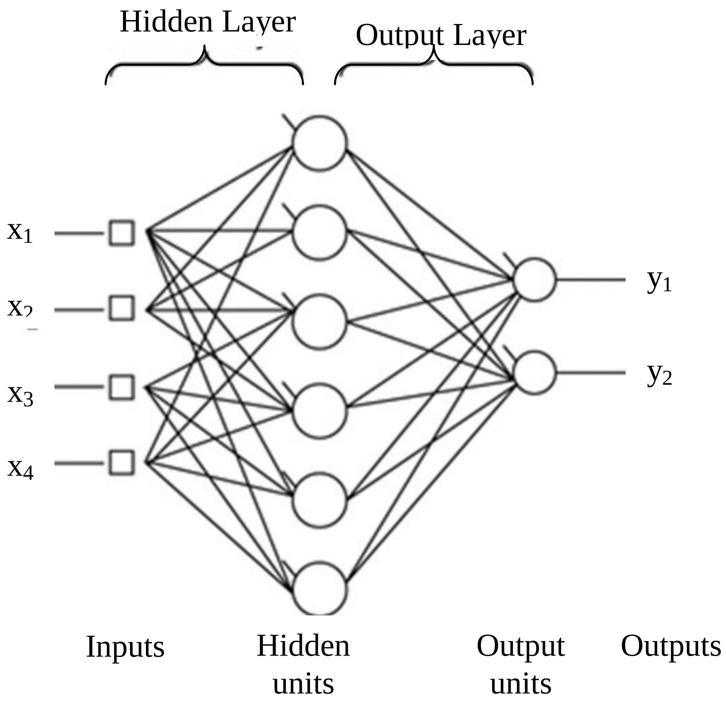


Figure 3.5: Multi-layer perception model.

The multilayer perception network given above has three layers with one hidden layer. x_1,x_2,x_3,x_4 are input values and $y_1$and y_2 are the outputs of the network.

The are two crucial aspects of the multi-layer perception. First one is that nonlinear processing elements (PEs) that have a nonlinearity which has to be smooth (the logistic function and the hyperbolic tangent are the most widely used). Second aspect is that MLP has massive interconnectivity that is any element of a given layer feeds all elements of the next layer (Principle et al., 1999).

Principle et al. (1999) have indicated that multi-layer perception are generally trained by backpropagation algorithms.

3.5.2 Backpropagation algorithm

The Backpropagation algorithm propagates the error through the network and allows adjustions of the hidden nodes, i.e. processing elements. The multi-layer perception is trained by error correction learning, so it means that desired for the system should be known.

Learning by error correction is done in the following way:

At PE i at the iteration n, the system generates response $y_i(n)$ and the difference between system response, $y_i(n)$ and the desired response $d_i(n)$ for a given pattern gives us the instantaneous error, $\varepsilon_i(n)$, which is defined by equation 3.4:

$$\varepsilon_i(n) = d_i(n) - y_i(n) \tag{3.4}$$

According to the *gradient descent learning theory*, each weight in the network can be adjusted by modifying the present value of the weight with the terms that is proportional to the present input and error at the weight, that is defined as in equation 3.5:

$$w_{ij}(n+1) = w_{ij}(n) + \eta\delta_i(n)x_i(n) \tag{3.5}$$

where $\delta_i(n)$ is the local error and can be easily computer from the system error, ε_i, at the output PE or can be computed as the weighted sum of errors at the internal PEs. The constant η is the step size and called the learning rate. This procedure is defined the backpropagation algorithms.

The backpropagation algorithm computes the sensitivity of the cost function with respect to each weight in the network, and updates each weight proportional to the sensitivity. This procedure is very easy because it can be implemented with local information and requires just a few multiplications per weight. Due to the fact that gradient descent procedure uses only local information, it can be stuckt at local minima. Principle et al. (1999) have indicated that the procedure is inherently noisy because usage of poor estimate of gradient may cause slow convergence.

Momentum learning is the improvement to the straight gradient descent in the sense that a memory term (the past increment to the weight) is used to speed up and stabilize convergence. By momentum learning, the weights are updated as following in equation 3.6:

$$w_{ij}(n+1) = w_{ij}(n) + \eta\delta_i(n)x_j(n) + \alpha(w_{ij}(n) - w_{ij}(n-1) \tag{3.6}$$

Where α is the momentum of learning. The momentum rate is taken to be between range 0.1 to 0.9 in most literatures (Çelebi and Bayraktar, 2008; Ghate and Dudul, 2010).

Weights in the network during the training process can be updated in two ways, on-line and batch training. In on-line training, the weights are updated continuously each time as each pattern is input to the network. On the other hand, in batch learning we present all the patterns in the input file (an epoch), accumulate the weight usdates, and then update the weights with the average wieght update. Principle et al. (1999) have noted that online learning and batch learning are theoretically equivalent, however, on-line learning is superior in tough problems where many similar input-output pairs exits. Moreover, Haykin (1999) states that it is less likely that backpropagation algorithm to get stuck in a local mnima with the on-line mode of training.

Backpropagaion starts by loading an initial value for each weight (normally a small value) and proceeding until some stopping criterion is met. The three most common stopping criteria are to cap the number of iterations, to threshold the output mean square error, or to use cross validations. The cross validation is the most powerful of the three types of criteria because it stops the training at the point where the best generalization is obtained that is the performance in the test set. The training data is divided into two sets: training set and cross validation set. The cross validation set is used to see how the trained network is doing, for example, every 100 training epochs cross validation set test the network training performance. When the performance starts to degrade in the validation set, the training should be stopped (Alpaydın, 2004; Haykin, 1999; Principle et al., 1999).

The fundamental of any iterative training process is the measurement of the learning process. The learning curve can be good estimeate of the training process. Learning curve shows the change of mean squared error with the number of epochs. The difficulty of the task and how to control learning can be judged from the learning curve. If the learning curve is straight, it means that learning rate should be increased. Meanwhile, if the learning curve is oscillating up and down, it means that step size shoulf be decreased. When the learning curve stabilizes at an error level that is not acceptable, the network topology, like altering number of node in hidden layer, or number of hidden layers, or the training prosedure like other more sophisticated search techniques.

3.6 Types of Activation Functions

Activation function, shown as $\varphi(v)$, defines the output of a neuron in terms of the induced local field *v*. Here are three basic types of activation functions (Haykin, 1999):

1. Threshold Function: For this type of the activation function in equation 3.7, described in Figure 3.6,

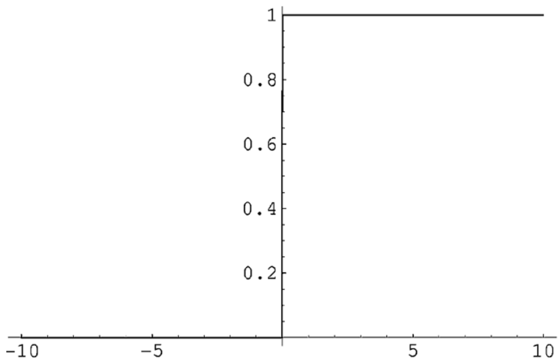


Figure 3.6: Threshold function

$$\varphi(v) = \begin{cases} 1 \; if v \geq 0 \\ 0 \; if v < 0 \end{cases} \quad \textbf{(3.7)}$$

is used. In engineering literature, this form of threshold function is commonly referred to as Heaviside function. Correspondingly, the output of neuron *k* employing such a threshold function is expressed as in equation 3.8:

$$y_k = \begin{cases} 1 \; if v_k \geq 0 \\ 0 \; if v_k < 0 \end{cases} \quad \textbf{(3.8)}$$

where v_k is the induced local field of the neuron; that is as in equation 3.9:

$$v_k = \sum_{j=1}^{m} w_{kj} x_j + b_k \quad \textbf{(3.9)}$$

Such a neuron is called in the literature as the MCCulloch-Pitts model in recognition of the leading work done by McCulloch and Pitts. Within this model, the output of the neuron takes on the value of 1 if the induced local field of that neuron is non negative or 0 if negative. This statement describes the all-or-none property of the McClloch-Pits model.

2. Piecewise-Linear Function: For the piecewise-linear function described in Figure 3.7, the following equation 3.10 is used:

$$\varphi(v) = \begin{cases} 1, & v \geq +\frac{1}{2} \\ v + \frac{1}{2}, & +\frac{1}{2} > v > -\frac{1}{2} \\ 0, & v \leq \frac{1}{2} \end{cases} \tag{3.10}$$

Where the amplification factor inside the linear region of operation is considered to be unity. This form of an activation function may be considered as an approximation to a nonlinear amplifier. The following two situations may be viewed as special forms of the piecewise-linear function (Haykin,1999):

a. A linear combiner arises if the linear region of operation is maintained without running into saturation.
b. Piecewise-linear functions if the amplifier factor of the linear region is made infinitely large.

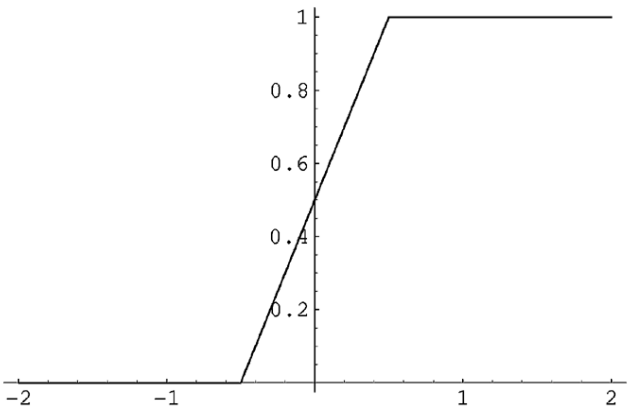


Figure 3.7: Piecewise-linear function

3. Sigmoid Function: This function, which has s-shaped figure, is by far the most common forms of activation function used in the architecture of artificial neural networks. Sigmoid function can be seen in Figure 3.8 below:

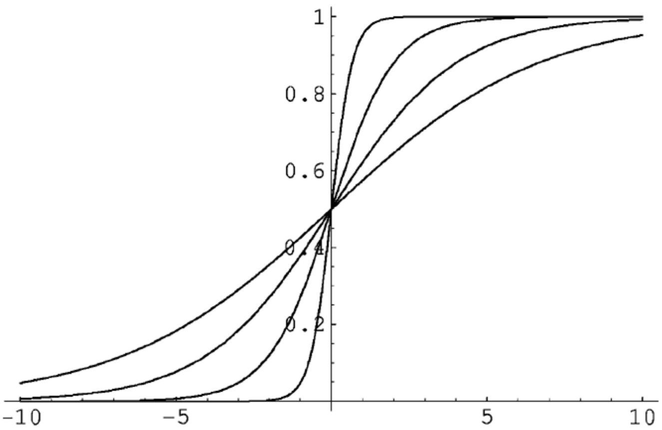


Figure 3.8: Sigmoid function

Sigmoid function is defined as a strictly increasing function that exhibits a graceful balance between linear and nonlinear behavior (Haykin, 1999). An example of sigmoid function is the logistic function, defined by equation 3.10 below:

$$\varphi(v) = \frac{1}{1+\exp(-av)} \qquad \textbf{(3.11)}$$

Where *a* is the slope parameter of the sigmoid function. By varying the parameter *a*, we obtain sigmoid functions of different slopes, as ilustrated in Figure 3.8. Actually, the slope at the origin equals $\frac{a}{4}$. In the limit, as the slope parameter approached infinity, the sigmid function becomes simply thresholf function taking the value of 0 or1, a sigmoid function assumes continuous range of value from 0 to 1. Note also that the sigmoid function is differentiable, whereas the threshold function is not (Haykin, 1999).

The activation functions defined above the range from 0 to +1. It is sometimes desirable to have the activation function with range from -1 to +1, in which case the activation function takes an antisymmetric form with respect to the origin. In other words,the activation function is an odd function of the induced local field. Specifically, the threshold function in Equation 3.11 is now as

$$\varphi(v) = \begin{cases} 1 & if v \geq 0 \\ 0 & if v = 0 \\ -1 & if v < 0 \end{cases} \qquad \textbf{(3.12)}$$

Which is known as the signum function. For the corresponding form of a sigmoid function we may use the hyperbolic tangent function, defined by equation 3.12:

$$\varphi(v) = \tanh(v) \qquad \textbf{(3.13)}$$

Letting the activation function of the sigmoid type to assume negative values as above have some analytic benefits.

4. LITERATURE REVIEW

Inventory control also known as stock control and inventory management is one of the important techniques of operational management and it plays important role in the company management. A well designed approach to inventory management might have very important influence on the company competitiveness. To achieve their goals companies should apply optimization and multi-criteria decision making methods because of the huge amount of inventory items which needs a great attention to classify these items into different groups. That is different groups might require application of different management tools and policies. ABC inventory classification based on the Pareto principle takes into account only one criterion.

There are many contributions to multi-criteria inventory classification in the literature. Multiple criteria classification necessitates complex decision making tools (Chen, 2011). For inventory control with many different types of inventory, multi-criteria decision making methods have been developed for ABC classification. In literature, various methods include fuzzy analytical hierarchy process (AHP), data envelopment analysis (DEA), case based distance model, the weighted linear optimization, the joint criteria matrix, the clustering procedure, the particle swarm optimization method, principle component analysis, genetic algorithms, artificial neural network approaches and etc.

4.1 Inventory Classification Techniques

Inventory classification techniques have been developed for the various industries where large companies accommodate large quantities of inventory items. The literature on inventory classification includes different methodologies for multi-criteria inventory classification.

The literatur review hase been done on spare parts inventory classification based on multi criteria decision makng techniques and short summary of the main research studies can be seen in Table 4.1.

Table 4.1: Academic Researches about Multi Criteria Inventory Classification.

#	Citations	Technique(s) used	Evaluation criteria	Application Area
1	Braglia et al. (2004)	Reliability Centered Maintenance (RCM) with AHP.	Main criteria:spare part plant criticality, spare supply characteristics, inventory problems and usage rate Sub criteria: quality problems, lead time, internal repair, price, space required, obsolescence, deterioration problems, production loss, redundancies, etc.	Spare parts classification in paper operating company
2	Cakir and Canbolat (2008)	Fuzzy AHP	Price/cost, annual demand, blockage effect in case of stock out, availability of substitute material, lead time	Designing a decision support system in a small electrical appliances company
3	Simunovic et al. (2008)	a. Traditional ABC b. Multi Criteria ABC with AHP	a1. annual cost usage b1. annual cost usage, criticality, lead time1, lead time 2	Classification of spare parts in agricultural machine.
4	Çelebi et al. (2008)	Multi Criteria ABC with AHP and DEA	Main criteria Criticality, value usage, unit cost, lead time Criticality sub criteria: penalty cost, substitutability, commonality	Classification of maintenance spare parts in light railway systems

Table 4.1: Academic Researches about Multi Criteria Inventory classification (contd.).

#	Citations	Technique(s) used	Evaluation criteria
5	Sönmeztürk et al. (2008)	Multi Criteria ABC with Fuzzy AHP	Main criteria Criticality, value usage, unit cost, lead time Criticality sub criteria: penalty cost, substitutability, commonality
6	Hadi-Venchen and Mohamanghasemi (2011)	MC ABC with Fuzzy AHP and DEA	Annual dollar usage, limitation of warehouse usage, average lot cost, lead time
7	Ramanathan (2006)	Weighted linear optimization based on DEA	Average unit cost, annual dollar usage, critical factor, lead time
8	Ng (2007)	Ng model	
9	Partovi and Anandarajan	Multi Criteria ABC a. Genetic Algorithm b. Artificial Neural network	Unit price, ordering cost, demand range, lead time
10	Li and Kuo (2008)	Enhanced Fuzzy Neural Network	Part factor, demand factor, time factor, sale factor,associated factor
11	Chen et al. (2009)	Back-propagation neural network (MBPN) and moving fuzzy neuron network (MFNN)	Large variation of demand, long purchasing lead time, necessity to the operation machine, cost

4.2 Multi Criteria Decision Making (MCDM)

Multiple criteria decision making methods are separated into two basic groups: multiple attribute decision making methods and multi objective decision making methods. The first method is used to select the best alternative among the many feasible alternatives where each alternative is defined by some quantitative and/or qualitative criteria or attribute. The second method deals with two or more objective functions subjected to some constraints.

Multi criteria decision aid (MCDA) theory includes multi criteria ABC classification. According to Roy (1996), multi criteria decision is the aid to assist a single decision maker to choose, rank or sort alternatives within a finite set according to two or more criteria. MCDA begins with process of defining objectives, arranging them into criteria, identifying all possible alternatives and measuring the consequences of each alternative on each criterion. A consequence is a direct measurement of the success of an alternative against a criterion and usually is an objective physical measurement that includes no preferential information (Chen et al., 2008).

The main difficulty with many existing MCDA methods lies in the acquisition of the decision maker's preference information in the form of values or weights. Case based reasoning is an approach to finding preferential information using cases selected by the decision maker. Chen et. al. (2008) have proposed case-based distance model for multiple criteria decision aid problem which allows usage of any finite number of criteria in ABC analysis. In this approach criterion weights and sorting thresholds are generated mathematically based on the decision maker's assessment of the case set and thus, problems related to direct acquisition of preferences information are avoided. In their evaluation, the criterion includes annual unit cost, annual dollar usage, critical factor (criticality) and lead time.

4.3 Analytical Hierarchy Process (AHP)

Analytical hierarchy process is the most widely used techniques in almost all sectors such as economy, traffic, agriculture, information technology and many others. AHP was developed by Thomas Saaty (1980) and it has a tree like structure that contain main goal at the top of the hierarchy which is first level, then followed by criteria and sub-criteria levels. At the bottom of the hierarchy tree various alternatives are present that are compared according to above criteria within the tree levels. Analytical hierarchy process has been widely used for classification of inventory in several fields. Partovi and Burton (1993) applied the analytical hierarchy process to inventory classification in order to include both quantitative and qualitative evaluation criteria.

Braglia et al. (2004), has proposed a multi attribute decision making classification model as a tool for spare parts inventory management in a paper operating company by utilizing two different techniques: the reliability centered maintenance (RCM) and the AHP. The Reliability centered maintenance analysis defines a decision diagram which guides decision maker toward the best criticality classification for each type of spare part and several AHP models are implemented and integrated with each node of the decision tree. By these technique, Braglia et al. (2004) have avoided using over-complex and unmanageable decision diagrams and have taken into account numerous potential attributes influencing the spares inventory management policy in an easy and rational manner. In their developed model, evaluation criteria taken into consideration are group under four main classes: spare part plant criticality, spare supply characteristics, inventory problems and usage rate. Several of sub criteria analyzed in the model are quality problems, lead time, internal repair, price, space required, obsolescence, deterioration problems, production loss, redundancies, etc.

Çakir and Canbolat (2008) have proposed an inventory classification system based on the fuzzy analytic hierarchy process by integrating fuzzy concepts with real inventory data and designing a decision support system assisting a sensible multi-criteria inventory classification. Their study was conducted in a small electrical appliances company and they validated the design of the proposed multi-criteria inventory classification system and its underlying fuzzy AHP model.

Simunovic et al. (2008) has applied and compared the traditional ABC inventory classification, based on one criterion, and inventory classification based on multi criteria decision making technique, which is analytical hierarchy process on spare parts in agricultural machine. For traditional inventory classification Simunovic et al. (2008) has analyzed inventory with ABC analysis based on annual cost usage. On the other hand, evaluation criteria involved in multi criteria inventory classification with AHP methodology were annual cost usage, criticality and lead time1 and lead time 2.

The AHP methodology includes pair wise comparison of criteria, but not the pair wise comparison of alternatives.

AHP analysis is also widely used in inventory management of railway systems. For example, Sönmeztürk et al. (2008) have classified maintenance spare parts inventory using multi criteria ABC analysis and fuzzy AHP. They used fuzzy analytic hierarchy process model to define the major criteria used in inventory management in the railway organization and estimated level of significance of the criteria. Then, they classified the spare parts inventory according to following criteria: lead time, unit cost, value usage and criticality. They used importance scale from 1 to 9 as in most evaluation by AHP process models. AHP has been praised for its ease of use and its inclusions of group opinions; however, the subjectivity resulting from the pair-wise comparison process of AHP poses problems (Yu, 2011).

4.4 Data Envelopment Analysis

Ramanathan (2006) has developed a weighted linear optimization model that is based on the concept of data envelopment analysis. In his work, the weighted additive function is used which covers all the performances in terms of different criteria of an item and optimization linear model is defined for each item. The model developed by Ramanathan (2006) generated the optimal inventory score for each item as well as weighted factor values (weights) for all the criteria. However, in a model with large inventory items his method is time consuming. Several researchers have modified and improved Ramanathan’s model. For example, Zhou and Fan (2007) have extended Ramanathan’s model by incorporating some balancing features for multi-criteria ABC inventory classification. They developed more reasonable and encompassing index, since it used two sets of weights that are most favorable and least favorable for each item and the purpose of the model was to aggregate multiple performance scores of an item with respect to different criteria into a single score for the subsequent ABC inventory classification (Zhou and Fan, 2007).

Ng (2007) has improved the method developed by Ramanathan by just simplifying the model which was easier to apply. In his approach, Ng (2007), instead of using linear optimizer, has transformed and normalized all the criteria data for each alternative to the scale from 0 to 1. After that, partial averages of the transformed criteria values are found and maximal partial average value for each item is chosen to rank the items by the chosen value. At last, traditional ABC analysis could be applied.

Data envelopment analysis is also widely used with ABC analysis of inventory in railway systems management. Çelebi et al. (2008) have extended the classical ABC analysis by developing a multi-criteria inventory classification approach for supporting the planning and designing of a maintenance system for railway spare parts inventory. In their research, the criticality model is proposed and ABC analysis is performed according to the chosen criteria of criticality, lead times, value usage and units costs. The weights of the evaluation criteria were generated by a DEA like linear optimization to avoid the subjectivity on weights assignments. The DEA model used is a simple input-oriented single input, multi output model which is very similar to the model developed by Ramanathan (2006).

A similar method for estimation of weights of criteria as was done by Çelebi et al. (2008) have been performed by Hadi-Venchen and Mohamanghasemi (2011) in their examination of multiple criteria ABC inventory classification model. They have integrated fuzzy analytical hierarchy process (FAHP) and data envelopment analysis and applied their methodology with real case study. The proposed FAHP-DEA methodology used the FAHP to determine the weights of criteria, linguistic terms such as Very High, High, Medium, Low and Very Low to asses each item under each criterion and the data envelopment analysis (DEA) method to determine the values of the linguistic terms and the simple additive weighing method to aggregate item scores under different criteria in an overall score for each item.

Data envelopment analysis (DEA) maximizes the artificial inventory score that is used to classify each inventory item. Unlike AHP, the weights given to classified criteria are solved automatically when the DEA model is optimized. Like statistical clustering technique, this model must be reprogrammed and solved whenever a new inventory item is introduced (Yu, 2011).

4.5 Artificial Neutral Network

ABC analysis has been popular and effective method used to classify inventory items into specific categories that can be managed and controlled separately. Apart from traditional ABC analysis, multi criteria ABC analysis has gained much attention and many different models and techniques have been developed to improve classification of inventory in multi item operations and processes. Advancement of computer technology has enabled usage of complex database methodologies that helps manage the decision-making process. Moreover, human involvement in the decision making process has been reduced and thus it enhances the accuracy and consistency of the decision making process, at the same time decreases the processing time. One of the methods for multi-criteria inventory management has been artificial intelligence (Guvenir and Erel, 1998). Many researchers have compared artificial-intelligence based classification techniques with traditional multiple discriminant analysis. One of artificial intelligence based technique is an artificial neural network (ANNs) which is used widely for classification processes. Examples of these artificial neural network models include support vector machines (SVMs), back propagation networks (BPNs), and the k-nearest neighbour (k-NN) algorithm, etc. (Yu, 2011). Neural network models have been used and developed to classify and control inventory. The ANNs have two major strengths over the more traditional model fitting techniques. First, ANNs are capable of detecting and extracting nonlinear relationships and interactions within predictor variables. Second, the inferred relationships and associated estimates of the precision of the ANN are not related to the several assumptions about the distribution variables (Partovi and Anandarajan, 2002).

Artificial neural network have been used for classifications purpose, as well as for forecasting problems in a variety of applications. They are very helpful for finding nonlinear surfaces and separating the underlying patterns. Paliwal and Kumar (2009) have performed a comprehensive study on neural network articles, categorizing the application of networks into categories: accounting and finance, health and medicine, engineering and manufacturing, and marketing. Accounting and finance is the category with the greatest number of applications, especially with regards to bankruptcy prediction, credit evaluation, fraud detection and property.

Partovi and Andarajan (2002) have performed artificial neural network for ABC classification for stock keeping units in a pharmaceutical company. They compared the two learning methods in ANN: namely back propagation and genetic algorithm. Their results showed that there was no significant difference between two learning methods used to develop ANN. The evaluation criteria used in the models were unit price ($/unit), ordering cost ($/order), demand range (units/year) and lead time (days). The results showed that neural network based classification models have a higher predictive accuracy than conventional multiple discriminant analysis (MDA) technique.

Li and Kuo (2008) have developed an enhanced fuzzy neural network (EFNN) based decision support system for managing automobile spares inventory in a central warehouse. In their system, the EFNN is utilized for forecasting the demand for spare parts.

Chen, et al. (2009) have proposed moving back-propagation neural network (MBPN) and moving fuzzy neuron network (MFNN) to effectively predict the critical spare parts requirements in a wafer testing factory.

Other than classification problems, artificial neural network are used in clustering of inventory. For example, Malakooti and Raman (2000) have utilized unsupervised learning clustering ANN with variable weights for clustering alternatives and used feed-forward for the selection of the best alternatives for each cluster of alternatives. In their model, the learning mechanism of ANN was a generalized Euclidean distance where by changing its coefficients new formation of clusters of alternatives was achieved. The main difference between their approach of multi criteria clustering and the general clustering method lies in the fact that their model deals with criteria that have to be optimized and the selection of the best alternative.

The use of artificial neural networks in classification problems has become popular among several fields. ANN are used in classification of three phase induction motors (Ghate and Dudul, 2010), in planning, scheduling and analyzing a flexible manufacturing systems (Kurt, 2003). Other researches include the benchmark of the developed classification method to see their effectiveness, for example, Yu (2011) has performed comparison of three different artificial intelligence techniques with the multiple discriminant analysis techniques developed by Reid (1987), Flores et. al. (1992), Ramanathan (2006) and Ng (2007). In his study, techniques used were back-propagation network, support vector machine and k-nearest neighbor's methods. Back propagation networks (BPN) are the most widely used for classification method for training an artificial neural network. A BPN utilizes supervised learning methods and feed forward architecture to perform complex functions such as pattern recognition, classification and prediction. Support vector machine (SVM) were developed by Vapnik et al. (1995) and it employs structural risk minimization rather than the empirical risk minimization used by conventional neural networks. SVM use linear model to implement nonlinear class boundaries by the nonlinear mapping of input vectors into a high-dimensional feature space (Yu, 2011). K-nearest neighbors method is a non-parametric technique for classifying observations (Cover and Hart, 1967). Computations of the measure of distance or similarity between observations are conducted prior to classification and then, newly introduced item is classified to the group where the majority of k-NNs belong. As a result of his study Yu (2011) have concluded that forecasting accuracy of classification with artificial intelligence methods perform better than traditional MDA techniques like traditional ABC (Reid, 1987), AHP (Flores et al., 1992), optimal score (Ramanathan, 2006), and scaled score (Ng, 2007).

Ghate and Dudul (2010) have employed neural network classifier in their research and have proposed optimal multi layer perception (MLP) network model for classification model. Also, they have compared classification efficiency with different network structure, the Kohonen self-organizing map (SOM) network. The main purpose of their study was to reveal the importance of selection of significant inputs and selection of neural network parameters, which make the structure compact, and create highly accurate networks. In their work they have applied different input network parameter and compared the classification accuracy to determine the optimum network model. Results of network models with different transfer functions, momentum and learning rates have been evaluated and the optimum model resulted to be multi layer perception network with hyperbolic tangent function (Ghate and Dudul, 2010).

Traditional ABC analysis should be replaced with multi-criteria classification approaches in order to manage inventory more efficiently. Multi-class classification utilizing multiple criteria requires techniques capable of providing accurate classification and processing a large number of inventory items. AI-based classification techniques such as back-propagation network and genetic algorithm have proven to be efficient methods for classifying inventory items. The use of these techniques will improve the effectiveness and efficiency of inventory management (Yu, 2011).

4.6 Genetic Algorithms

Guvenir and Erel (1998) have applied genetic algorithm (GA) for multi criteria inventory classification and have compared their results with classification efficiency of analytical hierarchy process. Genetic algorithms are general-purpose search algorithms that use principles inspired by natural population genetics to evolve solutions to problems. The basic idea in genetic algorithms is to maintain a population of knowledge structures (also called chromosomes) that represent candidate solutions to the current problem. A chromosome is a sequence of genes. The population evolves over time through competition and controlled variation. The initial population can be initialized using whatever knowledge is available about the possible solution. Each member of the population is evaluated and assigned a measure of its fitness as a solution. After evaluating each structure in the populations, a new population of structures is formed in two steps. First, structures in the current population are selected for replication based on their relative fitness. High-performing structures might be chosen several times for reproduction, while poorly performing structures might not be selected at all. Then, the chosen structures are modified using idealized genetic operators to form a new set of structures for evaluation. The primary genetic search operator is cross-over operator, which combines the features of two parent structures to form two off springs similar to their parents. The role of the cross over operation is to form new fit chromosomes from fit parents. There are many possible ways of crossover: the simplest swaps corresponding segments of a string, list or vector representation of parents. In generating new structures for testing, the cross over operator usually draws only on the information present in the structures of the current knowledge base. If specific information is missing, because of storage limitations or loss incurred during the selection process of a previous generation, then crossover is not able produce new structures than contain it. A mutation operator which changes one or more components of a selected structure provides the way to introduce new information into the knowledge base. In most cases, mutation serves as a secondary search operator that ensures the ability to reach all points in the search space. The resulting off springs are then evaluated and inserted back into the population. This process continues until either, an absolute fittest structure is detected in the population or a predetermined stopping criterion (e.g., maximum number of generations or maximum number of fitness evaluation) is reached (Guvenir & Erel, 1998). In their model, Guvenir and Erel (1998) have proposed a new cross over operation, called continuous uniform crossover, such that it produces valid chromosomes given that the parent chromosomes are valid. One other research which compared classification performance of different artificial intelligence techniques on classificiaton of inventory with multiple criteria ABC analysis was done by Partovi and Anandarajan (2002) on stock keeping unit in pharmaceutical industry. Partovi and Anandarajan (2002) have compared classification performance of back-propagation network and genetic algorithm with multiple discriminant analysis technique. Their study has revealed that both artificial neural network models had higher predictive accuracy than MDA. The results also indicated that there is no significant difference between the two learning methods used to develop the artificial neural network.

4.7 Other Multi Criteria Classification Techniques Widely Used in ABC Analysis

4.7.1 Ng model

Ng- model was developed by Ng (2007) and his model is simple classification model used for multiple criteria ABC classification. According to model all the measurements are transformed to 0-1 scale according to below equation 4.1:

$$\frac{y_{ij} - min_{i=1,2,\ldots j}\{y_{ij}\}}{max_{i=1,2,\ldots j}\{y_{ij}\} - min_{i=1,2,\ldots j}\{y_{ij}\}} \quad \textbf{(4.1)}$$

For more detailed explanation of model please see Ng (2007). More detailed research on Ng-model for inventory classification has been done by Hadi-Vencheh (2009) where he has developed a nonlinear programming model which determines a common set of weights for all items. The developed model incorporated multiple criteria for ABC classification and at the same time maintained the effects of weight in the final solution. The evaluation criteria analyzed in the model were annual unit cost, annual dollar usage and lead time.

4.7.2 Case based model for multi criteria abc classification

Chen et.al. (2008) have introduced a case based multi criteria ABC analysis that improves traditional ABC analysis by accounting for additional criteria such as lead time and criticality in order to provide more managerial flexibility with inventory classification. The main problem with many existing multi criteria decision aid methods lies in the acquisition of the decision maker’s preference information in the form of value or weights Chen et.al. (2008). Case based reasoning is an approach to finding preferential information using cases selected by the decision maker (Jacquet-Lagreze et al., 2004; Slowinski, 1995). The choice of cases may involve: past decisions taken by the decision maker; decisions taken for a limited set fictitious but realistic alternatives; decisions rendered for a representative subset of alternatives under consideration, which are sufficiently familiar to the decision maker that they are easy to evaluate v Chen et.al. (2008). A main advantage of case based reasoning is “decision makers may prefer to make exemplary decisions than to obtain them in terms of specific functional model parameters” (Doumpous, 2002). The model uses decisions from cases as input, preferences over alternatives are represented intuitively using weighted Euclidean distances which can be easily understood by a decision maker.

4.7.3 Abc fuzzy classification

Chu et al. (2008) have purposed a new inventory control approach called ABC–fuzzy classification (ABC–FC), which can incorporate manager’s experience, knowledge, and judgment into inventory classification and can be implemented easily. Fuzzy classification is a technique that uses the available information in a set of independent attributes to predict the value of a discrete or categorical dependent attribute. The main difference of their paper from earlier works is as follows: (1) the approach can handle any combination of item attribute information that is important for managerial purposes (e.g., the criticality of a stock-out, order size requirement of the item); (2) manager’s preference for grouping based on operational performance can be accommodated; (3) fuzzy statistical discrimination criteria are considered; (4) our ABC–FC approach can be easily implemented on the spreadsheet, which is more accessible to practitioners.

5. AN ARTIFICIAL NEURAL NETWORK MODEL FOR ABC CLASSIFICATION OF SPARE PARTS IN RAILWAY SYSTMES

The usage of artificial neural network in inventory management especially, in spare part classification is more precise than the traditional classification techniques. We will extend the model developed by Çelebi et al. (2008) for classification of railway spare parts by the classical ABC analysis and developing multi-criteria inventory classification approach. The model will be developed for the inventory classification using ABC analysis and ANN approach for spare parts in light railway system. Maintenance operations play crucial role in the performance of railway vehicles and they are very important for railway services and very significant to provide uninterrupted and high quality service to passengers. To insure efficiency in railway services availability and reliability on spare part control method should be provided by maintenance operations.

5.1 Identification of Evaluation Criteria for ANN Model

Traditional ABC analysis for inventory classification is solely based on annual dollar usage (Chen et. al., 2008). Although annual dollar usage is an important criterion in inventory management systems, other criteria may also deserve management's attention in classification of inventory. Chen et. al. (2008) states that classification obtained from traditional ABC analysis sometimes should undergo further adjustments. For example, annual dollar usage of inventory may not be significant, but inventory's stock-out penalty cost may be sufficiently high and vice versa. In such cases, inventory manager has to pay attention to other significant criteria, which will influence his/her decision on how much inventory to keep in stock or which inventory should gain close stock control. These criteria for inventory control, suggested in different literature, may be ordering cost, demand distribution, obsolescence, criticality of part, lead time, number of requests, substitutability, scarcity, perish ability, durability, stock ability, reparability, criticality, commonality, etc. (Flores and Whybark, 1986, 1987; Cohen & Ernst, 1988; Partovi and Anandarajan, 2002; Ramanathan, 2006; Zhou & Fan, 2007; Chen et al., 2008; Hadi-Vencheh, 2010).

Evaluation attributes suggested by Partovi and Anandarajan (2002) for evaluation of spare parts model by ANN are unit price, ordering cost, demand and lead time. These criteria suggested by Partovi and Anandarajan (2002) may be sufficient for the evaluation, however, criteria used by Çakır abd Canbolat (2008) given in Table 5.1 seems more efficient to be chosen for the evaluation for spare part inventory.

Table 5.1: Subject criteria definitions (Çakır and Canbolat, 2008).

Criterion name	Definition
Price/cost	Price of the inventory item in case it is outsourced and cost in case it is manufactured in the company
Annual demand	Annual demand of the inventory item in the production process
Blockade effect in	Critical threat on the stability of production process when the item is case of stockout not available for production right on time
Availability of the substitute material	Availability and conformance of substitute materials
Lead time	Time required to get the item available on hand for production. Delivery time in case the item is outsourced and production time in case the item is manufactured
Common use	Common usage of the item in the product family

According to the nature of inventory items and industry, the criteria by which inventory is classified will have different weights. In the real world, prioritizing of the weights of criteria is always subjective that is depending on the conditions governing on industry and market, inventory managers usually assign different weights to the criteria of their inventory (Hadi-Venchen and Mohamadghasemi, 2011). For instance, when items are guaranteed to be supplied at certain time, the weight of lead time criteria is set lower than other criteria during evaluation. Moreover, it is stated that importance of criteria might differ among several industries (Hadi-Venchen and Mohamadghasemi, 2011). In hospitals, inventory's expiration date criteria will be important criterion; however, for manufacturing industry like cement this criterion will be insignificant. Similar is applicable for maintenance spare parts inventory in railway systems, there are criteria that are more important than several others. Çelebi et al. (2008) have evaluated spare parts inventory in light railway systems according to following criteria: criticality, value usage, unit cost and lead time.

Value usage criterion, similar as in classical ABC classification, is known as annual value usage where items are evaluated through their annual monetary usage values which are the products of annual usage quantities and the average unit prices of the items. The value of an item represents the common control characteristic of items and it is avoided to stock large amounts of high value items. Usually, a high value item favors controlling of inventories of that item more carefully. Thus, this criterion is positively related to the importance of the item because items with the highest annual monetary value receive the most attention where low monetary value items are controlled routinely (Çelebi et al., 2008).

Lead time criterion plays an important role among several factors that can affect the management of the inventory. Lead time of spare part item includes the time between placing order at the supplier of the item and the moment the item is available for usage. Both the length of the lead time and its variability might be crucial in maintaining adequate supply of an item without excessive costs (Flores and Whybark, 1985). The period of lead time is important since it directly establishes the stock levels of items with unknown demand and states the response time to a crisis. The variability affects the amount of safety stock necessary to provide the level of desired service. For instance, items with long lead times may incur financial losses as a result of possible interruption of maintenance operations or large inventory stocks. Usage of time as a measure basis of lead time criterion constitutes a common basis for all items in the product catalog and also provides both the user and the supplier with a common means of setting the goals and for controlling the performance of operations.

Unit cost is another criterion and is important in control of items on strategic and tactical level. High value items necessitate decisions like setting up the incentives, developing coordination with suppliers and creating opportunity for high negotiation power. On the other hand, the replenishment arrangements with low price items have to be efficient so that the administrative costs do not increase unreasonably in proportion to the value of the items themselves (Çelebi et al., 2008).

The criticality analysis includes a substantive amount of subjective criteria used in evaluating the criticality of parts practice (Cohen et al., 1997). It is related to the results of damage of a failure and shortage of inventory and the possibilities to control the situation. The criticality was divided into three sub-criteria: commonality, substitutability and stock-out penalty which covers the essentials of railway maintenance systems (Çelebi et al., 2008).

a. Stock-out penalty criteria is the costs associated with the spare part being out of stock. The criticality of a spare part item is related to the consequences resulted from the lack of the item when it is necessary. In theory, stock-out penalty can be evaluated by estimation of downtime expenses caused by failure to be corrected by the use of the part. Yet, it is usually difficult to determine the degree for criticality in practice and it might be sufficient to define a few degrees of criticality for practical purposes. One known approach is to relate criticality to the time in which the failure is required to be corrected. Huiskonen (2001) has mentioned three degrees of criticality in regards with consequences might be estimated as the following bases:

    - High: The failure needs to be repaired and the spare parts have to be provided immediately,
    - Moderate: The failure can be tolerated with temporary adjustments for a short time while spare part can be provided,
    - Low: The failure is not critical for the operation and it can be corrected after a long period of time when spare is supplied.

   The quantification of the given criticality degrees is done through assigning a penalty index (α_n) for each item n and setting it 5 for a high level item, to 1 for a low level and to 3 for a moderately critical item (Sönmeztütk et al., 2008).

b. Substitutability criterion is another important aspect of criticality. The substitution potential provides flexibility in response to problems, reducing the importance of the item in the spare parts inventory. If the item has a close substitute, more flexibility and reduced response time is possible, both of which reduce the criticality of the part.

Substitutability of spare part has a direct influence on the purchasing decisions which may affect both efficiency and effectiveness of maintenance operations. One dimension of substitutability is technical. Technical similarities between spare parts inventory can allow mutual substitution avoiding any loss of function or suitability. Higher level of substitutability might enable lower levels of stock-out risks that when there is lack of one type of item, it can easily be replaced with interchangeable item. However, in practice there are some situations where replaceable item is suitable as an alternative for the item in question and yet it may cost so much higher that it becomes a poor substitute (Çelebi et.al, 2008).

Other dimension of substitutability is connected to the availability of different suppliers who can provide the similar product with very little or zero quality and cost difference. Among the wide range of maintenance spare parts there are typically two type, standard type parts and special parts. Standard parts are widely used by many users and are readily available from many suppliers; on the other hand, special parts are specifically tailored for and used by a particular use only. The availability for standard parts is always good and there are stocks of these parts at different levels of the supply chain and the suppliers are eager to cooperate with the users as volumes are high and they offer economies of scale. However, for user-specific parts quite the opposite is valid: suppliers are not interested to keep low volume special items in stock and user is responsible from the availability and control of the special items himself. Similar to stock-out penalty index, a substitutability index might be a method of determining the substitution availability of an item. A substitutability index of $\gamma_i = 5$ implies that item i is fully differentiated, $\gamma_i = 1$ implies that it has perfectly substitutable products and $\gamma_i = 3$ implies that item *i* is moderately substitutable.

c. Commonality criterion is another aspect of criticality which is the measure of how many uses there are for a spare part item. If the item is used in many different vehicles or maintenance types, it might be important to devote extra attention to it and for management purposes it could be classified in a group of A items. The usage of common parts can be beneficial in terms of risk sharing and substantial savings up to can be achieved by use shared stocks compared to using separate stocks (Kranenburg and Van Houtum, 2007). Moreover, when the same spare part item can be used with several maintenance types or vehicles, the system allows the possibility of economies of scale since a common component can be supplied in large volumes (Çelebi et al., 2008). However, a situation when common components are out of stock will have higher influence on the maintenance system and as a result the maintenance schedules that have a use of shared component will be delayed or changed. Measurement of commonality is not so easy. Though there are several types of commonality indexes available in literature (Lyly-Yrjanainen et al., 2004), a simple and useful measure is the number of different maintenance types that the use of the item is needed.

In our model criticality will be analyzed under above three sub-criteria: namely stock-out penalty, commonality and substitutability and other criteria: value usage, lead time and unit cost as was suggested by Celebi et al. (2008) in their evaluation of criticality of spare parts inventory model. The dimensions of criteria are summarized in Table 5.2 as chosen to be as was suggested by Çelebi et al. (2008).

Table 5.2: Dimension of evaluation criteria for classification of spare parts.

Criteria	Dimension
Unit cost	Turkish lira (TL)
Lead time	Weeks
Value usage	Unit
Commonality	Number of different usage
Substitutability	1. Low 3. Medium 5. High
Stock out penalty	1. Low 3. Medium 5. High

5.2 Network Building

Designing a neural network requires the determination of several parameters and variables that define the network and that influence the learning performance of the network.

The most important choice is the determination of the necessary architecture, the number of layers, the number of processing elements in every layer and the number of synapses that are connected to each node. Moreover, it is also important to choose the appropriate activation function of the hidden and output nodes, the training algorithm, data transformation and normalization methods, training and testing sets and the performance criteria (Zhang et al, 1998).

5.2.1 Network structure

Artificial neural network structure may differ according how nodes and their connections are linked (Elmas, 2003). A network is said to be feedforward network if it contains no directed cycles back, otherwise it is called as recurrent network.

Most research models use a three layer network constituting one input, one hidden and one output layer (Partovi and Anandarajan, 2002; Çelebi and Bayraktar 2008; Turer et al., 2009). Hidden nodes with their activation function are required to introduce nonlinearity in the neural network. Most popular network model used for multi criteria inventory classification is multi layer perception network (Partovi and Anandarajan, 2002; Çelebi and Bayraktar 2008; Turer et al., 2009).

5.2.2 Multi layer feedforward networks

Networks that consist of a set of neurons that are logically arranged into two or more layer are known as multi layer feedforward neural networks (Masters, 1993). The architecture of the multi layer network constitutes of one input and one output layer and additional one or more hidden layer which have computation nodes that are called hidden neuron or processing elements.

Hidden units with their activation function are used to capture the nonlinear structures in the system observed (Qi and Zhang, 2001). The ability of hidden neurons to extract nonlinear structures is particularly important when the systems under consideration is complex, if the size of the input layer is large and the underlying relationship between the input and output units are undeterminable.

A Multi layer feedforward neural network, especially the one hidden layer feedforward type is the mostly used model form in literature (Masters, 1993). For the classification of spare part invetory in railway system Multi-layer perception (MLP) model was chosen because it has simple architecture and algorithm (Turer et al., 2008). Moreover, it has been successful in solving approximation and classification problems (Çelebi and Bayraktar, 2008). Generally, MLP consists of a group of nodes which form an input layer, one hidden layer and last, an output layer.

The popularity of the one hidden layer feedforward networks come from the fact that they are universal approximators (Zhang, 2003). Regarding the two hidden layer networks, it has been stated in the literature that two hidden layer network slows the learning rate and affect the overall performance of the network (Şahin, 2002; Masters, 1993).

5.2.3 Learning rule in neural network

All knowledge in an artificial neural network is encoded in its interconnection weights among the neurons calculated by the learning process. A weight represents the strength of associated among connected features, concepts propositions or events during a training phase. A neural network learns by adjusting according to learning method by which the interconnection weights are changed appropriately. As a result, learning can be explained as the modification in weights in the layers in order to perform the required performance criteria (Haykin, 1999).

Learning method for the artificial neural network model was taken as supervised learning, specifically back-propagation algorithm as was suggested by Partovi and Anandarajan (2002) for ABC classification in their network building method. The multilayer perception back propagation model suggested by Partovi and Anandarajan (2002) is given below Figure 5.1:

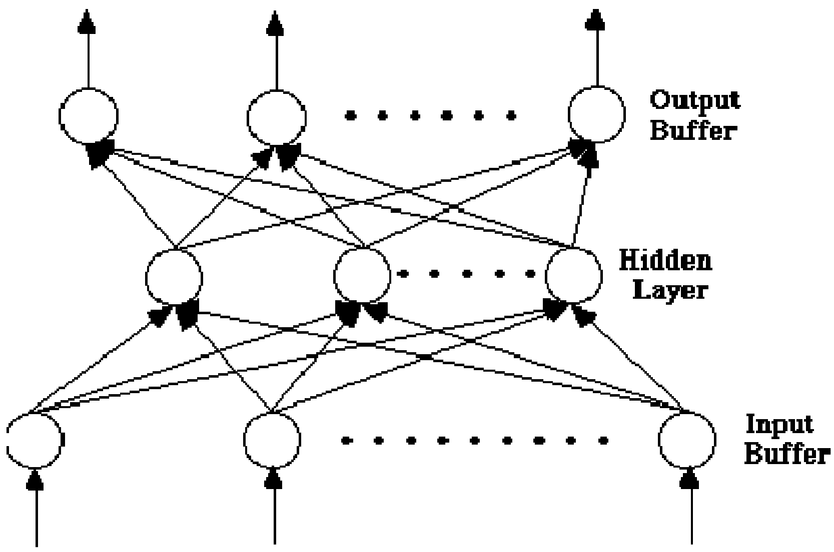


Figure 5.1: Artificial neural network structure (Partovi and Anandarajan, 2002).

Hidden layers are connected to previous layer by means of synapses and also to output layer. In the hidden layer, activation function is generally hyperbolic tangent or logistic sigmoid (Çelebi and Bayraktar, 2008; Turer et al. 2009; Partovi and

Anandarajan, 2002; Aksungur and Kavlak, 2009). MLP performs approximation or classification; as a result the output of neurons activation function can be linear or nonlinear.

The model will contain one input layer, one hidden layer and one output layer. In our model there will be six criterion nodes, which are stock out penalty, commonality, substitutability, value usage, unit cost and lead times, in the input layer and three nodes for the output layer, namely inventory items A, B and C. Total data contained in the evaluation consist a values of 71 data set of different spare parts. Complete data set has been divided into two subsets, the training and the test set, where 85 % (70% training and 15% cross validation data) of the total data is used as training set and the rest 15 % of the total data used as testing set to evaluate the accuracy and generalization ability of the network. In the evaluation data set of 60 data of the spare parts was used as a training set and for validation of the developed neural network sample of 11 data of spare parts was used as a testing set.

5.2.4 Choice of transfer function

The transfer function is important and it depends on the nature of the output of the network. There are a number of choices including the step function, sigmoid function, hyperbolic tangent function and linear function among others. Turer et.al. (2009) have preferred to utilize hyperbolic tangent function for their ANN for supplier evaluation process. Furthermore, it has been specified in several examination of different modeling of artificial neural network approaches with MLP for classification problems that hyperbolic tangent function, in equation 5.1, results in more efficient and reliable solution because of its nonlinear behavior which allows networks to model nonlinear mappings. Also, it has a larger range [1,+1] in contrast to [0,1] as in logistic sigmoid that corresponds to a greater weight range (Celebi and Bayraktar, 2008).

$$tahn(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad \textbf{(5.1)}$$

$$x = \sum_{i=1}^{n} x_i w_i \qquad \textbf{(5.2)}$$

On the other hand, sigmoid tangent function is also widely used in classification models. For example, Partovi and Andarajan (2002) have used sigmoid transfer function in their ANN model as well as Aksungur and Kavlak (2009). The sigmoid transfer function is illustrated in eq. 5.3

$sigm(x) = \frac{1}{1+e^{-x}}$b **(5.3)**

$x = \sum_{i=1}^{n} x_i w_i$ **(5.4)**

5.2.5 Determination of hidden layer number and nodes in hidden layer

The aim of neural network is to detect the hidden relationship in the data and to perform nonlinear mapping between input and output variables which is know as training of network. As have been mentioned before, the neural network structure chosen for classification model includes one input, one hidden and one output layers, which was chosen for its simplicity. The number of nodes in the input and output layers are predefined according to the evaluation criteria and classification model. Yet, the number of nodes in the hidden layer has to be determined by trial and error. The number of neurons in the hidden layer was calculated based on five-fold-cross validation method which was suggested by Çelebi and Bayraktar (2008). In five-fold cross validation method, training data is divided into five disjoint subsets and one of these subsets is hold as validation data and the rest four subsets are utilized as training data set. The network is run five times by 1000 iterations and by changing the number of nodes from two to twenty in the hidden layer. The results of training gave the minimum of MSE values for both training and cross validation data which were noted. Then, the first subset input back into training set and second subset is chosen as validation set and training is repeated again by noting the minimum MSE for the second subset. This procedure is performed till all five subsets are used as validation sets and we result with data of minimum MSE for five subset sets for different number of nodes in hidden layer (Çelebi and Bayraktar, 2008). MSE, mean squared error function calculates the accuracy rate of the configuration model by determining the difference between the output of the model and the desired output as can be seen in equation 5.5.

MSE= $\frac{1}{n} \times \sum_{i=1}^{n}(d_i - y_i)$ **(5.5)**

Finally, the average of minimum MSE for five subsets is calculated and the network model with hidden nodes which has the least average of minimum MSE for cross validation data is chosen as the optimum configuration (Çelebi and Bayraktar, 2008). To sum up, by five fold cross validation method the number of nodes the hidden layer is calculated by a stepwise training of five subsets from training data in order to select the model which minimizes the generalization error.

For the determination of hidden nodes in the hidden layer in classification of spare part inventory, evaluation data for the five fold cross validation method consists of 60 spare parts data, of which 48 of data is taken as training and 12 data as validation set. The structure of neural network which will have the minimum of Mean Squared Error (MSE) at the end of training process will be chosen for the evaluation of the testing data. Data for 11 spare parts is left out for the testing process to check the validity of the trained network.

5.2.6 The stopping criteria, learning rate and momentum

The learning process should be stopped when network reaches a predefined criterion. Generally mentioned in literature, the training is stopped either when training performance reaches a target performance or when a specific epoch number is reached. In our multi layer perception network model, the stopping criterion for the training process is taken to be the minimum of mean square error (MSE) or a network reached certain number of epochs. In their evaluation , Çelebi and Bayraktar (2008) have stopped training of the neural network as epoch number has reached 1000 iterations or the MSE for cross validation set is decreased below the threshold value of 0,01.

The learning rate of the network is dependent on the step size and momentum in each layer. Ghate and Dudul (2010) have evaluated optimal parameters for the multi layer perception model with three layers and have found that optimal step sizes should to be 1.0 and 0.1 for hidden and output layers respectively and moreover, for the momentum for hidden and output layers the optimal values to be 0.7 as can be seen in the below Figures 5.2 and 5.3:

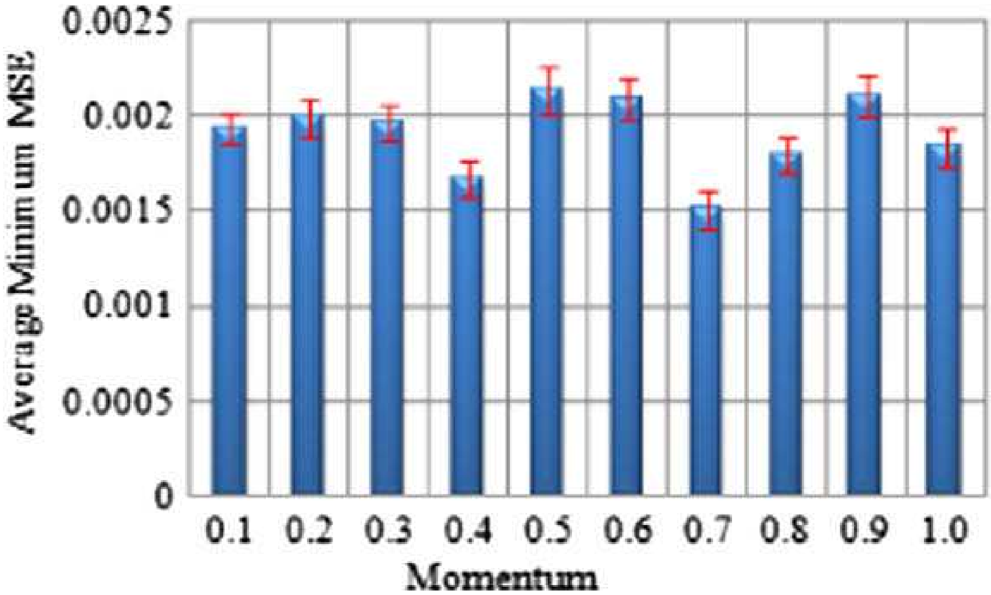


Figure 5.2: Variation of average minimum mse on training dataset with step size of hidden layer (Ghate and Dudul, 2010).

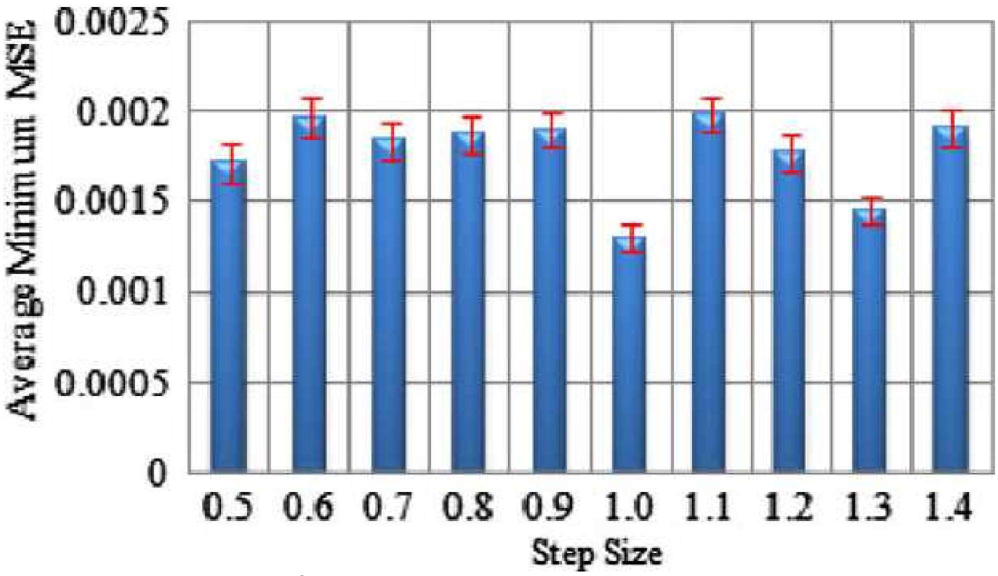


Figure 5.3: Variationof average minimum mse on training dataset with momentum output layer. (Ghate and Dudul, 2010)

To sum up, in the evaluation of spare parts inventory by ABC classification with neural networks, the network parameters for the neural network was chosen as mentioned below in Table 5.3.

Table 5.3: Network parameteres for the ann model

Learning Method	Supervised
Learning model and type	Multi layer perception with back propagation algorithm
Performance Function	Mean Square Error (MSE)< 0,01
Number of Hidden Layer	1
Number of Neurons in Hidden Layer	To be determined by Five Fold Cross Validation Method
Transfer Function	1. Hyperbolic Tangent 2. Sigmoid
Number of Input Neurons	6 (Lead time, Value Usage, Unit Cost, Stock-out Penalty, Substitutability and Commonality
Number of Output Neurons	3 (A,B,C)
Hidden Layer Step Size	1.0
Output layer Step Size	0.1
Momentum for Hidden and Output Layers	0.7
Number of Maximum Epoch	1000 iterations
Learning mode	Sequential (on-line)

5.3 Neurosolutions 5.0 Software Used for Classification Model

Neurosolutions 5.0 software was used to run the neural network model for railway spare part inventory classification using ABC analysis. Neurosolutions 5.0 is incorporated package in Excel and it has very convenient tools such as Neurosolutions for Excel, Neurobuilder and Neuroexpert. The data set of 71 spare parts was entered in Excel file as can be seen in Table A.1 in the App. A.1

For the modeling the data has been randomized before the building neural network. As was mentioned in previous sections, 85% (70% training and 15% cross validation)of data was selected as training and 15% of data was used as testing data as was suggested by Ghate and Dudul (2010). Evaluation data for 48 spare parts was chosen as training, 12 data as cross validation and 11 data for spare parts was taken as testing to validate to results of the training process.

5.4 Training Process by Neurosolutions

As was mentioned in previous sections, the number of hidden neurons in the hidden layer is determined by five-fold cross validation method.

5.4.1 Five-fold cross validation method

The training data set is divided into 5 disjoint subsets; one of these subsets is retained as cross validation data and other are selected of training data subset. The training data has been divided into 5 subsets as can be seen in following Table A.2 in App. A.2.

For a different number of hidden neurons, the network is run 5 times by 1000 iterations by choosing one subset as cross validation data from the divided 5 subsets till all subsets has been used as validation data at once. The minimum of MSE data for training and cross validation data is recorded for each subset. At the end, we end up with minimum MSE data of training and validation data for each five subsets for nodes from two to twenty in the hidden layer. The results of training process for five fold cross validation method can be seen in Table B.1 and Table B.2 in App. B.1

The average of minimum MSE of five subsets is calculated for each number hidden neuron in the hidden layer and results are presented in Table 5.4 and Table 5.5. The example of calculaion of average minimum MSE can be seen in App. B.2.

The neural network was run for two different activation functions, hyperbolic tangent and sigmoid functions.

As can be seen from Table 5.3, the neural network with 8 nodes in hidden layer gives the least MSE data as 0,03736, in the five-fold cross-validation method for hyperbolic tangent function.

Table 5.4: Average minimum mse for training and cross validation data with hyperbolic tangent function.

	AVERAGE OF MINIMUM FOR MSE	
# Of Nodes	Training Data	Cross Validation Data
2	0,085196608	0,060057349
3	0,104065011	0,080212494
4	0,098453339	0,063394639
5	0,103797196	0,058799254
6	0,108533403	0,051304857
7	0,124705773	0,078014818
8	**0,117554536**	**0,037361614**
9	0,119401166	0,070503936
10	0,102225101	0,050416581
11	0,122653927	0,067211223
12	0,121354667	0,054275056
13	0,150951974	0,099301261
14	0,1308181	0,087902196
15	0,14452272	0,092461572
16	0,131054975	0,109547602
17	0,136192376	0,091467151
18	0,163597183	0,08304966
19	0,138571933	0,079926288
20	0,130147528	0,088493876

Similarly, the neural network with sigmoid transfer function has least minimum average MSE at 5 nodes in the hidden layer as 0,029434 as a results of five fold cross valdation method.

Table 5.5: Average minimum mse for training and cross validation data with sigmoid function.

	Average Of Minimum Mse	
# Of Nodes	Training Data	Cross Validation Data
2	0,012802124	0,039772375
3	0,010029668	0,036783332
4	0,010363869	0,037398745
5	**0,008184525**	**0,029434182**
6	0,007971996	0,034192718
7	0,007298897	0,030355518
8	0,007575996	0,034793708
9	0,007062787	0,036389971
10	0,00716824	0,033420092
11	0,006585697	0,030453519
12	0,006237987	0,037203207
13	0,007348075	0,036478183
14	0,007059592	0,031441525
15	0,006556774	0,034703476
16	0,0068532	0,036517449
17	0,006869229	0,036944207
18	0,006975038	0,035741524
19	0,006835798	0,035016715
20	0,006801946	0,036829793

5.4.2 Training proposed artificial neural network model

After the number of hidden neurons in hidden layer has been determined to be 8 and 5 for hyperbolic tangent and sigmoid function models respectively, networks have been trained with 60 spare part data (48 training and 12 cross validation data) 2 times each by 10 runs and 1000 iterations with randomized initial weights in order to avoid network to be stuck at local minima. As can be seen from the results of training network with hyperbolic tangent function has minimum mean squared error is of 0,003122 and sigmoid function network has minimum mean squared error of 0,022120 in Table 5.6.

Table 5.6: Results of training for hyperbolic tangent and sigmoid transfer functions

Best Networks	Hyperbolic Tangent	Sigmoid Function
Run #	1	2
Epoch #	490	1000
Minimum MSE	0,003122	0,022120

For network with hyperbolic tangent function, training has been stopped at 1st run and 490th iteration. Similarly, training for network with sigmoid function has been stopped at 2nd run and 1000th iteration.

As a result, we can see that neural network with hyperbolic tangent function has minimum MSE of 0,003122 compared to 0,022120 of the sigmoid function and it is chosen to be most optimal network model for the spare part multi criteria ABC classification. The neural network model with hyperbolic tangent activation function has minimum MSE when hidden layer has eight nodes. The architecture of the proposed artificial neural network model is given in Figure 5.4.

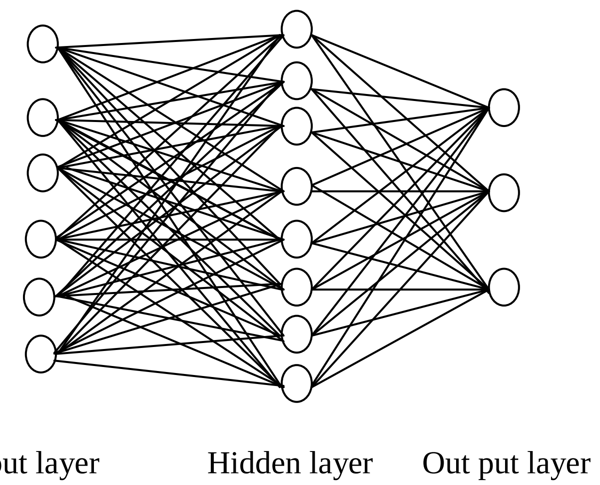


Figure 5.4: Proposed artificial neural network model.

The network chosen consists of 6 nodes in input layer, 8 nodes in hidden layer and 3 nodes in output layer.

5.5 Testing of the Neural Network

In order to validate the training results, 11 spare parts data, given in Table 5.7, previously not introduced to network were used as testing.

Table 5.7: Data for 11 spare part used for testing of the network.

ID no.	Stock-out Penalty	Substitutability	Value usage	Lead Times	Unit Cost	Commonality	A	B	C
61	3	3	2	5	1484,71	8	0	1	0
7	2	5	850	1	0,465	21	0	0	1
71	1	5	1	1,5	105	8	0	0	1
1	5	5	10	3	22	12	0	1	0
3	5	5	6	1	64,53	11	0	1	0
56	5	5	14	4	32,21714	6	0	1	0
62	1	5	3	1	1,22	14	0	0	1
40	1	5	211	2	1,1	3	0	0	1
58	2	5	5	1	5,85	11	0	0	1
53	5	3	2	5	2963,1	5	0	1	0
22	5	1	9	24	4174,138	1	1	0	0

The results of overall testing of the network are given in Table 5.8.

Table 5.8: Results of testing process for proposed neural network.

Results for overal network after testing	
	Hyperbolic tangent
MSE	0,006483
NMSE	0,01037
r	0,9999
% ERROR	1,038739

The results of neural network model for multi criteria ABC classification of 11 testing data can be viewed in Table 5.9. The calculation of the MSE, NMSE, R and % Error values can be in App. B.2.

Table 5.9: Output matrix of the neural network for 11 testing data.

Output / Desired	*A*	*B*	*C*
A	1	0	0
B	0	5	0
C	0	0	5

The output matrix gives the ratio of output of the model and the actual desired data.

The performance of the network over the 3 output desired parameters A, B and C set can be seen in Table 5.10:

Table 5.10: Performance of classification of network for 11 testing data according to 3 desired output sets, a, b and c.

Performance	A	B	C
MSE	0,002802827	0,001439141	0,001760844
NMSE	0,033914208	0,005804534	0,007102071
MAE	0,051057237	0,03160245	0,036232916
Min Abs Error	0,006789058	0,000959325	0,013045936
Max Abs Error	0,055555555	0,055553772	0,055555538
r	0,999999898	0,999969019	1
Percent Correct	100	100	100

As can be seen from performance measurements the correlation, r, for A set is 0,9999, for B set 0, 99996 and for C set is 1,00. The overall correlation (r) of the testing is 0,9999 for the testing data. The mean absolude error for the overall testing network is obtained to be 1,038 % which shows that developed neural network has results in acceptable solution.

The comparison of neural network output and desired data are given in Table 5.11.

Table 5.11: Desired data and output of neural network

A	B	C	A Output	B Output	C Output
0	1	0	-0,0555551	1,05555374	-0,0555548
0	0	1	-0,0555556	0,01518085	0,98695406
0	0	1	-0,0555556	0,01518085	0,98695406
0	1	0	-0,0555551	1,05555377	-0,0555548
0	1	0	-0,0555551	1,05555377	-0,0555548
0	1	0	-0,0551998	1,05206318	-0,0555555
0	0	1	-0,0555556	0,01518085	0,98695406
0	0	1	-0,0555556	0,01518094	0,98695402
0	0	1	-0,0555556	0,01518005	0,98695279
0	1	0	-0,0551977	1,05203963	-0,0555555
1	0	0	0,99321094	0,00095933	-0,0555555

6. CONCLUSION AND RECOMMENDATIONS

Spare parts inventory management has always been a very important part in almost all organizations. In railway systems management, main attention is given to maintenance operations, because it directly influences the performance of railway vehicles and play a crutial role in railway services to provide uninterrupted and high quality service to passengers. Except from preventative maintenance, the demand of spare part for maintenance tasks is generally random; thus, quick and reliable management of the spare parts inventory is an important factor for the successful implementation of the maintenance process. The purpose of this thesis is to investigate multi criteria ABC inventory classification of spare parts in railway systems, based on one method of artificial intelligence, which is an artificial neural network.

The traditional ABC classification method uses only the unit price and the annual usage of inventory items in ranking. Yet, in some cases, the classification done using only these two criteria turns out to be insufficient. On the other hand, the new method enables the integration of multi criteria into ABC classification. Firstly, multiple criteria have been determined from a various criteria mentioned in various literatures and most important ones have been chosen as evaluation criteria for the ABC classification model of spare parts inventory. The classification model on railway systems spare parts have been performed based on the following criteria: annual unit cost, lead time, value usage, substitutability, commonality and stock-out penalty. Then, in order to estimate the priorities among the criteria, the optimum network architecture and network parameters is determined and neural network is built. The chosen network constitutes of one input layer, one hidden layer and one output layer. The suitable network algorithm is chosen as supervised learning network type which is multi-layer perception based on back-propagation algorithm. The number of nodes in the hidden layer have been estimated by five-fold cross validation method, which is a good estimation method widely used in literature. The network have been trained with two different activation functions, which are hyperbolic tangent and sigmoid functions, in order to compare classification accuracies and choose more suitable network model. The number of nodes in the hidden layer has been found to be eight for hyperbolic tangent neural network and five for neural network with sigmoid transfer function. It was found that hyperbolic tangent function gives more precise results with MSE of 0,003122 compared to the MSE of 0,022120 of the sigmoid function and it is chosen to be most optimal network model for the spare part multi criteria ABC classification. The developed artificial neural network architecture has 6-8-3 structure, which constitutes of six nodes in the input layer, eight nodes in the hidden layer and three nodes in the output layer. Six nodes in input layer represents the evaluation criteria of classification model which are annual unit cost, lead time, value usage, subctitutability, commonality and stock-out penalty. The three nodes in output layer are the three group of A,B and C item classes.

The training of network has been done with 60 data of different spare part items and classification efficiency of the developed ANN was tested by 11 data spare part items. The evaluation data was taken from company performing maintenance operations of light railway system.

The comparison of actual data and output of the network is given in Table 5.9.

The testing of developed neural network for classification spare parts items gave mean squared error of the overall network as 0,006483 and overall error as 1,038 %. The results also indicate that neural network has been trained enough which can be seen from the classification accuracy of 99% of 11 data used for testing the developed artificial neural network.

Inventory management in Turkish Railway Systems is not widely investigated area in the literature. The previous researches on multi criteria inventory management have been done using analytical hierarchy process and data envelopment analysis (Çelebi et al., 2008; Sönmeztürk et.al., 2008). The main contribution of this thesis is that different technique has been applied, which is artificial neural network, for classification of spare parts inventory of light railway systems based on multi criteria ABC analysis. The developed network has shown good classification accuracy and this network can be used for classification of other real-world inventory data in railway systems maintenance.

REFERENCES

Aksungur, S. and Kavlak, K., 2009: "Realization Of The Collision-Free Motion Of A Scara Robot In An Environment With Obstacles Using Artifical Neural Networks And Genetic Algorithm", *Journal of Technical-Online*, Vol. **8**, No. 2, pp. 112-126.

Apaydın, E., 2004: Introduction To Machine Learning, The MIT Press, London, England.

Braglia, M., Grassi, A. and Montanar, R., 2004: "Multi-attribute Classification Method for Inventory Management", Journal of Quality in Maintenance Engineering, Vol. **10**, No.1, pp. 55-65.

Bonnett, Clifford F., 2005: Practical Railway Engineering.London,GBR: Imperial College Press.

Cakir, O. and Canbolat, M.S., 2008: "A web-based decision support system for multi-criteria inventory classification using fuzzy AHP methodology", *Expert Systems with Applications*, **35**, 1367–1378

Chen, Y., Li, K.W., Kilgour, D.M. and Hipel, K.W., 2008: "A case-based distance model for multiple criteria ABC analysis", *Computers & OperationsResearch,* **35,** 776 – 796.

Chen, F.L., Chen, Y.C. and Kuo, J.Y., 2009: "Applying Moving Back PropagationNeural Network and Moving Fuzzy Neuron Network to PredicttheRequirement of Critical Spare Parts", *Expert Systems with Applications*, Vol. **36**, No. 7, pp. 4358– 4367.

Chen J-X., 2011: "Peer-estimation for multiple criteria ABC inventory classification, *Computers and Operations Research*, Vol. **38**, No. 12, pp. 1784-1791.

Chu, C.W, Liang, G.S and Liao, C.T., 2008: Controlling Inventory by combining ABC analysis and fuzzy classification, *Computers and Industrial Engineering*, **55**, 841-851.

Çelebi, D. and Bayraktar, D., 2008: "An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information", *Expert Systems with Applications*, **35,** 1698–1710.

Çelebi, D., Bayraktar, D. and Aykaç, D.S., 2008: Multi Criteria Classification for Spare Parts Inventory, CIE.

Çınar, D., 2007:Hidroelektrik enerji üretiminin hibrid bir model ile tahmini, *Master Thesis*, ITU Institute of Science and Technology, Istanbul.

Cohen, M. A and Ernst, R., 1988: "Multi-item classification and genetic inventory stock control policies", *Production and Inventory Management Journal*, Vol. **29,** No.3, pp.6-8.

Cover, T. M., and Hart, P. E., 1967: "Nearest neighbor pattern classification", IEEE Transactions on Information Theory, **13**, pp. 21–27.

Diakoulaki, D., Zopounidis, C., Mavrotas, G. And Doumpos, M., 1999: "The use of a preference disaggregation method in energy analysis and policy making", *Energy–The International Journal*, Vol. **24**, No. 2, pp. 157-166.

Doumpos, M., 2002: Multicriteria Decision Aid ClassificationMethods. Secaucus, NJ, USA: Kluwer Academic Publishers, p. 4.

Elmas, Ç., 2003: Yapay Sinir Ağları, Seçkin Yayınevi, Ankara.

Ernst, R. and Cohen, M.A., 1990: "Operations Related Groups (ORGs): A Clustering Procedure for Production/inventory Systems", *Journal of Operations Management*, Vol. **9**, No.4, pp. 574-598.

Fisher, R.A., 1936: "The use of multiple measurements in taxonomic problems." *Annals of Eugenics,***7**, pp. 179–188.

Flores, B.E. and Whybark, D.C., 1985: "Multiple Criteria ABC Analysis", *Journal of Operations and Production Management*, Vol. **6**, No. 3, pp. 38-46.

Flores, B. E., Olson, D. L., and Dorai, V. K., 1992: "Management of multicriteria inventory classification", *Mathematical and Computer Modeling*, Vol. **16**, No. 12, pp. 71–82.

Guvenir, H.A. and Erel, E., 1998: "Multi-criteria inventory classification using a genetic algorithm", *European Journal of Operational Research*, **105**, pp. 29–37.

Ghate, V.N. and Dudul, S.V., 2010: "Optimal MLP neural network classifier for faultdetection of three phase induction motor", *Expert Systems with Applications*, **37**, pp. 3468–3481.

Haykin, S., 1999: Neural Networks: A Comprehensive Foundation, Prentice Hall, New Jersey, USA.

Hadi-Venchen, A., 2010: "An improvement to multi criteria ABC inventory classification", *European Journal of Operational Research*, **201**, pp. 962-965.

Hadi- Venchen, A. and Mohamadghasemi, A., 2011: "A fuzy AHP-DEA approach for multiple criteria ABC inventory classification", *Expert Systems with applications*, **38**, pp. 3346-3352.

Huiskonen, J., 2001: "Maintenance spare parts logistics: Special characteristics and strategic choices", *International Journal of Production Economics*, **71**, pp. 125-133.

Jacquet-Lagreze, E. and Siskos, Y., 1982: "Assessing a set of additive utility functions for multi-criteria decision making: the UTA method, *European Journal of Operations Research*, Vol. **10**, No.2, pp. 151-164.

Kranenburg, A.A. andVan Houtum, G.J., 2007: "Effect of commonality on spare parts provisioning costs for capital goods", *International Journal of Production Economics*, Vol. **108**, No. 1-2, pp. 221-227.

Kurt, A., 2003: “Flexible manufacturing system design using simulation artificial neural network”, *J. Fac. Eng Arch. Gazi Univ*., Vol. **18**, No. 2, pp. 31-38.

Leenders, M.R., Fearon, H.E. and England, W.B., 1989: Purchasing and MaterialManagement, Irwin Inc.

Li, S.G. and Kuo, X., 2008: “The Inventory Management System for Automobile Spare Parts in a Central Warehouse”, *Expert Systems with Applications*, **34**, pp. 1144-1153.

Lyly-Yrjanainen, J., Lahikeinen, T. and Paranko, J., 2004: “Cost effect of component commonality: Manufacturing cost perspective. NordDesing 2004, Product design in Changing Environment 18-20 August 2004, Tampere, Finland.

Masters, T., 1993: Practical Neural Network recipes in C++, Academic Press, Inc.

Malakooti, B. and Raman, V., 2000: "Clustering ans Selection of Multiple Criteria Alternatives Using Unsupervised and Supervised Neural Networks", *Journal of Intelligent Manufacturing*, **11**, pp. 435-451.

McCulloch, W. and Pitts, W., 1943: "A logical calculus of ideas immanent in neural activity", *Bulletin of Mathematical Biophysics*, **5**, pp. 115-133.

Ng, W.L., 2007: **"**A Simple Classifier for Multiple Criteria ABC Analysis", *European Journal of Operational research*, **177**, pp. 344-353.

Nieddu, L. and Patrizi, G., 2000: "Formalmethods in patternrecognition:A review", *European Journal of Operational Research*, **120**, pp. 459-495.

Paliwal, M., and Kumar, U.A., 2009: "Neural networks and statistical techniques: A review of applications", *Expert Systems with Applications*, **36**, pp. 2–17.

Partovi, F.Y. and Anandarajan, M., 2002: "Classifying inventory using an artificialneural network appoach”, *Computer and Industrial Engineering*, **4**1, pp. 389-404.

Partovi, F. Y., and Burton, J., 1993: "Using the analytic hierarchy process for ABC analysis.", *International Journal of Operations and Production Management*, Vol. **13**, No. 9, pp. 29–44.

Petrovic, D. and Petrovic R. 1992: "Sparta ii: Further development in an expert system for advising on stocks of spare parts", *International Journal of of Production Economics*, Vol. **24**, No. 3, pp. 291-300.

Principle, J.C., Euliano, N.R. and Lefebvre, W.C., 1999: Neural and Adaptive Systems. Fundamentals Through Simulations, John Wiley&Sons, New York, USA.

Qi, M. and Zhang, G.P., 2001: "An investigation of model selection criteria for neural network time series forecasting", *European Journal of Operational Research*, Vol. **132**, No. 3, pp. 666-680.

Ramanathan, R., 2006: “ABC inventory classification with multiple-criteria using weighted linear optimization”, *Computers & Operations Research*, **33,** pp. 695–700.

Reid, R. A., 1987: "The ABC method in hospital inventory management: A practical approach", *Production and Inventory Management Journal*, Vol. **28**, No. 4, pp. 67–70.

Roy, B., 1996: Multi Criteria Methodology for decision aiding. Dordrecht: Kluwer.

Saaty, T.L., 1980: “The Analytic Hierarchy Process, New York, NY:McGraw-Hill.

Siskos, J., Lochard, J. and Lombardo, J., 1984: “A multi criteria decision-making methodology under fuzziness: Application to the evaluation of radiological protection in nuclear power plants”, North-Holland, Amsterdam, pp. 261-283.

Schalkoff, R., 1997: Artificial neural networks, McGraw-Hill International series, New York.

Sharaf, M.A. and Helmy, H.A., 2001: “A classification model for inventory management of spare parts”, Proceedings of 7th International Conference on Production, Industrial Engineering, Design and Control, Alexandria, Egypt, 13-15 February, **1**, pp. 375-382.

Shen, L., Tay, F.E.H, Qu, L. and Shen, Y., 2000: “Fault diagnosis using rough sets theory”, *Computers in Industry*, **43**, pp. 61-72.

Sherbrooke, C.C., 1968: “METRIC: a multi-echelon technique for recoverableitem control”, *Operation Research*, **16**, pp. 122-141.

Simunovic, K., Draganjac, T. and Simunovic, G., 2008: “Application of Different Quantitative Techniques to Inventory Classification”, *Technical Gazatte*, Vol. **15**, No. 4, pp. 41-47.

Slowinski, R., 1995: “Rough set approach to decision analysis”, *AI Expert Magazine*, Vol. **10**, No. 3, pp. 18-25.

Sönmeztürk, G.I., Şafak, M. E., Celebi, D. and Bayraktar, D., 2008: “Bulanık AHY ile Yedek Parça Stoklarının Sınıflandırılması”, Yöneylem Araştırması ve Endüstri Mühendisliği 29.Ulusal Kongresi.

Şahin, Ş.Ö., 2002: Yapay sinir aglari yardimi ile dinamik bir senaryo analizi, *Master Thesis*, ITU, Institute of Science and Technology, Istanbul.

Tosun, S., 2007: Classification with artificial neural networks and data envelopment analysis: application with student performance. *Master Thesis*, ITU Institute of Science and Technology, Istanbul.

Turer, S., Ayvaz, B., Bayraktar, D. and Bolat, B., 2009: “Tedarikçi Değerlendirme Süreci için bir Yapay Sinir Ağı Yaklaşımı: Gıda Sektöründe bir uygulama”,*Endütsri Mühendisliği Dergisi*, Vol. **20**, No. 2, pp. 31-40.

Tsumoto, S., 1998: “Automated extraction of medical expert system rules from clinical databases based on rough set theory”, *Information Sciences*, **112**, pp. 67-84.

Vapnik, V. N., 1995: The nature of statistical learning theory. New York: Springer-Verlag.

Viale, J. D., 1996: Inventory Management : From Warehouse to Distribution Center.Menlo Park, CA, USA: Course Technology Crisp.

Wei, S., Zhang, J. and Li, Z., 1997: "A Supplier-Selecting System Using a Neural Network", IEEE International Conference on Intelligent Processing Systems, pp. 468-471.

Yu, M.C., 2011: "Multi-criteria ABC analysis using artificial-intelligence based classification techniques", *Expert Systems with Application*, **38,** pp. 3416-3421.

ZhangG.P., Patuwo, B.E. and Hu, M.Y., 1998: "Forecasting with artificial neural network networks: The state of the art", *International Journal of Forecasting*, **14**, pp. 35-62.

Zhang, G.P., 2003: **"**Time series forecasting using a hybrid ARIMA and neural network model", *Neurocomputing*, **50**, pp. 159–175.

Zhang, Y. and Wan, X., 2007: "Statistical fuzzy interval neural networks for currency exchange rate time series prediction", *Applied Soft Computing*, **7**, pp. 1149-1156.

Zhou, P. and Fan, L., 2007: "A note on multi-criteria ABC inventory classification using weighted linear optimization", *European Journal of Operational Research,* **182**, pp. 1488–1491.

Zopounidis, C., 1998: Operational Tools in the Management of Financial Risks, KluwerAcademicPublishers, Dordrecht.

APPENDICES

APPENDIX A.1 Evaluation data for Spare parts in Excel file

Table A.1: Evaluation data for spare parts in excel file.

ID no.	Stock-out Penalty	Substitutability	Value usage	Lead Times	Unit Cost (TL)	Commonality	Class
1	5	5	10	3	22	12	B
2	5	5	230	1	4,92	10	B
3	5	5	6	1	64,53	11	B
4	1	3	1000	3	2,99	16	C
5	5	5	170	3	8,59	4	A
6	2	5	400	1	0,213	20	C
7	2	5	850	1	0,465	21	C
8	5	1	5	28	610,131	2	A
9	3	5	500	1	0,38	22	C
10	2	3	33	18	321,7395	18	C
11	2	5	775	1	0,582	23	C
12	2	3	150	4	1,43568	5	B
13	1	5	1	2	65	19	C
14	1	5	25	3	34,5	14	B
15	5	5	2	5	850	3	A
16	2	5	100	3	10,5	7	B
17	3	1	1	12	101,8708	17	C
18	3	3	10	8	87,02062	6	B
19	3	5	250	4	17	13	B
20	5	5	13	3	24	8	B
21	3	5	22	4	110	15	B
22	5	1	9	24	4174,1376	1	A
23	1	5	80	1	0,75	24	C
24	5	3	37	5	0,3	9	B
25	5	5	235	1,5	28,5	1	A
26	5	1	1	16	1257,783	1	A
27	5	1	3	14	7386,3504	2	A
28	3	5	91	3	57	11	C
29	5	1	4	16	11330,6156	3	A
30	5	1	3	16	23500,5892	4	A
31	5	5	3	12	803,112	5	A
32	5	3	9	18	127,935	6	A
33	5	1	3	12	14660	7	A
34	2	5	3	2	43,488	12	C
35	5	1	1	12	15075,4436	8	A
36	5	1	1	12	836,4884	9	A
37	1	5	14	2	50,92	13	C

Table A.1: Evaluation data for spare parts in excel file (contd.)

ID no.	Stock-out Penalty	Substitutability	Value usage	Lead Times	Unit Cost (TL)	Commonality	Class
38	5	3	2	8	766,774	10	B
39	3	1	54	12	167,22	2	B
40	1	5	211	2	1,1	3	C
41	4	3	5366	10	7,837	1	A
42	5	1	2	30	24255,294	1	A
43	1	5	40	1	7,25	2	C
44	1	5	2100	1	0,97	3	C
45	5	1	1	12	7797,4206	1	A
46	3	1	15	14	2529,271	2	B
47	5	1	20	14	403,705	2	A
48	5	1	20	14	350,406	3	A
49	5	1	20	14	447,741	1	A
50	1	5	4	2	25,5	13	C
51	5	2	140	3,5	15,2	3	A
52	3	3	32	5	4,237	10	C
53	5	3	2	5	2963,1	5	B
54	5	3	3	5	71,972416	1	A
55	5	5	16	3,5	293	2	A
56	5	5	14	4	32,217138	6	B
57	5	3	14	2	196,36	4	B
58	2	5	5	1	5,85	11	C
59	1	3	3	4	175,02	9	C
60	3	1	6	5	418,3	7	B
61	3	3	2	5	1484,71	8	B
62	1	5	3	1	1,22	14	C
63	1	5	20	1	5,6	12	C
64	5	2	157	14	60,211025	1	A
65	1	3	11	10	138,25	2	B
66	1	5	40	1	4,352	7	C
67	3	3	1	8	1337,7	5	B
68	3	5	9	10	350	4	B
69	1	1	337	30	325,325	3	B
70	1	5	217	4	5,5	6	C
71	1	5	1	1,5	105	8	C

APPENDIX A.2 Randomized training data divided into 5 disjoint subsets.

Table A.2: Randomized training data divided into 5 disjoint subsets.

Set	ID no.	Stock-out Penalty	Substitutability	Value usage	Lead Times	Unit Cost	Commonality	A	B	C
1	47	5	1	20	14	403,705	2	1	0	0
	44	1	5	2100	1	0,97	3	0	0	1
	65	1	3	11	10	138,25	2	0	1	0
	70	1	5	217	4	5,5	6	0	0	1
	66	1	5	40	1	4,352	7	0	0	1
	18	3	3	10	8	87,02062	6	0	1	0
	43	1	5	40	1	7,25	2	0	0	1
	48	5	1	20	14	350,406	3	1	0	0
	38	5	3	2	8	766,774	10	0	1	0
	20	5	5	13	3	24	8	0	1	0
	26	5	1	1	16	1257,783	1	1	0	0
	68	3	5	9	10	350	4	0	1	0
2	32	5	3	9	18	127,935	6	1	0	0
	4	1	3	1000	3	2,99	16	0	0	1
	10	2	3	33	18	321,7395	18	0	0	1
	55	5	5	16	3,5	293	2	1	0	0
	17	3	1	1	12	101,8708	17	0	0	1
	6	2	5	400	1	0,213	20	0	0	1
	59	1	3	3	4	175,02	9	0	0	1
	15	5	5	2	5	850	3	1	0	0
	9	3	5	500	1	0,38	22	0	0	1
	60	3	1	6	5	418,3	7	0	1	0
	36	5	1	1	12	836,4884	9	1	0	0
	64	5	2	157	14	60,211025	1	1	0	0
3	42	5	1	2	30	24255,294	1	1	0	0
	39	3	1	54	12	167,22	2	0	1	0
	33	5	1	3	12	14660	7	1	0	0
	13	1	5	1	2	65	19	0	0	1
	28	3	5	91	3	57	11	0	0	1
	50	1	5	4	2	25,5	13	0	0	1
	2	5	5	230	1	4,92	10	0	1	0
	5	5	5	170	3	8,59	4	1	0	0
	29	5	1	4	16	11330,616	3	1	0	0
	25	5	5	235	1,5	28,5	1	1	0	0
	34	2	5	3	2	43,488	12	0	0	1
	11	2	5	775	1	0,582	23	0	0	1

Table A.2: Randomized training data divided into 5 disjoint subsets (contd.)

Set	ID no.	Stock-out Penalty	Substitutability	Value usage	Lead Times	Unit Cost	Commonality	A	B	C
4	54	5	3	3	5	71,972416	1	1	0	0
	69	1	1	337	30	325,325	3	0	1	0
	51	5	2	140	3,5	15,2	3	1	0	0
	30	5	1	3	16	23500,589	4	1	0	0
	46	3	1	15	14	2529,271	2	0	1	0
	67	3	3	1	8	1337,7	5	0	1	0
	21	3	5	22	4	110	15	0	1	0
	23	1	5	80	1	0,75	24	0	0	1
	16	2	5	100	3	10,5	7	0	1	0
	14	1	5	25	3	34,5	14	0	1	0
	8	5	1	5	28	610,131	2	1	0	0
	12	2	3	150	4	1,43568	5	0	1	0
5	41	4	3	5366	10	7,837	1	1	0	0
	57	5	3	14	2	196,36	4	0	1	0
	45	5	1	1	12	7797,4206	1	1	0	0
	52	3	3	32	5	4,237	10	0	0	1
	63	1	5	20	1	5,6	12	0	0	1
	27	5	1	3	14	7386,3504	2	1	0	0
	37	1	5	14	2	50,92	13	0	0	1
	49	5	1	20	14	447,741	1	1	0	0
	35	5	1	1	12	15075,444	8	1	0	0
	24	5	3	37	5	0,3	9	0	1	0
	31	5	5	3	12	803,112	5	1	0	0
	19	3	5	250	4	17	13	0	1	0

APPENDIX B.1 Results of five fold cross validation method

Table B.1: Minimum mse for training data for five subsets with different hidden node numbers.

No. Node	Set No.	Min MSE
2	1	0,08815
	2	0,06739
	3	0,13478
	4	0,0445
	5	0,09115
3	1	0,117
	2	0,10056
	3	0,13108
	4	0,04473
	5	0,12695
4	1	0,09932
	2	0,11548
	3	0,1527
	4	0,04596
	5	0,0788
5	1	0,10625
	2	0,14047
	3	0,14856
	4	0,05011
	5	0,0736
6	1	0,14437
	2	0,1707
	3	0,12214
	4	0,03823
	5	0,06723
7	1	0,11787
	2	0,19249
	3	0,1443
	4	0,07113
	5	0,09774

Table B.1: Minimum mse for training data for five subsets with different hidden node numbers (contd.).

No. Node	Set No.	Min MSE
8	1	0,07779
	2	0,19322
	3	0,09057
	4	0,10379
	5	0,12241
9	1	0,1219
	2	0,14422
	3	0,16835
	4	0,07616
	5	0,08637
10	1	0,12905
	2	0,14588
	3	0,14726
	4	0,02551
	5	0,06343
11	1	0,15411
	2	0,16615
	3	0,16694
	4	0,02693
	5	0,09915
12	1	0,11062
	2	0,17151
	3	0,12117
	4	0,11553
	5	0,08794
13	1	0,16306
	2	0,20393
	3	0,14529
	4	0,13174
	5	0,11075
14	1	0,15419
	2	0,17794
	3	0,14864
	4	0,05335
	5	0,11996

Table B.1: Minimum mse for training data for five subsets with different hidden node numbers (contd.).

No. Node	Set No.	Min MSE
15	1	0,1487
	2	0,19374
	3	0,16906
	4	0,10186
	5	0,10925
16	1	0,13792
	2	0,16596
	3	0,13202
	4	0,12651
	5	0,09286
17	1	0,11358
	2	0,14936
	3	0,17715
	4	0,1146
	5	0,12627
18	1	0,15
	2	0,22412
	3	0,2271
	4	0,10334
	5	0,11343
19	1	0,15252
	2	0,15401
	3	0,201
	4	0,12044
	5	0,06489
20	1	0,14244
	2	0,17321
	3	0,12773
	4	0,08764
	5	0,11971

Table B.2: Minimum mse for cross validation data for five subsets with different hidden node number.

No. Node	Set No.	Min MSE
2	1	0,005085
	2	0,004325
	3	0,001183
	4	0,140905
	5	0,148789
3	1	0,000588
	2	0,090422
	3	0,001209
	4	0,154522
	5	0,154322
4	1	0,001157
	2	0,001736
	3	0,001411
	4	0,181418
	5	0,131251
5	1	0,003889
	2	0,076077
	3	0,00256
	4	0,112092
	5	0,099379
6	1	0,005218
	2	0,012562
	3	0,002006
	4	0,097729
	5	0,139008
7	1	0,003707
	2	0,090419
	3	0,005994
	4	0,094414
	5	0,195539

Table B.2: Minimum mse for cross validation data for five subsets with different hidden node number (contd.)

No. Node	Set No.	Min MSE
8	1	0,003047
	2	0,004106
	3	0,002985
	4	0,094213
	5	0,082456
9	1	0,004907
	2	0,004451
	3	0,003486
	4	0,105173
	5	0,234503
10	1	0,002915
	2	0,054784
	3	0,003956
	4	0,104444
	5	0,085984
11	1	0,002936
	2	0,004287
	3	0,039104
	4	0,097593
	5	0,192136
12	1	0,003986
	2	0,004837
	3	0,004139
	4	0,087576
	5	0,170837
13	1	0,079069
	2	0,054462
	3	0,003323
	4	0,127702
	5	0,23195
14	1	0,003745
	2	0,07637
	3	0,023162
	4	0,099809
	5	0,236425

Table B.2: Minimum mse for cross validation data for five subsets with different hidden node number (contd.)

No. Node	Set No.	Min MSE
	2	0,054655
	3	0,054023
	4	0,193675
	5	0,156212
16	1	0,003561
	2	0,077017
	3	0,063359
	4	0,216191
	5	0,187609
17	1	0,004463
	2	0,093997
	3	0,004591
	4	0,200056
	5	0,154228
18	1	0,054184
	2	0,00446
	3	0,054077
	4	0,094209
	5	0,208318
19	1	0,020275
	2	0,084552
	3	0,004016
	4	0,154283
	5	0,136505
20	1	0,004325
	2	0,054723
	3	0,004049
	4	0,189607
	5	0,189766

Table B.3: Calculation of average minimum mean squared error for two node of cross validation data.

NO. NODE	SET NO.	MIN MSE
2	1	0,00508
	2	0,00432
	3	0,00118
	4	0,14091
	5	0,14879

Average minimum MSE for node 2=$\frac{\sum_{i=1}^{5} MSE_{min}}{5}$ i=1,2,3,4,5:set number

Average minimum MSE for node 2=

$$\frac{(0{,}00508+0{,}00432+0{,}00118+0{,}14091+0{,}14879)}{5} = 0{,}060057349$$

Calculation of average minimum MSE for other nodes is done in similar way and the results are given in Table 5.5 and Teble 5.6 for tangent and sigmoid functions respectively.

APPENDIX B.2 Formulas for calculation of network performance measurements.

MSE: The mean squared error is simply two times the average cost (see the access points of the ErrorCriterion component.) The formula for the mean squared error is:

$$MSE = \frac{\sum_{j=0}^{P} \sum_{i=0}^{N} \left(d_{ij} - y_{ij}\right)^2}{N\,P}$$

Where P=number of output processing elements

N=number of exemplars in the data set

y_{ij}=network output for exemplar I at processing element j

d_{ij}=desired output for exemplar I at processing element j

NMSE: The normalized mean squared error is defined by the following formula:

$$NMSE = \frac{P\ N\ MSE}{\sum_{j=0}^{P} \frac{N \sum_{i=0}^{N} d_{ij}^{\ 2} - \left(\sum_{i=0}^{N} d_{ij}\right)^2}{N}}$$

Where P=number of output processing elements

N=number of exemplars in the data set

MSE=mean squared error

d_{ij}=desired output for exemplar I at processing element j

r: the correlation coefficient between a network output x and and a desired output d is:

$$r = \frac{\frac{\sum_i (x_i - \bar{x})(d_i - \bar{d})}{N}}{\sqrt{\frac{\sum_i (d_i - \bar{d})^2}{N}} \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{N}}}$$

The correlation coefficient is confined to the range [-1,1]. When r =1 there is a perfect positive linear correlation between x and d, that is, they covary, which means that they vary by the same amount. When r=-1, there is a perfectly linear negative correlation between x and d, that is, they vary in opposite ways (when x increases, d decreases by the same amount). When r =0 there is no correlation between x and d, i.e. the variables are called uncorrelated. Intermediate values describe partial correlations. For example a correlation coefficient of 0.88 means that the fit of the model to the data is reasonably good.

% Error: The percent error is defined by the following formula:

$$\%Error = \frac{100}{N\ P} \sum_{j=0}^{P} \sum_{i=0}^{N} \frac{\left|dy_{ij} - dd_{ij}\right|}{dd_{ij}}$$

Where P=number of output processing elements

N=number of exemplars in the data set

dy_{ij}=denormalized network output for exemplar i at processing element j

dd_{ij}= denormalized desired output for exemplar i at processing element j

APPENDIX B.3 Network building with Neurosolutions software.

The selection of training, cross validation, testing, input and desired data on Excel worksheet can seen in the Figure B.1.

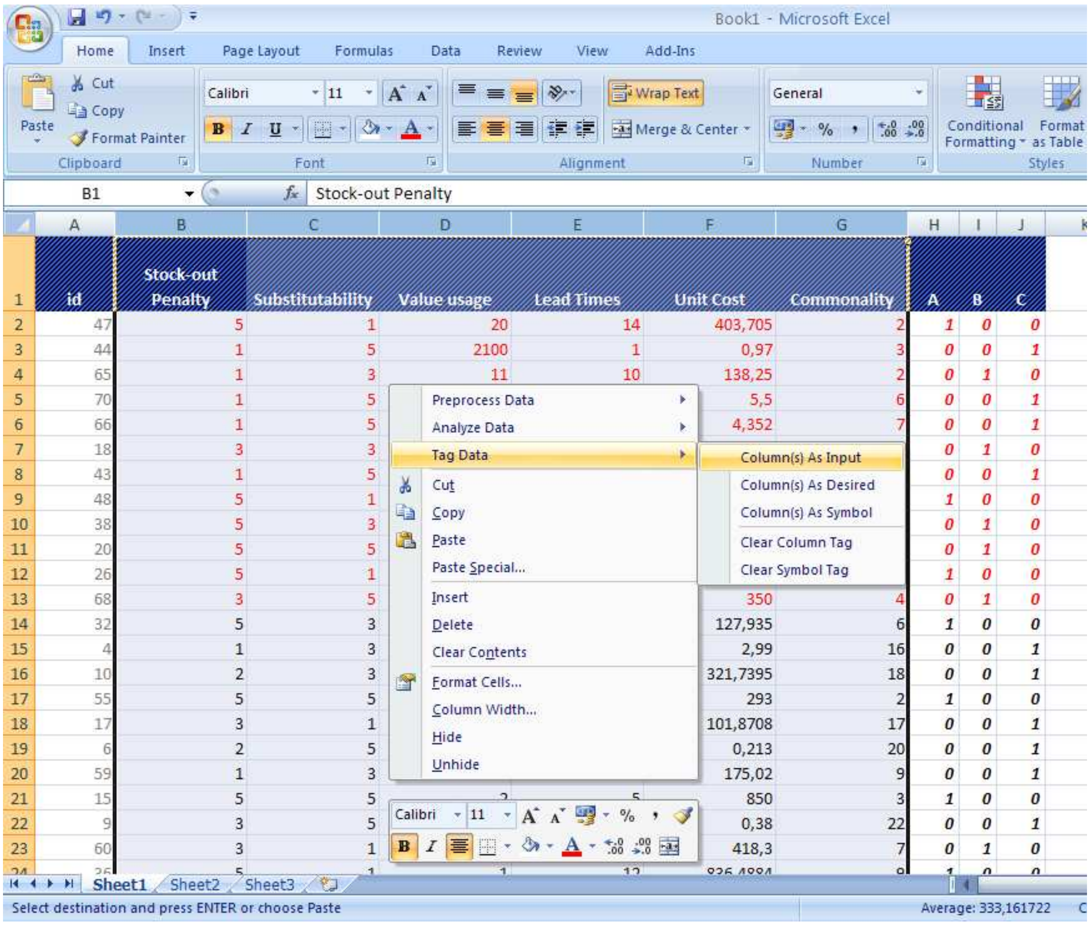


Figure B.1: Selection of input and desired parameters in Excel.

After selecting 6 input and 3 output criteria, by similar way 48 data is chosen as training, 12 of data is selected as cross validation and 11 data for spare parts is chosen as testing data for the network model. Then Neurobuilder tool from Neurosolutions menu was chosen to design the network. The starting menu of Neurobuilder can be seen in Figure B.2.

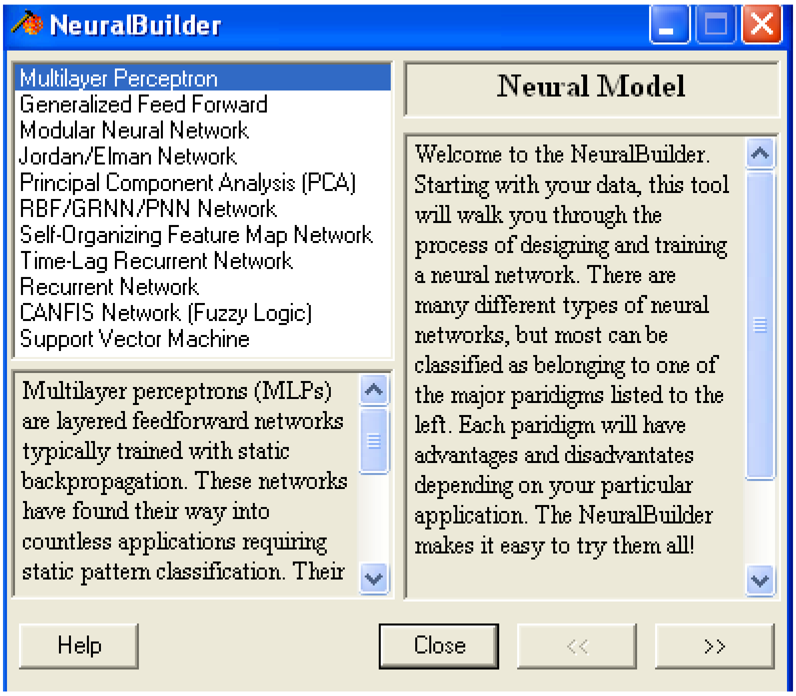


Figure B.2: NeuroBuilder tool.

In the starting menu the learning model for neural network should be chosen. After selecting the learning model of neural network >> sign is ticked and in the next section number of hidden layer is selected as can be seen in Figure B3:

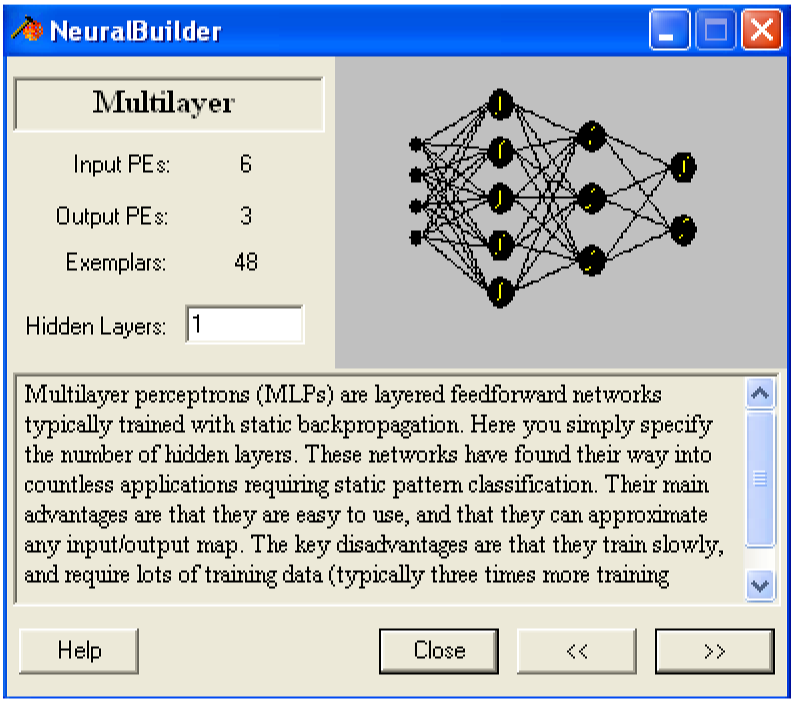


Figure B.3: Selection of Hidden Layer number in Neurosolutions.

After selecting the hidden layer number, the number of hidden neurons, activation function, step size and momentum rate in the hidden layer is selected as can be seen in Figure B.4:

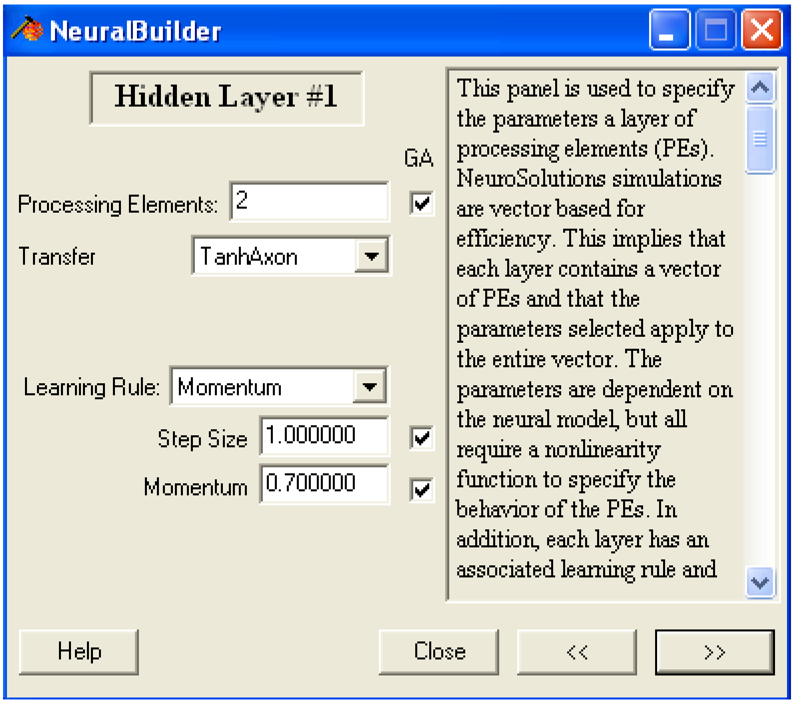


Figure B.4: Selection of training parameters for Hidden Layer in Neurosolutions.

In the similar way, the training parameters like activation function, step size and momentum are selected for the output layer as can be viewed in Figure B.5 below.

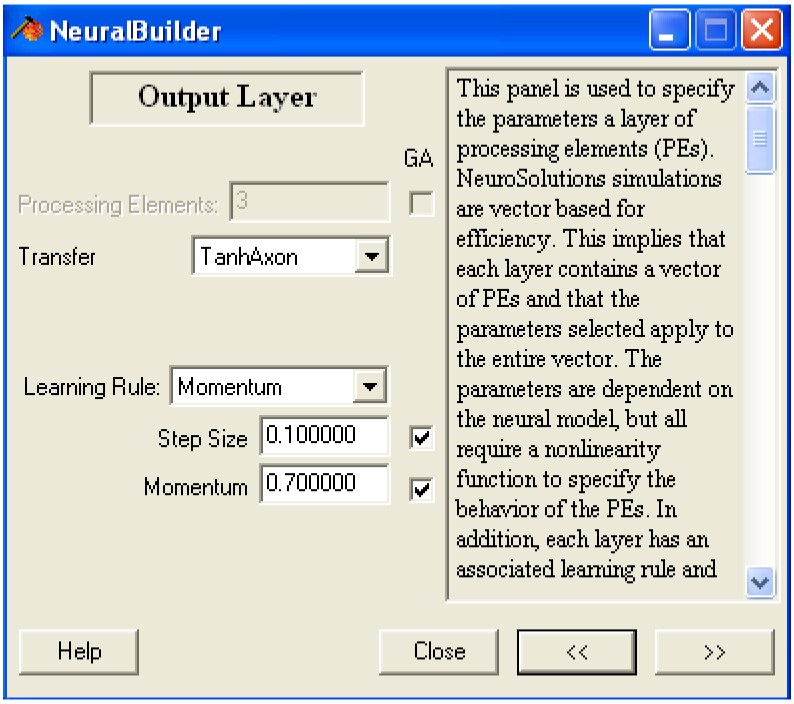


Figure B.5: Selection of training parameters for Output Layer in Neurosolutions.

In the least step of network building in Neurosolutions,in Figure B.6, the maximum number of epochs, stopping performance criteria and weight update methods are selected.

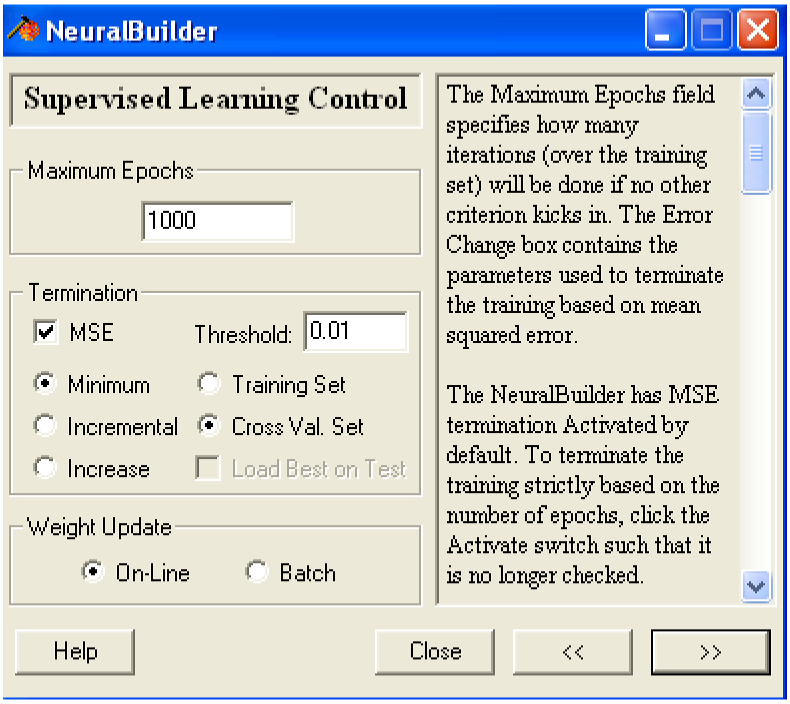


Figure B.6: Selection of performace criteria in Neurosolutions.

After choosing all required parameters with the network structure, probe configuration, in Figure B.7, is specified is order to view required solutions of the training.

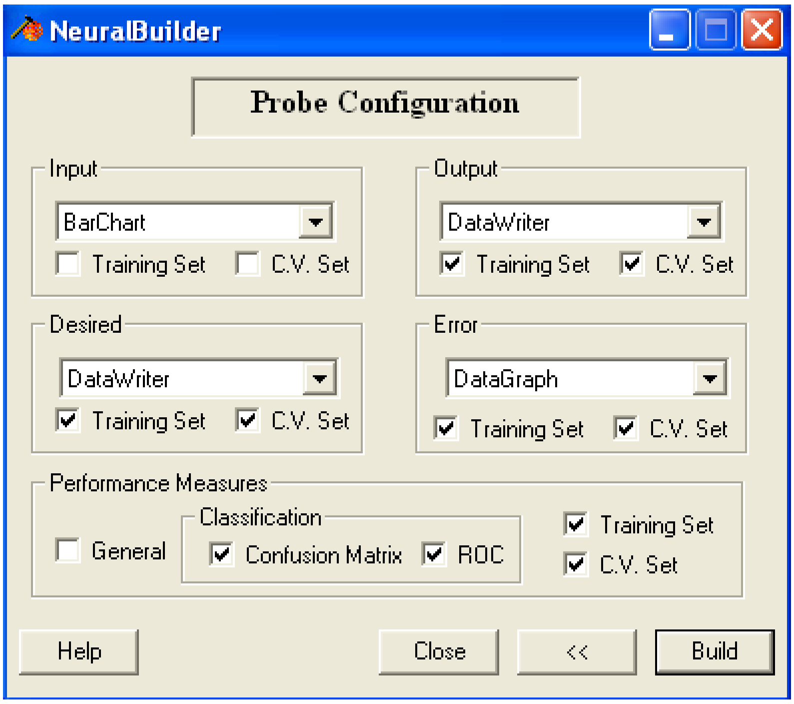


Figure B.7: Setup for Probe Configuration in Neurosolutions.

Finally, after finishing selection of network structure and learning parameters, “Build” key is clicked and network is built as can seen in Figure B.8.

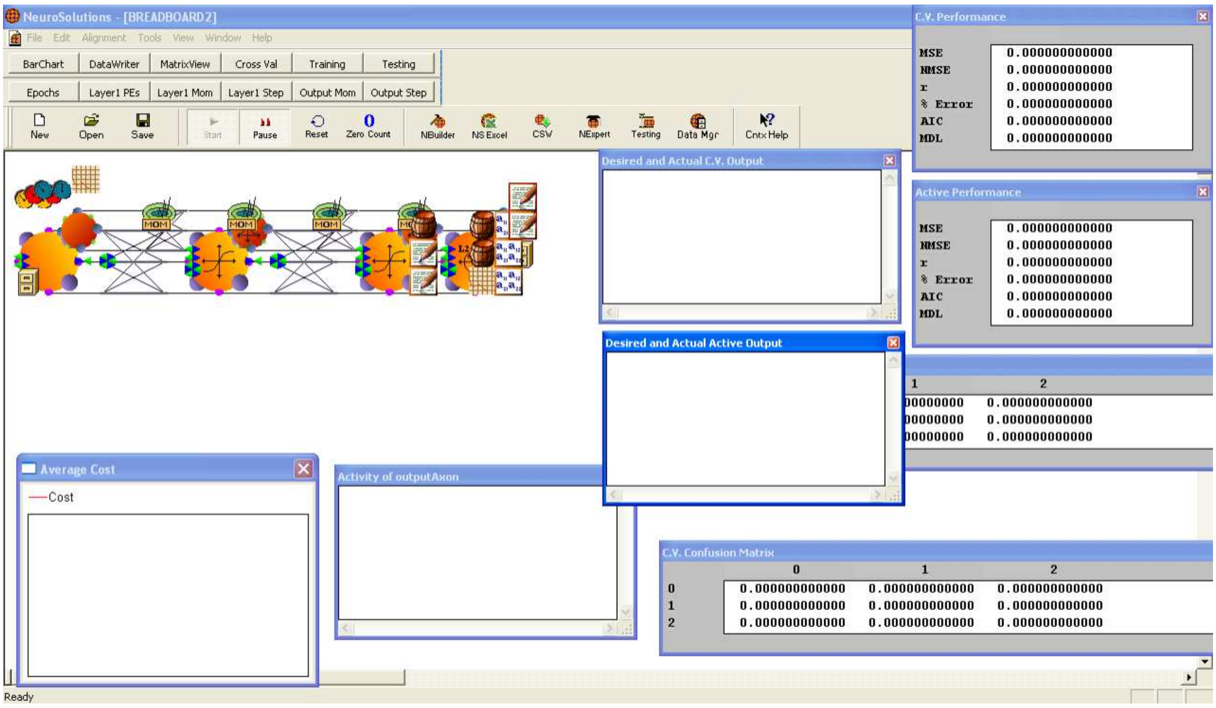


Figure B.8: NeuroSolutions Network Structure.

Training of the network is started from NeuroSolutions tool in Excel, the number of max epochs and runs is specified in the opened icon, which is displayed below in Figure B.9.

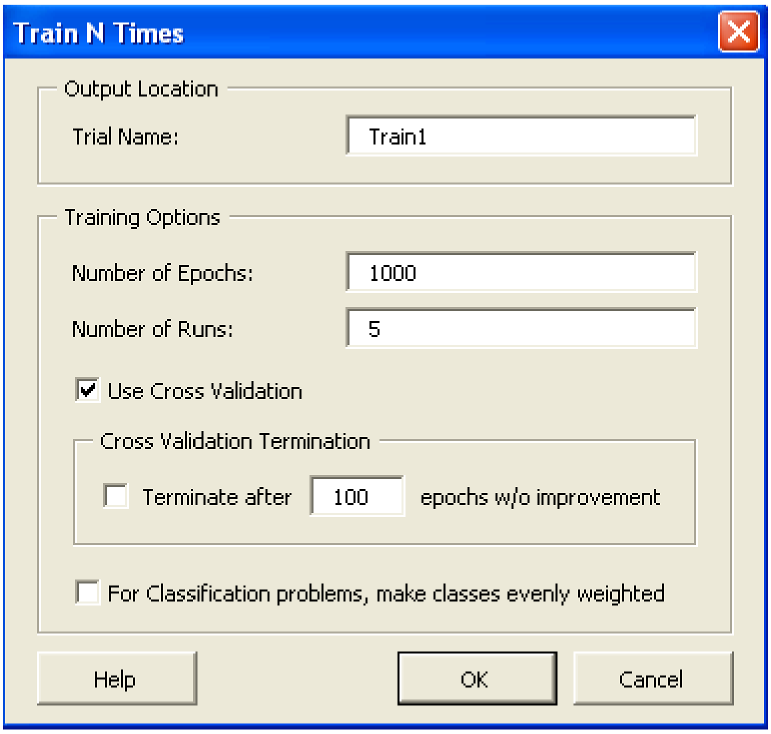


Figure B.9: Training Process in Neurosolutions.

CURRICULUM VITAE

Candidate's full name: JarkynJumabaeva

Place and date of birth: Bishkek-Kyrgyzstan 18.08.1985

Permanent Address: Ogombaeva Street, 179. Bishkek, Kyrgyzstan

Universities and
Colleges attended: Middle East Technical University