T.C.

ALTINBAŞ ÜNİVERSİTESİ

Electrical and Computer Engineering

PATTERN RECOGNITION USING NEURAL NETWORKS

Oday Mohammed Ahmed

Master Thesis

Prof. Dr. Oğuz Bayat

İstanbul, 2019

PATTERN RECOGNITION USING NEURAL NETWORKS

by

Oday Mohammed Ahmed

Electrical and Computer Engineering

Submitted to the Graduate School of Science and Engineering

in partial fulfillment of the requirements for the degree of

Master of Science

ALTINBAŞ UNIVERSITY

2019

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

_____

Prof. Dr. Oğuz BAYAT

Supervisor

Examining Committee Members (first name belongs to the chairperson of the jury and the second name belongs to supervisor)

Prof. Dr. Osman Nuri UÇAN	School of Engineering and Natural Sciences, Altinbas University	_____
Prof. Dr. Oğuz BAYAT	School of Engineering and Natural Sciences, Altinbas University	_____
Prof. Dr. Hasan Hüseyin BALIK	Air Force Academy, National Defense University	_____

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

_____

ASST. PROF. DR. ÇAĞATAY AYDIN

Head of Department

Approval Date of Graduate School of Science and Engineering: ____/____/____

_____

Prof. Dr. Oğuz BAYAT

Director

DEDICATION

To the one who made me learn how to succeed, to the one who I miss him in the middle of difficult times, to the one who the life decided to take him before I have enough from his caring, to my father may he rest in peace. And to the one who my word racing to describe her caring, to the one who tried her best to teach me, support me, and did the best to make me to what I am now, to the who I swim in the sea of tenderness that she has to cure my wounds, To my mom who I ask Allah to give the longest healthy life. To my brother and sisters. To everyone who at least taught me at least a letter …

I give you this thesis in a form of appreciation and I may ask Allah to accept this work.

ACKNOWLEDGEMENTS

I wish to express my acknowledgements to my supervisor, Prof. Dr. Oğuz Bayat who was abundantly helpful and offered invaluable support with his sincerity and belief in me.

ABSTRACT

PATTERN RECOGNITION USING NEURAL NETWORKS

Ahmed, Oday Mohammed

M.Sc., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Prof. Dr. Oğuz Bayat

Date: June 2019

Pages: 51

Facial recognition is also one of the main research topics due to its different applications such as security systems, medical systems, entertainment, etc. Face recognition: natural, robust, and non-intrusive. The preferred method for human identification. To confirm or determine the identity of the applicant, a wide variety of systems demand reliable personal identification schemes. The aim of these plans is to ensure that the services rendered are available to only a legitimate user and nobody else. Secure access is provided, for example, to buildings, computer systems, laptops, mobile telephones and cameras. In the absence of robust personal recognition systems, these systems are vulnerable to the will of an impostor. The human-face identification system using artificial neural networks has been developed and shown in this article that the visual recognition rate of 40 people is 87.5 percent for the 400 frame values in the AT&T database.

Keywords: Pattren recognition, Neural network, Kohonen, Self-Organized map, Classification.

ÖZET

SİNİR AĞLAR KULLANARAK ÖRNEK TANIMA

Ahmed, Oday Mohammed

Yüksek Lisans, Elektrik ve Bilgisayar Mühendisliği, Altınbaş Universitesi,

Danışman: Prof. Dr. Oğuz Bayat

Tarih: Haziran 2019

Sayfalar: 51

Güvenlik sistemleri, tıbbi sistemler, eğlence vb. Çeşitli uygulamaları nedeniyle yüz tanıma da ana araştırma konularından biri olarak tanımlanmıştır. Tercih edilen insan tanımlama yöntemi yüz tanıma yöntemidir: doğal, sağlam ve müdahaleci olmayan. Çok çeşitli sistemler, talep edenin kimliğini onaylamak veya belirlemek için güvenilir kişisel tanımlama şemaları gerektirir. Bu programların amacı, yalnızca meşru bir kullanıcının ve başka hiç kimsenin sunulan hizmetlere erişmemesini sağlamaktır. Örneğin, binalara, bilgisayar sistemlerine, dizüstü bilgisayarlara, cep telefonuna ve ATM'lere güvenli erişim dahildir. Bu sistemler, sağlam kişisel tanıma sistemleri yokluğunda bir sahtekârın iradesine karşı savunmasızdır. Bu makale, yapay sinir ağları kullanan insan yüz tanıma sistemini geliştirdi ve gösterdi; bu, 40 birey için yüz tanıma oranının, AT&T veritabanında yüzde 87,5 ile 400 kare için sonuç gösterdiğini gösteriyor.

Anahtar kelimeleri: Örnek tanıma, Sinir ağı, Kohonen, Öz-Organize harita, Sınıflandırma.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

1. INTRODUCTION

1.1 BACKGROUND

Growing global markets and booming e-commerce could help competitors improve warehouse efficiencies with growth in the horizon alone. Efficient warehouse transaction also enables new companies to benefit from the ever-expanding pick-up of warehouses, as customers around the world are increasingly needed to reduce their costs to stay competitive. In warehouses there are many crucial operations, including orders fulfillment, reclassification and shipping. Bartholdi and Hackman state that the most important warehouse operation is the collection of orders [1]. The collection is the searching and preparation of the ordered product prior to shipment. The focus is on selecting the main time requirement to improve the operation of the most important warehouse.

The most frequent time a worker spends on a job is the travel time when selecting products [1]. Touring all aisles, harvesting areas and even finding storage units of warehouses (SCU's) around the warehouse according to the pickup plan. Travel can also be referred to as the most feasible step [1]. Since many people move to different products where SKU is located, total system performance, employee use and direct picking costs can be dramatic. A pickler also has direct connection to the number of journeys performed by a picker in the storage. For the efficiency of a warehouse, choice systems play a crucial role.

One of the most popular selection schemes today is to select areas. An area is an area for which a picking employee selects the SKUs within its areas. The allocation of workmen to areas can improve the selection of performance measures. The employees are generally allocated to areas in front of the warehouse, where space and inventory are very small compared to the whole warehouse. The same applies to the reserve area, containing the remaining SKU inventory. By placing high-use SKUs in the forward field, it reduces the travel time for employees. This has to do with reducing the journey time due to the fact that the more popular SKUs are located in a little area and greater proximity to the shipping and receivers. This thesis would most likely apply to quickly selected warehouses in which containers visit each area before leaving the system. The

approach is applied. The fast-selection shops have an area with less popular SKUs maintained in smaller quantities in order to speed up the selection process [1].

Alongside the methodology for future reservations / areas in which each SKU is allocated. SKU placement in designated warehouse locations is called slotting. When picking up activities, time and money can be saved by strategically adapting the warehouses to the company's order profiles. As mentioned above, one of the strategies used in warehouses is the use of working areas. Slotting also has to deal with a number of methods and problems, and that is why it is studied so well. If a specific warehouse dedicated to search time savings are available, changing demand and order profiles can be used to avoid travel savings. In addition, literature also mentions the concept of re-slotting, a SKU movement based on new information on the order profile. Another topic is not just where, but when practitioners should redeploy to benefit from travel savings. The placement of a slotted warehouse is not identical to the allocation of an random assignments will influence the routing schedules. The slotting concept was described by Yingde and Smith based on the correlations of order frequency [2]. If, for instance, SKU A is controlled 90% of the time by SKU B, these products are kept in the same region. This reduces the number of visited theoretical zones and reduces the time required to choose all the orders in both SKUs and comply with them.

Efficient slotting was observed to optimize picking and maintain available and readily accessible most popular items for quicker ordering [3]. A slot can help a few things right. If SKUs are logically established mathematically whether or not they move quickly or rarely, the organization of any warehouse will be more efficient. Although devoted storage can be performed at random, the systemic implementation of the collection strategy / order can lead to optimum picking operations. For instance, if shipping and receipt are closest to fastest moving products, they have a significantly greater turnover and lower photo cost due to the reduction in travel times.

Proper slots will enable all working categories, including the handling of materials, collector and management required for the operation to be utilized efficiently and effectively. This is achieved through improved collection, travel reduction and efficient warehouse management. The location of the slot has many aspects. Depending upon the time of year and promotion currently available,

some high-season products do not need to be in the forward and booking areas. For example, in the third and fourth quarters of the year, decoration in Christmas is much higher than after the beginning of the year.

Some practically and theoretically applicable slotting methods have been adopted. The difference is, however, that savings are lost by rearranging or "re-storing" SCUs; for decades the slotting concept has been studied for warehouse operations and will be discussed in depth in the literature review. The theory shows that better results are obtained;

1.2 MOTIVATION

The methodology of this thesis may remove a few limitations from current slotting. The first thing that needs to be done, as stated earlier, is to pay attention to what these SKUs do not order. The benefit of SKU alignment is that savings on travel can be seen to reduce travel distances within each area when product is mathematically slotted and clustered. The re-slotting problem versus the expected savings are yet another limitation of the current slotting methods. Many warehouse solutions are more costly than travel savings, but make no sense in reorganizing them. In the literature review sections, these examples are examined in more detail. This thesis permits the rearrangement of new data and the grouping through the self-organizing map at any time after addition.

Space and working time are maximized in the future ideal state. Looking at the ordering of the goods, the warehouses can improve the ordering instead of what is ordered (pure request). The products are shipped together.

A SOM is an attractive clustering method because the number of partitions for massive datasets can help find the best solution. Artificial Neural networks are very sturdy. It is well-suited for larger data sets that provide better solutions, neural data networks with large amounts of data

perform much better and is ideal for warehouses with different product requirements as well as numerous order details.

The usage and workload balance for SOMs are considered where the clusters could excessively weigh. The balance of workload and travel distances must be considered in order to determine correct zones. Areas within a forward area refer to Picker boundaries. The visual nature of the output data is one advantage of SOMs. You can easily see the SKUs ordered and the SKUs not ordered. Visible lines for output may be drawn on the basis of the SOM clusters. In Figure 1 the above appears.

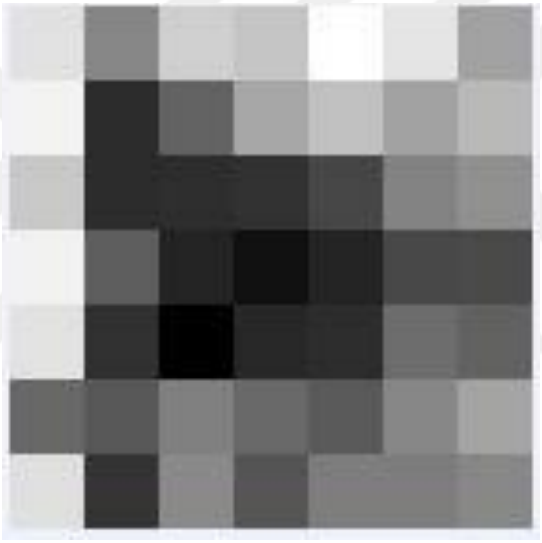

Figure 1.1: SOM visual output example

Figure 1.1 is an instrument SOM, and each square is a node. Figure 1.1 Topographic SOM dimensions are determined by the processing element, neuron or node. 100 working elements, neurons or nodes are provided when a 10x10 matrix is used. Light denser nodes are ordered more often together, and those ordered less frequently are darker. This figure facilitates the viewing of certain clusters within the data. In this paper, methods for SOM clustering SKU with other methods are developed and compared. The efficiency of the methodology can be assessed by comparing

categorization and area development with pick speed. The information on the order profile or which a "Command Similarity Grouping" can be used to develop an optimized slotting method.

In fact, money determines usually how and which changes are to be made. An undertaking can use technology to help select products through automation. These typically require a considerable investment, however. If some algorithms modified conventional picking, a low change in process could lead to the maximum savings. By reducing the dramatic changes, it would be easier to obtain a worker buy-in, making everyone happier when management achieves better savings process.

2. LITERETURE REVIEW

2.1 SLOTTING

As selection time is mostly a way to improve the pick efficiency, the handling of materials would be a good way. Many warehouses use techniques to reduce the gap. Special and different storage methods (specific storage, AS / RS, etc) and routing technology: Stock-Keeping Unit (SKU) positioning in the ware home, the slotting. This includes positioning and direction of the SKU; type of racks;

Many factors must be taken into account to correctly settle in a certain warehouse. Bond quotes a number of other factors, including product weight, size, speed, receipt, shipment, product marketing, and many others, that interact with slot optimization [3, 4]. The slotting sections for products have to be reassessed following a set time period, usually quarterly, given the fluctuation in demand. The techniques presented in this literature review are mathematical heuristics and optimization.

2.1.1 Heuristics

Cube by order index (COI) is commonly used as one of the oldest known slotting methods for many other methods in the literature. The Haskett [5] cube by order method takes account of popularity, SKU and slot density. The writers discussed in their Petersen and Kelley study [6] how effective slot methodology is used and how it further influences the performance of slot scooters. By changing factors and limitations, Petersen and Kelley [6] were attempting to achieve the best results. Depending on how the products are stored in the warehouse, different performance measures may lead to different results.

The research showed that the COI is best suited to the internal strategy: the popular SKUs are densely packaged and are close to their delivery and acceptance site. For the trans-aisle and diagonally, SKU popularity worked best. Figure 2.1 shows this situation.

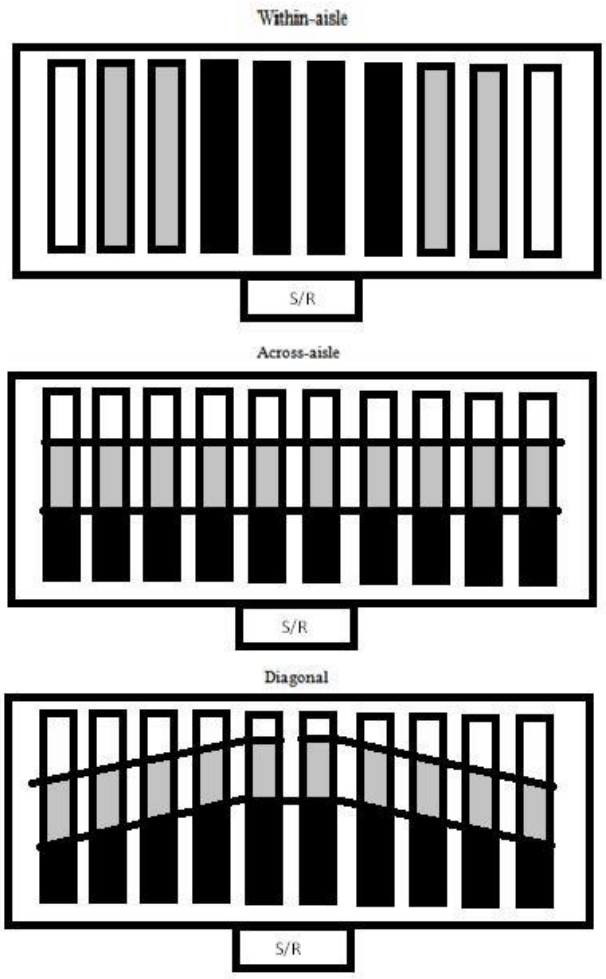

Figure 2.1: Various layouts of COI with various densities of SKU

A lower ICS, indicating more COI indexation as well as shipping and receiving point are the black shaded sections on the islands. This study seems to resolve the need for innovative and detailed methods in a relatively straightforward warehouse design.

Several years later, Mantel, Schuur and Heragu [7] developed a new strategy to attract SKUs to the best places in the warehouse. The COI was comparable with the order-oriented slot technology. The following assumptions were included in the order-based slotting (OOS): an empty warehouse, several waiting orders and various routing approaches. You are going through the convention on warehouse setup with some common features. Aisles are one aisle for each two racks parallel and

double-sided. SKUs are all of the same shape and size, distance in the horizontal direction is to be minimized and the first shipping / receiving point is to be achieved.

They discussed the "S-shaped," a routing strategy, with 26 shaft changes and thirty shaft lengths being optimised (snaking up and down the shaft). The objective was, as previously stated, to reduce the total pick / route time. In the x-direction distance (horizontal) 19 percent and in y distance (verso) 12 percent decreased (32 changes on the aisles, 96 wingspans holding one SKU to 26, 84 respectively).

It was also stated that the OOS problem has been solved by two sub-problems in Mantel, Schuur and Heragu [7]: where to store SKU and a track to optimize road distance. Their approach is a "S-shaped" routing strategy (snouting up and down). The results will enhance the overall routing strategy by OOS and therefore reduce costs compared to COI. Based on the results, OOS has shown that other heuristic methods can be implemented while university researchers use the traditional IOC.

The actual vertical position of a slotted product off the floor also reduces or increases the selection time. Many hand-held tasks cannot be avoided in the case of warehouse operations; products remain turned and moved even in certain almost fully automated picking areas. With regard to ergonomics, slots can help not only minimize selection times, they can also improve the total selection ergonomics. In a paper by Jones and Battitste the authors talked of a ' golden zone ' of typically 40 to 60 inches on ergonomics and slot [8]. The range for men is lower at the top and the range for women lower. The height of this area has improved selection times because employees who bend or reach the zone can tire and increase selection time and create other problems.

2.1.2 Metaheuristics

Genetic algorithms (GAs) are another method used to optimize slotting. In a Wu and paper. Al. [9] Authors use GAs to resolve an automated system storage and recycling problem. The weight

distribution of SKIs in the racks will be a smaller sub-problem each evening. They found that the total harvest period was reduced by 8 per cent from the current layout to the GA design. Another practitioner, Zu, discovered that GAs can also improve slots [10]. As with the example above, a bi-objective, mathematical analysis bubble-free model has been developed to balance two goals: reducing the center of gravity in the SKU and reducing total time for selection. They saw a decrease of 43,7 percent compared to the original situation. However, the more significant or desirable variable for each target could be weighed in this particular case, for example 0.3 or a total weight of 0.7. Instead of an optimum center of gravity the optimum slots would be "weighed." In general, GAs showed a clear value as regards slotting optimization.

Kim and Smith's slotting methodology [11] tackled a problem similar to that that resolved in this thesis. Many warehouses know the data and the situation using light-to-light techniques for a warehouse management system. Product picking is supported together with WMS pick-to-light systems. The worker is generally given a visual indication to help in the process speed and limit errors. Five main areas of concern were mentioned by the authors.

1. Slotting: allocation of SKUs across the facility to different locations

2. Assignment of pickers: assign pickers to certain cells that are separated from the areas called zones

3. Carbonization: SKUs are placed in cartons.

4. Scheduling of boxes: sequence of boxes

5. Routing: all order sequences to be selected.

It assumed 2, 4 and 5 were out of their control and concentrated on the first operation if three operations were completed. There are metaheuristic methodologies developed and an optimum MIP model of mathematics is designed to be a benchmark (Most decent quarter slotting, suitable slotting and simulated ring slotting). However, the focus of their approach appeared to be the

recreating metaheuristic simulation. The implementation time of each method discussed raised one of the interesting points; some methods produce better results faster than others.

A real-life problem was the focus of this study and significant improvements were found from 5.5 to 27 percent, based on factors such as the number of cartoons, numbers and meta-heuristic results. The main thing is that they saw a total improvement in the picking time regardless of what type of technology they used. The total time of completion of all operations shall be Makespan. There are obvious warehouse improvements that can be made with many slotting techniques.

Yingde and Smith developed a simulated slot technology called ant-colony-slot-exchange optimization (ACO-SE) [2] similarly to simulate annealing techniques. A heuristic approach to reaching an initial viable solution reduces the simulation time of the problem. This enables solutions to optimize the picking wave process, better than random assignment by generating an initial viable solution. The authors only simulate time-and time-start zone for selecting SKU. The author has compared his ACO-SE-methods with previously studied COIs and simulated glutting techniques using the same information as in the previous study, besides examining selectors and returning products [11].

They found important results following their simulations (with Visual Basic Programming). In comparison to the COI method (regardless of problem-size in small to medium dimensions) the Authors have found a maximum reduction of 25.71% over the time of picking with an average of 18.28 to 18.78%. There were two significant results from the comparison between SA and ACO-SE: first, an improvement of 1.89 percent-20.23 percent; and secondly, the SA methodology results showed that there was no significant proximity to the SKUs but only in the same areas of seizure.

For the work of this thesis, the final idea of carbonization is relevant. In order to optimize the one another, the authors stress the need to combine carbonization and slotting. The study demonstrates that universities and the realm use similar systems with invaluable slotting technology (Pick-to-Pick systems and WMS).

2.1.3 Slotting on Affinity

A more recent Kofler et method. al also takes into account product affinities, and similar to the aim of the Thesis [12]. The "M-SLAP" procedure considers the variation of multiple years with the location problem of the storage location. For all reasons, whether seasonal or market variations should be considered, the method of this approach takes into account fluctuating demand. Not only do they consider the cost savings for each method of selection, but also the selection which should be made or restored in their analysis. M-SLAP appears to be some kind of technique for risk mitigation.

Even a few SKUs can produce significant results. Authors talk about moving. However, for optimum results, you could possibly move thousands of products, but this is far less probable in practice. For instance, they had a storage room with 5,832 pallets, in which the aisle changes could be reduced by 97% depending on the method, as their storage was empty. The relocation of the warehouse or its reloading as the writers call it can achieve this improvement. It was not an economic choice based on how much time it took. Another technique suggested is to change the operation over time and to progressively restore the products to new places. The previously discussed concept of cure was the last method to be tested. This test has shown improvements in the different areas and general changes in the aisle. In this study, the M-SLAP method, which uses healing / setting methods, does not work in real terms, when taking into consideration the more academic re-warehousing assumption. This makes the changes more realistic, which can enhance overall work. In this case it is obvious how the performance of your warehouse can really affect slotting.

2.2 SOM (SELF-ORGANIZING MAPS)

For the grouping into physical and immaterial object topographies, many methods were developed. In this literature review section [13], numerous applications for clustering are available for all areas such as robotics, pattern recognition, grading, etc. Organized maps or SOMs come from Kohonen

and are referred to as Kohonen's self-organized maps [14]. These maps differ from common methods of clustering, like k-means and classic statistical clusters. A self-ordering map is a neural network type that divides the data among various clusters using the random weight of the vector.

There is a potential for 1D, 2D, and 3D data clusters in order to extract knowledge very efficiently from a particular dataset. But 3-D SOMs are seldom used as discrete solutions, like 1, do not have space for a 2-dimensional solution. This technology has shown that SOMs can meet certain traditional methods of clustering, including ANOVA and k-means. In this section, a number of cases in comparison to other clustering benchmarks techniques and the performance of SOMs have been investigated.

2.2.1 Clustering of Physical Distance

The author [13] in this case is to assign a certain field of demand to supply centers in Spain / Portugal. There are several key assumptions about problem modeling in this article. The first is to fix all applications and the second is to guarantee the demand. The fact is that slots provide data and demand forecasts is different from the number of warehouse operations. The idea in the paper of Lozano and. Al. was a p-median / TSP problem. There has been considerable localization research [13].

The new section of the methodology for classical localization is the use of identified clusters from an SOM. A unique approach to medialization is the concept investigated by the authors. The cost is proportional to the distance from the Euclidean. It takes more time and more money for visits to a number of zones. The various solution, which brought similar results, which were also nearly optimal, was another important finding found among writers in [13]. Finding the optimal solution could include further classifications or limitations.

Like the location on a cluster basis, with a demand center to allocate supply center to the problem, authors Hsieh and Tien work on solving a problem using Kohonen SOL and the Math Model with

a view to minimizing the entire weighted distance. However, when there is a high number of supply and demand centres, the models are much harder to provide solutions in their research. They concentrate on major localization issues. Matlab uses parameters like learning rates, round and periods that affect SOM results directly in the preparation of experiments. This study was conducted in 2004 before computer abilities were dramatically increased for 12-13 years. It is worth mentioning. They talk about runtime and performance on the basis of certain parameters for long. Even people with optimal solutions have shown better results compared to other heuristics on the self-organized feature map, after comparing the already successful research on their field.

SOMs are applicable in many areas, including the supply chain, as can be seen in this paper and the findings.

2.2.2 Data Clustering

The map is an efficient tool for data groups and for other purposes it handles the physical distance. The data analysers and decision makers could use 2-D and 3-D self-ordaining maps more efficiently and easily in an article written by Rushmeier and Almasi [16]. Data segmentation can help to filter in real time large amounts of data. The pragmatism of this article is an academic view of the optimal solution and not what is in practice if any technical application is implemented. The authors stress that developing SOMs and successful marketing campaigns are important to political leaders in presenting these data, and how that promotes the successful development of a marketing campaign [16].

This thesis looks at the best way to slot. This practice must be adopted with realistic hypothesis in addition to purchases from warehouse managers and major players. A key factor that is discussed in the paper is control of the input selection role of data analysts [16]. In this thesis, the SOM main input are items in the sequence of an order or SKUs containing information on order. There is an important mention of the noise of the role of certain models, which the authors of this paper emphasized. The use of 2D and 3D models in data submissions was also an important point found

in research [16]. The report notes that 2-D maps are worth a solution so that feasible solutions for small data sets, with 100 records, can be limited to the best solutions.

Data clouds subsequently overlap the presentation, introducing the 3-D model. Since slots can have a 2-dimensional application in a warehouse, 3-dimensional modeling is the best opportunity for a successful slot redesign. It is very difficult to see more than three attribute dimensions. The authors have demonstrated that creation of a successful model does not just fit into a dataset, but it also provides visual information, making changes based on these models easy for those.

Although SOMs can provide a good understanding of each set of data, the methodology employed can differ slightly to achieve the desired results. Vesanthropy and Alhoniemi and Orkun et al. [17, 18] are the two-tier clustering processes that originally produce an inherently raw cluster and then replace it. The authors of [17] discuss the importance of the dual approach in their case. The grouping of data, hierarchical and "partial" approaches can be achieved in several ways. A partial separation in the areas in which SKUs are classified according to order and other inputs to the model would be included in this thesis the approach would be.

Determine number of clusters, start weight clusters (auto-SOM), divide data and updates, if not upgrade, stop; otherwise, return to partitioning stage. If weights do not upgrade, stop. With this algorithm, they advise prototypes to cluster instead of data, as even with few samples it can become a long template. One of the data collections tested was the face of a clown (visual pattern recognition); all correct attributes, like the nose, right eye and mouth, were successfully clustered by SOM. They talk about an unnecessary additional cluster, however, that probably is because of that particular feature's small number of data points. This is important not just for SOMs in all neural networks. In order to achieve a certain efficiency, the number of data points must be high. The more clusters, the more sensitive they are to surface areas, is a positive result that the authors mention, but this thesis does not indicate a very high number of zones. In the identification and categorization of data correctly, the grouping methods were successful [17].

It is efficient to use an approach of two stages, including those used by SOMs to collect and organize multi-disciplinary data, in accordance with the techniques of Vesanto, Alhoniemi and Orkun [17]. In this case, the types of fragmentation mechanisms will be analyzed by a volcanic ash test. Authors recognize that nonlinear or linear issues can be solved by SOMs. Classification problems solving problems include these types of problems. Two different kinds of SOMs were involved. The first "high" training and a finer formation, both consisting of an input model with 13 functions. After 500 studies they found a global 0.004572 measurement error on a map (the methodology chapter will discuss this type of error). These significant results were found not only in error but also better than standard statistical models like ANOVA. SOM could identify the same important surface texture types as ANOVA found in the example in particular. Another way to enhance or compare efficiency of SOMs is to compare additional methods, including k-means. Although a number of other applications for a SOM and several others show that SOMs are useful for practically more concrete applications, society, not just a special organization, can influence.

Mostafa offers work that contributes to identifying the environmental factors most significantly. SOMs [19] are studied. Biocapacity as a benchmark can be evaluated for sustainability in this study for each respective country. This was the first such study to explore this particular problem using SOMs, taking into account attributes such as literature, GDP and export data for trends in biocapacity. Three distinct clusters were established. The "temperature maps" are displayed visually, and combined, allowing for easy visualization of input factors after differentiating the cluster. The various ways in which data can be visualized are one of the advantages of SOMs and the idea is constantly demonstrated in [19].

The author performed k-means scoring and ANOVA statistical analysis before the results were discussed. It was found that better results than the k-means were achieved in the SOM. One of the author's important comments is on how the combination between GA and flawless modelling improves the SOM and how the proof is made of this concept. The author's recent applications for positive SOM clusters finds that the use of SOMs in this thesis is relevant and supportive for the

purpose of clusters. In this article, Mostafa uses SOMs to cluster organ donors ' motivations again in a paper [20] when they have been clustered with SOMs.

In [20] the author tries to determine why the Egyptians donate organs and what the decision could affect. In order to save lives, it enables marketing strategies and tactical profiling. Inputs like altruism and organ donation knowledge [20]. The data is split into four SOM groups. The author's software did not provide the number of clusters, as the number of clusters required is predetermined for certain SOM packages.

The introduction of high-dimensional data is one of the key features SOMs offer uncontrolled clustering algorithms. The k-means and SOM clusters produced the same results in this particular case. The author found that the key components of the donation of marketing organs are perceivable benefits, positive attitudes and intentions. Even though the results from the SOM were less effective, the study was still excellent as the conventional K-means are just as efficient.

As mentioned earlier, social efficiency SOMs have many areas. Social issues that are valid in other aspects of life are, in other cases, grouped into SOMs. Music companies can be classified according to the right genres to reach the correct public and optimize possible hearers for a genre of musicians. The genres of the music are classified as SOM [21] in Ahmad, Sekhar and Yashkar. The three classes were quite successful when they divided the genres by pitch, luminosity, tempo and enharmonically keys in three separate categories. Based on the cluster analysis, they achieved a precision rate of 81.8 percent compared to their expected general population. The use of SOMs and the methodology for clusters creates a simple and effective way to classify the music mathematically.

The more social applications for clustering were examined by Stambuk, Stambuk, and Konjevoda, like Ahmad, Sekhar and Yashkar [21]. A wider approach has been identified with the SOM to human religious motivation. The authors determined three clusters of religious motivation by employing ten information from 473 people of all ages and genders. The collected data was based on the intrinsic religious motivation of Hoge. The variance between 2 and 3 clusters was analyzed.

The three-cluster method was 8.1 percent acceptable. The writers show that the SOM not only performs well but that the graphic nature of the data representation makes it easily interpreted. Their data were tested against the main component analysis and hierarchy grouping techniques. The results found in countries such as the US (the survey was carried out in Croatia) are comparable, so this model can be used for the majority of places worldwide.

As the relevant and efficient slotting and clustering by SOMs are demonstrated in these papers, the current slotting methods of the warehouse should be demonstrated. SOM clusters have also proven that, if we compare conventional and older statistical methods, comparable and often even improved results are obtained. A comprehensive description of how a well-structured warehouse is being built, the relevant inputs, the number of clusters and other factors.

3. METHODOLOGY

The advantage of face recognition is a passive, non-intrusive identity control system. Many techniques for face recognition supervised and unattended were reported. Several algorithms have been used for facial recognition and can be divided into two methods: structural (appearance-based) and statistical (feature-based). In recognition of the face three technologies— PCA, ICA and SOM [1], [2]. Karhunen-Loeve's Transformation [8] is derived from the Primary Component Analysis (PCA). Because every facet of a photo set is represented in its own dimensions, PCA tends to find the maximum variation direction in the original photo space with its basic vectors. Normally, the new subspace is smaller. If image elements are seen as random parameters, vectors of the PCA basis are defined as vectors of the dispersion matrix. Independently analyzing components(ICA)[8] minimizes input data dependency in second order as well as higher order, and seeks to find the basis for (where projected) statistical independence.

Bartlett et al supplied the ICA two ICA Architectures–statistically independent basic pictures and Architecture II factorial code. Facial SOM analysis results better than those shown in this paper, PCA and ICA techniques. SOM is an uncontrolled learning process that maintains topology.

Basically there are two types of modes [5]:

- **Supervised**: The desired system response is provided by the teacher during supervised learning at every point when the input is used. In many situations of natural learning, this persistent mode is used. This learning mode requires a set of inputs and putting patterns called a training set.

- **Unsupervised**: Uncontrolled study algorithms use patterns that are typically redundant raw data without labels of membership. The network needs to identify any possible patterns, regulations, separating properties, etc. in this way of learning. When this is discovered, the network undergoes parameter modifications, which are called autonomy. Sometimes unattended learning is referred to as teacher-free learning. We use the non-controlled learning algorithm based on the neural network called the self-organizing map.

3.1 FACIAL ANALYTICS

Persons are often recognized through their faces and a similar recognition has now automatically emerged with recent computer technology. Algorithms initially used geometrical simple models but processes of recognition matured into an enhanced science of mathematical representation and communication.

Algorithms for recognition can be split into two major approaches,

- Geometric that looks at characteristics,
- The Photometric approach distils an image into values and compares the values with templates in order to eliminate differences. For our facial analysis, we use photometric approach.

Our project's basic block diagram is illustrated in Figure 3.1.

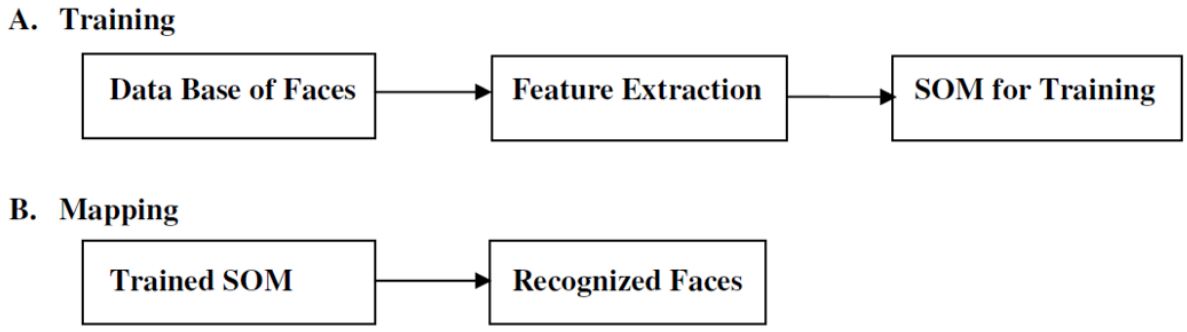

Figure 3.1: Face Recognition

1. Training

   Data Base of Faces: Each of 40 distinct persons contains ten different images. For certain people, the images were taken at various times, with different lights, face expressions (open / closed eyes, smiling / not smiling) and face details (glasses / no glasses). All pictures are takenS in an upright frontal position against a dark homogeneous background (with tolerance for certain side motion). The PGM files are available. Feature Extraction: Feature

removal is a special form of reduction of dimensionality. If the data input into an algorithm is notorely redundant and is too large to process (many, but not many data), it is converted into a reduced range of characteristics (also known as vector characteristics). The data input is transformed into a set of functions. It is called functional extraction. The set of functions is to extract the relevant data from the input data in place of full size input to accomplish the desired task with this decreased image. SOM Training: Input and target vectors are included in the recognition process. For SOM training, no target vector is required. A SOM can categorize training data without external monitoring leading to various Clusters or Classes being formed. Various clusters such as grid, hex and random are topological. In our project, we employ hex topology, which covers the maximum area of neurons trained.

2. Mapping

   Trained SOM: The input image mapping is carried out using trained database clusters. The application of the Euclidean Distance formula is the result of this match. Recognized Faces: The result of this process is the best match determined with the Euclidean distance formula. The actual recognized facial image is the minimum distance between an image input and the classifier or cluster.

3.2 A LOOK AT NEURAL NETWORKS

Neural Networks SOMs are a special class. We must first clarify the actual neural networks, their workings and the economic foundation, before we turn to a detailed description of the algorithm.

3.2.1 Linear Regression

One often has to predict or predict certain target variables in economics. Theory of forecasting has become an important field of research and is equally important in every field of economics specialization. Projections are usually carried out by linking one output variable and a series of variables observed x. The x set can also include lagging observations of the relevant target variable. This leads to the linear regression model in the simplest case:

$$yt = \sum_{k=1}^{k} \beta_k x_{k.t} + \epsilon_t \tag{3.1}$$

$$\epsilon_t \sim N(0, \sigma^2) \tag{3.2}$$

The random error ϵ_t is usually assumed to follow the normal distribution with mean zero and σ^2 variance (the variance is assumed to be continuous). Parameter sets of $\{\beta\}_{k=1}^{k}$ are the part that is estimated, and the target variable is denoted by y using the $\{\beta\}_{k=1}^{k}$. The linear regression model tries to detect the vector $\{\beta\}_{k=1}^{k}$, minimizing the differences of actual y to y predited values; in other words, the evaluation procedure is supposed to lead to a parameter set ϕ in parameters ϕ in order to minimize a function, when the y value is defined.

$$\min \phi = \sum_{t=1}^{T} \epsilon_t^2 = \sum_{t=1}^{T} (y_t - \hat{y}_t)^2 \tag{3.3}$$

Prevision is usually achieved by freezing the pre-estimated parameters and using them for extrapolating in the future. The autoregressive model of the form is the most common model used in literature for prevision.

$$Yt = \sum_{k=1}^{k} \beta_k y_t - k + \sum_{j=1}^{j} y_j x_{j,t} + \epsilon_t \tag{3.4}$$

The lags of the dependent variables y and independent variables x (weighted by the parameters) in this model k* are included. Naturally, this results in an estimated total of k*+ j* parameters. The degrees of freedom in this regard are reduced by adding additional lagged dependent variables to the model. The linear model has a closed-form solution that minimizes the amount of squared errors (or: differences) between y and ^y and is calculated very quickly. In contrast, when trying to capture nonlinearities in time series the linear model fights. We can, for example, observe long-term trends on Financial Markets, but in the short term, time series changes are often huge, e.g. due to bubbles that burst. The linear model meets its limits in this regard.

3.2.2 Nonlinear Models – GARCH

There are various nonlinear models that attempt to capture the true underlying process with an assumption of parameter and a particular function. The GARCH model, developed by Bollerslev

and Engle, is the most prominent example. This model directly affects the mean value of the target variable by the variance of the disturbance term. The variance is modeled on its own past values and the previous square prediction mistakes. The GARCH template can be written as simple as possible

$$\sigma_t^2 = \delta_0 + \delta_1 \sigma_{t-1}^2 + \delta_1 \epsilon_{t-1}^2 \tag{3.5}$$

Under the GARCH model y, the return rate for some assets is represented, the expected appreciation rate is present and t is assumed to be normally spread with a null mean and conditional (heteroscedastic) variance σ_t^2 .

The parameter describes the risk premium effect on the return; i.es. When the risk increases (β > 0), the investor demands a higher return on an asset. The vector parameter defines how the dependent variance evolves. A stochastic recurring model is the GARCH model. If σ_0^2 and ϵ_0^2 are given initial conditions and the α, β, α0, μ1 and β2 estimates are known, then the asset return is determined in full as a random shock, its own averages (by parameters), and a Risk Premium Effect. Random shocks are supposed to be distributed normally. The likelihood function for this problem is therefore very easy to set up.

$$L_t = \prod_{t=1}^{T} \sqrt{\frac{1}{2\pi\sigma_t^2}} \exp\left(-\frac{(y_t - \gamma_t)^2}{2\sigma_t^2}\right) \tag{3.6}$$

The hats indicate that the respective parameters are estimates. In general, the logarithm of the probability function is used by the Berndt-Hall-Hausman (BHH) algorithm for numerical reduction using some standard iterative procedure.

The probability function is often very difficult to minimize. The GARCH model is one of the biggest inconveniences. The GARCH model has the same problems as every other estimate of the probability parameter. In addition the GARCH approach is also very restrictive since this approach is confined to a well-defined distributor and a clear parameter set. The method of estimating is not always compatible with interpretable parameters. Nonlinearities are explicitly considered by the GARCH processes. This can be seen from the way the dependent variance is driven. It is driven

by the non-linear transformation of his own past values and a non-linear transformation by past forecasting errors. The variance is easy to use as a determinant of asset returns as risk plays an integral part in financial markets ' price pricing and, of course, in dynamic predictions.

3.2.3 Polynomial Models

We try to deal with the disadvantages of GARCH using polynomial approximation methods. The methods, including neural networks, can approximate the unknown non-linear process with less restrictive semi-linear assumptions. The functional forms here are indicated but the polynomial degree or the amount of neurons is not. It is not possible to give the numbers of parameters only as they are normally, e.g. for GARCH models, an easy interpretation of the estimate. This is why this type of model is known as a semi-parametric model.

$$g(x_{t-1}, x_{t-2}, \dots) = \sum_{i=1}^{\infty} a_i x_{t-i} + \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} b_{ij} x_{t-i} x_{t-j} + \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} \sum_{k=j}^{\infty} c_{ijk} x_{t-i} x_{t-j} x_{t-k} \dots. \quad (3.7)$$

This is the standard linear average in equations (3.7) (the single summing). The twin sum accounts for two variables lagging across products, etc. In order to avoid counting the sum indexed with j at I more than once, the sums indexed with k start at j etc. The function in Equation 3.7 provides a balanced sum of polynomial functions from the previous variables. With the degree of polynomial expansion, the number of parameters increases exponentially. This is known in the nonlinear approach as the curse of dimensionality. There is a decreasing degree of freedom in statistical estimates accompanying a increasing amount of parameters (and thus an improvement in the degree of accuracy). That's what a scientist has to deal with.

3.2.4 Neural Networks

Nonlinear statistical processes are a very efficient means of modeling the neural networks (NNs). As a rule, NN refers to a series of variables input to a set of a few variables in output. The method differs from the approaches described by the one or more layers used by an NN to break down variables in previous paragraphs. This change is done by a special function (so-called logistic

functions). Many different NN architectures are found in the literature. We will first give an understandable example of the design of the Feedforward Network. Our example of one hidden layer comprising of two neurons, one y output variable and three input variables x1, x2, and x3 is the simplest specification for a NN.

In this graph, the input variables, also referred to as Neuro input and Neuro output variable (NEuro outcomes) are combined with the hidden layer by means of synapses. The hidden weight of the input variables is attached and an output is generated. This operation is processed in parallel by the hidden layer to improve prediction (parallel processing). This is a copy of the human brain structure that involves a huge number of neurons in parallel treatment. The number of synapses is increasing significantly as the human brain develops. The brain processes the data in parallel when a sign (or: input neuron) is presented to the NN. Fortunately, it is enough to deal with a very small number of neurons for economic models and forecast purposes. Latent variables in the models are often important in econometrics. Usually the driving force in an economic process is this class of variables.

A range of subjective factors, including personal experience, education, intelligence, and culture, influence financial markets investors, although they are also influenced by laggards and current critical variables. All these factors are interconnected, and relationships with standard modeling strategies are usually very difficult to understand. NN is a natural tool for predicting investors ' decisions because it tries to clone signal processing in the human brain. Due to the complexity and the latent variables mentioned above, NN seems to have good performance in this field in this decision-making process. The NN-Learning Process is based on two principles: Rustichini et al. (2002).

- Principle of functional segregation: This principle basically expresses the idea that the entire brain is not required to take certain decisions or to process all brain functions.
- Functional Inclusion Principle: The principle states that various regional (brain) networks are used for various functions. However, the regions used in various networks overlap.

The authors argue in the same paper that people take decisions on the basis of approximations. Probably the most natural way to address forecasted problems is if this argument holds true.

a. Sigmoid Activation

The question is how such an NN can be formalized and how exactly the above-mentioned punch function can be formulated. When presented to the network, the information is processed in two different ways. The first is to form linear data combinations, and the second is to quash these linear transformations by a signmoid or logistic function. In many applications this function appears in relation to a triggering function in which the neuron is not triggered by small changes in the data if the system has extreme values. However, small changes can change when \threshold areas "are present and can therefore influence the system. The following way can be written a system with a threshold (and the most basic NN):

$$y = \frac{1}{1+e^{-x}} \tag{3.9}$$

This special form is also known as the MLP network. This is a feed-forward neural network. It is of special interest in economics because thresholds on financial markets are often decided. A trader may not affect his decision to sell or hold an asset if the company is still operating on a very high profit level, if the company announces less pro-t in one period.

This scenario becomes another story, when the company is already struggling to make money. In such a situation, the announcement would simply be much more important because it had become. This information is definitely being considered by the trader and is acting accordingly. Thus this very simple NN is a very good alternative to conventional prediction models of linear time series. There are a range of changes and complicated forms in the literature, but the simple MLPs that we have are enough to understand the basic ideas behind the NN approach.

b. Other Activation Functions

The revision of the activation function is a natural adjustment of the algorithm. In contrast to the above defined Logsigmoid function (logsigmoid input combinations are broken down in [0; 1]). "Linear combines of input within interval" [-1:1] are the Tansig or Tanh function (also known as the hyperbolic tangent function).

The Gaussian function is narrower than the logsigmoid function, and shows little or no signals when extreme values are included. (i.e. 5.-2.(;-1).). 2. In the interval between critical values like [-2; 2] the Gaussian active function is steeper than the logsigmoid function. Many other activation functions have been used in the literature and a rule of thumb, the one best to use, cannot be given. This issue is the subject of research for the researcher to decide. The network design and the selection of the activation function must fit the topic. The three functions presented here present a nice overview of the neural network's basic structure.

Different features can be identified by the Radial Base Functions (RBF) from the above described MLP. The two main variations are

- Maximum 1 hidden layer is available.
- The radio transmission system calculates the Euclidean distance between the input vector signal and the unit center.

The input neuron may be a linear regressor combination, but an input signal or a set of input coefficients are present. The input signals are identical to each neuron, so that each neuron is transformed in Gaussian by different means around k*. The input signals therefore have different centers for normal distributions. If you take and combine these various Gaussian transformations in linear way, the output or forecast results respectively.

The input variables are x and n, for which weight of $\{w\}_{i=1}^{i*}$was taken into account. The above are the linear transformation. k* different centers $\{\mu\}_{k=1}^{k*}$ are available. The Rk functions are obtained from k* with this information. Once the functions of the radial basis are identified, they are

combined to predict them. The R_k is then weighed by a parameter vector called $\{\tau\}_{k=1}^{k*}$. To that end the R_k is used. The optimization procedure embraces the adjustment of the parameter sets $\{w\}_{i=1}^{i*}$ and $\{\tau\}_{k=1}^{k*}$. Moreover, the centers of the transformation $\{\mu\}_{k=1}^{k*}$ are an argument for fitting the net. Between RBF and SOM there are a great many analogies. Both methods have their basis in measuring the distance from the center of a neural unit between certain data in the vector area. Blayo et al. (2003) recently attempted to combine these two methods to predict the DAX30 Index. Based on the data sets previously identified by SOM, RBFs were used for the development of local models.

3.3 PATTREN RECOGNITION

One of the few high-precision and low intrusive biometric techniques. One of the key steps in the image processing is the identification of patterns. The first step in recognition of patterns is to select a number of features or attributes to sort the pattern from a universe of features available.

The original pattern then has to be converted into a programmatic display. Functions in the data that are defined as corresponding to the pattern are searched after processing data to remove noise. The data are classified by similarity measurements with other patterns in the classification stage. The pattern recognition process ends when the data are labeled on the basis of their class membership.

In the face recognition system, we have a image database stored in the system. When a new image is obtained, the image database in the system is compared to the already saved one. We first tested the database of our students.

- Training uses input examples to build the map. It is also known as vector quantization as a competitive process.
- Mapping classifies a new vector automatically.

3.4 SELF-ORGENIZED MAP

The self-organizing maps are in their simplest interpretation a' elastic network' of points suitable for a certain distribution (Kohonen, 1997). This network has the interesting property of maintaining global ordering of the original distribution space by network nodes (index vectors), i.e. no' knotting' of the network. A knotted network would be difficult because traveling in one direction over the network would not ensure that the same point is not revisited in the design area. The preservation of the topology of this network leads to the map forming a non-parametric model for the design environment.

The nodes may be considered' Classifiers' of their regions in the design area once the network is ordered (or trained). When a design is given to the network, the node is activated that represents the domain area which is closest to that design. Due to the topological order of the network, it is possible to move to neighboring nodes within the design space to nearby regions. The network can be seen as local features, i.e. the network can be used to approximate the parameter relation for small regions with nearby traffic jams. Similarly, the SOM characteristics should be identified and interpreted by a domain expert, to identify and interpret the linear features found by PCA.

3.4.1 Introduction

Differentiating between data analysis and data mining has often been confused. The two terms are not precisely defined, though the expression exploratory analysis generally describes the whole process of data extraction while the expression "data mining," meaning the whole subgroup, is used to describe the phase of discovery. All of this can subsume the generic term discovery of knowledge in data. The aim is to detect information from the data structure that can be used to understand the underlying complexity of data-driven processes. It can be done either by visual or numerical inspection (or by a combination of the two). Lastly, a \classical ' statistical analysis and data processing methods such as the neural networks or a SOM are the key to this analysis.

In contrast a specific technique to make this analysis feasible is required for the visual inspection of the data structure. The methodological requirements are more sophisticated when multi-

dimensional data come into play. Dimensional decrease methods were widely researched in the field of data analytics, resulting in the principal component analyzes (CPA), cluster analyses (CA) and factor analysis (FA). Extensive data analysis research was performed on data reduction techniques. This thesis refers to CA algorithms that are grouped into similar patterns by the methodology of the Map of Self Organization. In turn, the (multivariate) data can be used for symbolic identification with the aim of reducing its dimensionality. For four different reasons the SOM approach is an attractive tool for data mining. Namely, the technique is

- A numerical method.
- A not parametric method.
- A method which does not require preliminary data distribution assumptions.
- A method that can detect unexpected data features due to the unattended learning nature.

All this provides the user with a great deal of flexibility. There is a virtue in the above-mentioned benefits of the technique: the user has no ex ante assumption of data structure, restrictions on parameters or any other odds or odds. Data mining from time series has mainly addressed two issues in recent years. Initially, to extract informative structures from time sere patterns, either in whole or in part similar, and, secondly, to identify full or in part periodic patterns in the time series, which may be used in conjunction with the initial one. The self-organizing maps of Kohonen are a very sophisticated way of recognize patterns and explore data. The remaining chapter describes in great detail the algorithm. Then we shall look at the possible use of SOMs for economic issues.

3.4.2 Mathematical Overview of SOM

Teuvo Kohonen, a graduated computer scientist who mainly works in artificial intelligence, developed in 1982 the fundamental algorithm for Self-Organizing Maps. It is one{ and probably the most popular{ approach to a technology that is commonly used in natural sciences as a way of classifying, organizing and visualizing large and multidimensional datasets called unattended neural networks. Financial market researchers are especially confronted with data which make the

filtering of informative structures difficult. Moreover, accessibility and data availability have increased considerably over the past two decades. For economists, it is a valuable development, but more advanced means of managing this growing quantity of information are needed. Kohonen SOMs describe a certain kind of neural artificial networks trained according to certain rules of education. The ANN adapts successively, on the basis of these rules, to the structure of the formation data set. Artificial intelligence methods are increasingly entering economic research, in particular in financial market quantitative analysis. The latter field of research has seen neural networks and genetic algorithms. These techniques are capable of simulating the behavior of financial market agents because a learning aspect is explicitly integrated into the mathematical algorithms. Self-organizing maps by Teuvo Kohonen belong to the class of methodology that has drawn much attention in recent years, both academic and non-academic. SOM is a special neural network class with fields of equipment ranging from mathematics and science to 3-D virtual generation and medicine. The algorithm concentrates on economists because it is a powerful device for analyzing time series and data mining. For both marketing and financial market modeling, the SOM methodology is particularly useful.

Deboeck and Kohonen (1998) published a first series of papers on the utilization of SOM models in finance. SOM models are particularly interesting to the financial industry as they provide a convenient solution for the processing of complex, high-dimensional time series data. Kohonen Maps, whose roots are in science, have not yet been deeply involved in economic research but have already introduced some interesting, innovative approaches. There is still very little economic literature in this area. Blayo et al. (2003) made influential contributions and so sought to find a probalistic implication for the prediction by SOM of a time series. In addition, Cottrell et al. (1998), provide a solution for a period-projected problem by training a SOM for the day-to-day use of energy for a given population and by using the identified sample prediction patterns (within 24 hours). A concrete example is provided in Resta (2002) for applying the method on financial markets. Principe et al. (1998) use Self-Organizing Maps to predict chaotically decisive time series, as the Lorenz equations generate. By using the pattern recognition method, you gain considerable success.

It is aimed at creating a compact representation of the X to RN space on the basis of the XT sample distribution of previous examples of design. This represents the mapping of a more compact (i.e. smaller) space from the design space. This mapping must no longer be linear, but the design space is assumed to be bound up and interpolated in the design space by a finite set of "codebook" vectors. These vectors represent prototypes of design and are arranged in a m-dimensional grid (m= N) in such a way that the neighboring codebook vectors are also' neighbouring' inside the design space with respect to the grid. The purpose of these codebook vectors is to' cover' design space in order to provide a proportional increase in the number of vectors covering a densely populated distribution area than in a densely populated distribution. The vectors of the codebook are determined by the sample distribution +/-XT. The closest codebook vector for each of these, m, has been modified in order to be closer to the given training vector, x, for this iteration:

$$m = m + \alpha(x - m) \tag{3.1}$$

The' learning rate' is where m is updates the index vector and α is a parameter which determines how much the index vectors can be changed. This is repeated and the rate of learning decreases α until an error has reached an acceptable level between index vectors and the sample distribution. It should be noted that this iterative process also provides noise resistance for the distribution of the sample.

3.4.3 Clustering Illustration

Two parameters (inner and outer bay mixer port diameter) of the combustion mechanism of the self-organization maps are used to illustrate. A small amount of noise is added to the diameters so that the original design settings can be changed a little. The illustration is further developed by adding dimensions to the training data successively. Initially, the design space determined by two mixer ports will be trained with the small (4 to 3 elements) SOM (Figure 3.2).

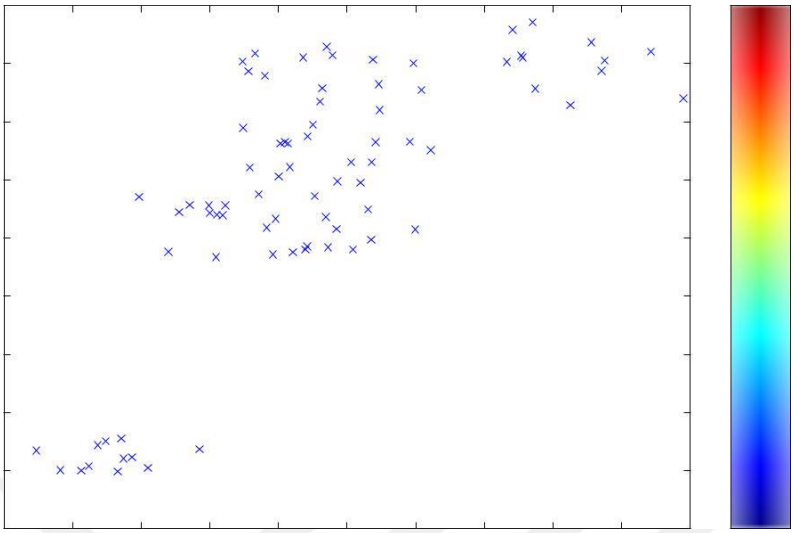

Figure 3.2: Mixer port diameters with noise added

a. Initial set up and training

Initialization of the SOM codebook vectors, mi, is pertinent to the 2D design space as a regular lattice (Figure 3.3). Then, every data point in the training set visits the SOM. Letx be the first point (Figure 3.3 shows the data point). In this case, point m6 is also identified as the closest (closest) lattice point. This grid is then moved closer to the data point (which in the grid is' more similar' to the data point):

$$m_6 = m_6 + \alpha(x - m_6) \tag{3.2}$$

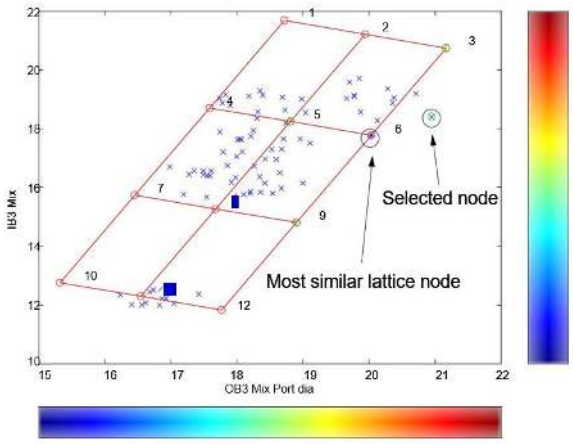

Figure 3.3: Initial Self Organizing Map lattice overlaid on design space

where α is slightly less than 1 set initially, and decreases over the entire set of training points with each iteration. The m7's neighbors are also moved similarly, but to a lesser extent, to ensure the lattice remains topologically correct. Only the immediate vicinity of m7 is updated in this situation:

$$m_3 = m_3 + \alpha(x - m_3) \tag{3.3}$$

$$m_5 = m_5 + \alpha(x - m_5) \tag{3.4}$$

$$m_9 = m_9 + \alpha(x - m_9) \tag{3.5}$$

α2 < α where. Figure 3.4 shows the result of this operation. This procedure is repeated several times on each point in the dataset (each time the whole training set is iterated, with α and α2 decreases each time) until the grating form converges into a display of the training data (Figure 3.5).

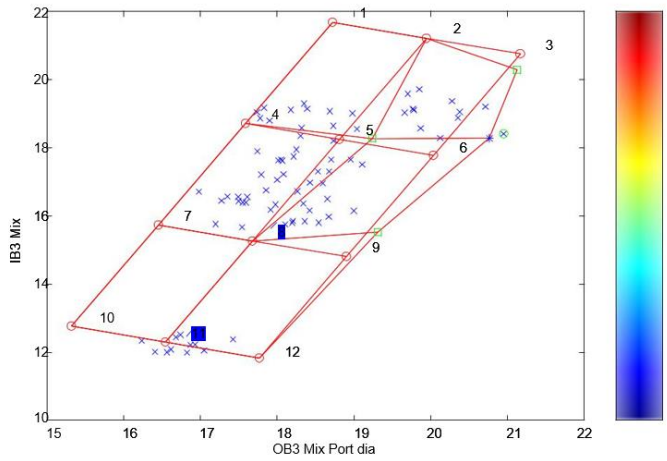

Figure 3.4: Movement of lattice points after the first iteration

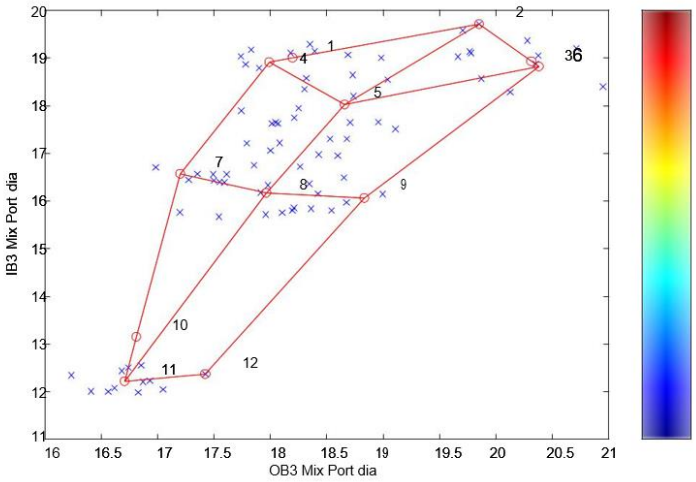

Figure 3.5: Position of lattice points after training

b. Generating the component maps

The maps of component projections can be visualized by the SOM. The projection of a particular component of codebook vectors, drawn on the corresponding point in the gate, is each of these maps. The projecting of that codebook component, mi, will be placed on a grid in each component

I in the design space (i.e. both in design parameters and assessments) of the SOM (in Figure 3.6 the color representing values, scales for which approximate bar scales are drawn, also alongside the component axes, in the 2D scatter plot). (see Figure 3.6)

The SOM maps are used for identifying clusters and producing heuristics. This illustration illustrates two design clusters, namely designs with large diameter mixer ports and those with low diameter mixer ports, which are present in the current 2D training set. Correlations are identified as follows from these component maps:

1. Compounds with very similar maps have global correlations (similarity is identified through the presence of some form of monotonous relationship between maps).
2. Local correlations are established between components on their component maps which share a similarly formed area. These links only apply to a certain subset of the entire design space.

The non-linearity of SOM does not necessarily result in a linear correlation between the correlations discovered by this principle. The ability to find local correlations is also an interesting property that is used to detect phenomena that do not exist across the entire design space.

In today's 2D illustration the inside and outside mixer port diameters are drawn from the 2 component maps (see Figure 3.7). These maps of components are examined in the original (training) data set for the clusters. A large distance between items 7, 8 and 9, with elements 10, 11 and 12, is available to check the right component map. On the left-hand component map, a similar but smaller effect is noticed. In this case, the first component should mainly be used to identify which cluster is a design to which the value of that component along with models from different classes is expected to be more discriminatory. In this case, therefore, there are two clusters that are based on where the SOM is extended by component two. As in this case the two components are the diameters of the mixer port, the two clusters are described as large-diameter and small-diameter designs. This suggests that designers only have to decide whether large or small ports are used instead of deciding independently on the size of both ports.

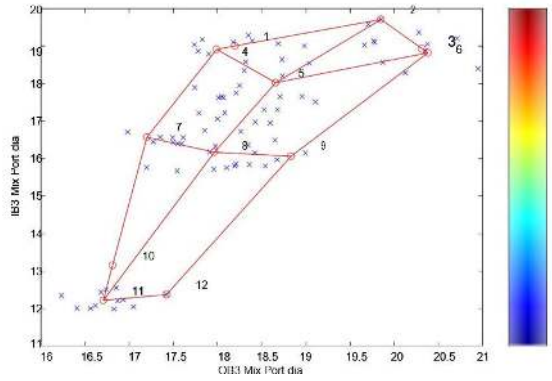

Figure 3.6: Trained SOM lattice overlaid on design space

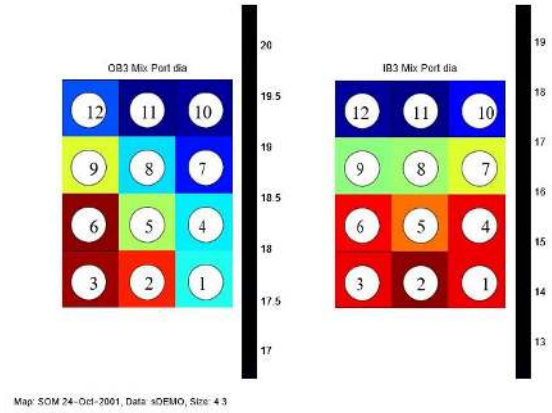

Figure 3.7: SOM component maps

3.4.4 Expanding to Higher Dimensions

In a case of reality, the design would be many more dimensions and therefore the SOM would have to design these larger dimensional spaces. The traditional plotting techniques no longer allow these spaces to be visualized. The illustration will initially be developed into a 3D illustration with a graphic illustration. After this, 4 dimensions and by extension, arbitrary dimensional spaces will be advocated in this illustrative approach.

a. Adding a third component

Each member of the training data set receives a third component, T30 (air temperature at the combustion entry) and a new network is trained (Figure 3.8). This new trained map does not necessarily resemble the previous map, since the second component will affect each node's two initial components. The first difference between the new map series (Figure 3.9) and the previous set is that two mappings in port diameter have rotated by 90 livres in gradient orientation. This is because of the initialization of the maps and the effect on this procedure of the T30 component. An initial examination of these three maps seems to indicate, due to the orthogonality of the gradient between the port diameter maps and the T30. This map set would either show that T30 is unrelated to port diameters or that the space description misses an important component, because it is assumed that there exists a relationship between it (direct and indirect). If the T30 part is relevant, the parameterization of the space is supplemented with the fourth component.

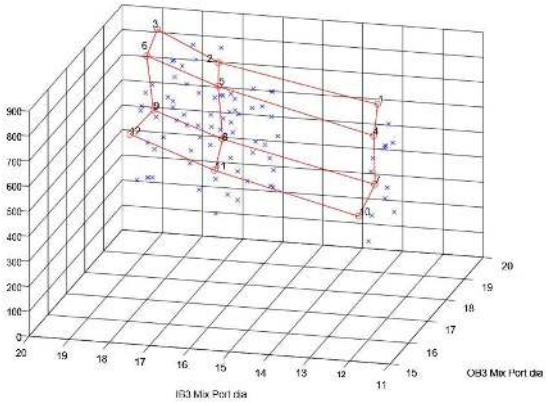

Figure 3.8: 3D plot

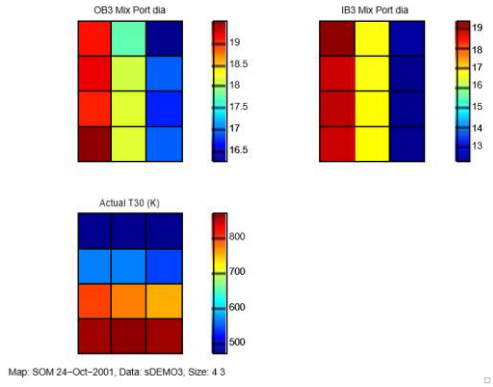

Figure 3.9: Component map projections

b. Adding a fourth component

A fourth component is now added, P30 (air pressure at incinerator input). These data can not be traced directly to a two-dimensional surface with four dimensions. Consequently, the component maps generated by the SOM are now needed in order to extract data relationships. The four

dimensional data set component maps are provided in Figure 3.10. Figure 3.10. The T30 and P30 are obviously similar to the maps of these components. The addition of P30 also led to a shift in port component diameter maps. This change identifies the potential connection between harbor diameters and pressure and therefore also between harbor diameters and temperature. A clear area of low diameter is shared between the two maps on the component maps. When this region is mapped on the other maps, it can be shown that a combustion system operates at a high temperature and pressure with low diameter mixer ports. In reversing, it is possible to see that the smaller values of the temperature and component maps are dominated by designs that have a larger port in diameter. Therefore a heuristics of design related to these elements could be:

As the diameter of the mixer ports is decreased, the combustor operates at higher temperature and pressure.

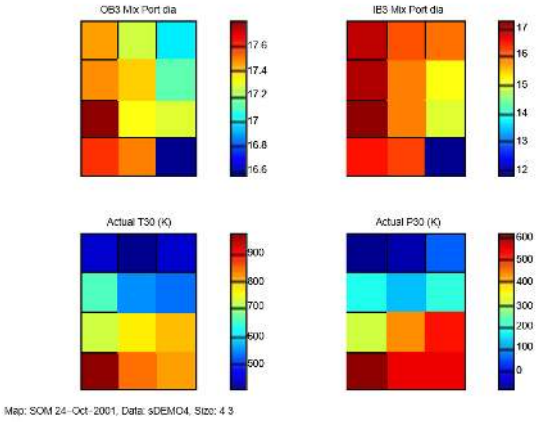

Figure 3.10: Component maps for the SOM representing a 4 dimensional space

Such heuristics are then given to a domain expert to check for precision and perhaps also to re-express heuristics so that it is more useful to a designer. The cause-specific guidance of the heuristic cannot be determined and therefore the domain expert must also clarify this.

c. Higher dimensions

The procedure applied in the preceding paragraph is extended for higher dimensional spaces. Similar map groups from a potentially large collection are now difficult to recognize. In the first place, the maps are compared in pairs and an equivalence value for each map pair is associated. There is a problem with' inverse relationships,' which are classified as different in comparison algorithms. By comparing the absolute gradients of the maps, the opposite correlation can be recognized by two maps that show opposite gradients.

3.4.5 Interpretation of The Component Maps

The objective of the component maps is to extract meaningful links between components. From previous approaches to this problem it was noted that there is a relationship between components if the SOM component maps "behave" similarly. However, all component maps are taken into account at once and therefore global trends are identified within the data. A comparison between maps for similar behaviors, extending the approach adopted in this study. Local phenomena can also be identified using a sufficiently sensitive similarity measure (i.e. relationships that only have in one part of the design space instead of the entire space). A modified Tanimoto metric determines the similarity between component maps that compares the maps on a lattice point by lattice point basis. Upon collection of all these metrics, the component maps can be clustered according to similarity. The clustering of design components shows the domain structure, which in turn can be used as a basis for creating a coarse domain model. The examination of the clustering also generates explicit domain knowledge, which was implicit in the distribution and perhaps unknown to field experts. The ability to identify component pairs with a connection between them is a huge computer saved when all the component pairs are tested by regression. This is a new way to study previous design data in order to generate explicit rules that explain component relationships.

3.4.6 Final Thoughts

For non-linear fields, the use of Self Organizing Maps has been shown. The benefits of using SOM are its resistance to noise from training data and its ability to train from the distribution on a

relatively small sample set (compared to other neural networks, such as multi-layer sensors). The SOM however produces a limited map that cannot be extrapolated far beyond its borders. In addition, if the distribution of the sample is small, the resulting grid is probably rough. The self-organizing map or SOM of Teuvo Kohonen is an unattended learning process that learns how patterns without class information are distributed. The Kohonen Self-Organizing Maps (SOM) were used as a neural uncontrolled learning algorithm in a large variety of models. The use of SOM as a feature extraction method in face recognition applications is a promising approach since learning is not monitored, and pre-classified image data are not required. In highly compressed representations of face images or their parts, the final classification procedure can be fairly simple with a small number of labeled training samples. The SOM differs from most classification and/or classification techniques as it topically offers classes. Similarity in input patterns retains the process output. The topological maintenance of the SOM process makes it particularly useful to classify data, which consists of many classes. The algorithm used for our Facial Analytics project consists of following steps:

Step 1 – START.

Step 2 – Initialize the Map for Clustering.

Step 3 – Set t = 0 and Repeat the following steps until t < e, where t is the iteration rate and e is the error rate,

Step 4 – Get the Best Matching Unit.

Step 5 – Scale Neighbors

Step 6 – Increase t by small amount.

Step 7 – END.

4. IMPLEMENTATION AND RESULTS

This chapter presents the extensive simulation results for methods investigated in this project is Self-Organized Map clustering method on AT&T dataset.

a. Comparative Results

With almost 400 images, we experimented here with 40 people variations. As shown in Figure 4.1, a preview picture of the Faces database is shown.

Figure 4.1: AT&T sample

The weight vectors are initialized in a number of ways. The first is to give random values for its data between 0 and every weight vector as illustrated in Figure 4.1.

1. There are fewer iterations needed to create a good map and the analysis can be saved some time.

2. Best Matching Unit.

This is a very easy step, simply by crossing over all weight vectors and calculating the distance from the chosen sample vector. The winner is the shortest length weight. When more than one is at the same distance, the weight of the winner will be selected randomly among the smallest. A number of mathematical means determine the distance.

b. Scale Neighbors

The adjacent weights can actually be divided into two parts: what weights are regarded as nearest and how much each weight is closer to the sample vector. The second part of the scaling of the neighbor is learning. The winners are much more like a rewarded sample vector. The neighbors are more like the sample vector. A feature of this process of learning is the more the next person is distant from the winner, the less he learns. The rate at which you can learn weight is decreasing and as much as you want can be determined. So the winner is found once a weight is determined and every person is found. Table 4.1 shows the experimental observations of data sets experiments as follows:

Table 4.1: Experimental Observations

Criteria	Values
Training images	40
Testing images	360
Learning coefficient	0.5
Iterations	2000
Recognition Rate	87.34%

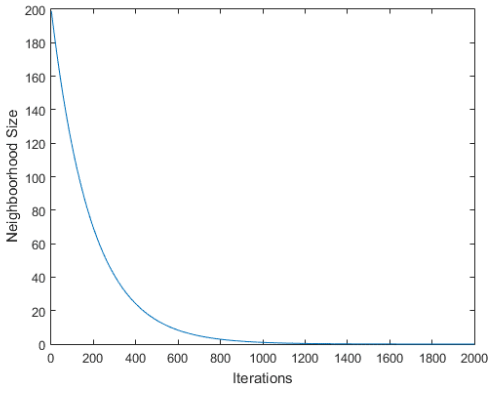

Figure 4.2: Error graph performance using SOM

5. CONCLUSION

In order to make the face matching in large databases, a novel self-organized retrieval system (SOM) is being suggested. The system offers a small subset of faces most similar to a certain query face from which users can easily check the images they correspond to. The system architecture includes two main elements. First, the system generally integrates multiple functional sets using several self-organizing maps. Secondly, a SOM is trained by the compressed feature vector to organize all face images in a database. The organized map can be used to efficiently identify similar faces to a query. SOM is a Face Recognition algorithm based on statistics.

An improved SOM method is suggested in this paper. For 40 persons' 400 AT&T database images, the highest average recognition rate of 87% is achieved, which takes place on only 30 images and is tested on other images. The proposed method is therefore an efficient process of face recognition.

5.1 SUGGESTIONS

In this research work, we perform the SOM algorithm training on AT&T. For future suggestions we want to perform below points:

- The current evaluation is based on single class problems for unsupervised learning, however under the real time scenario this will not be the case always. So we suggest evaluating the performance of proposed model by making it suitable for supervised learning.

Second point is, the consideration of more real time will be the interesting future direction for this research work.

REFERENCES

[1] J. Bartholdi and S. Hackman, Order Picking, 2014, pp. 137-.

[2] L. Yingde and J. Smith, "Dynamic slotting optimization based on SKUs correlations in a zone-based wave picking system," in 12th international material handling research colloquium , 2012.

[3] J. Bond, "Slot Leadership," Modern Materials Handling , 01 October 2015.

[4] R. Michel, "Slottings Elevated Place in an Omni-Channel World," Modern Materials Handling, November 2015.

[5] J. L. Haskett, "Cube-per-order index - a key to warehouse stock location," in Transportation and Distribution Management, 1963, pp. 27-31.

[6] C. Petersen and R. T. Kelly, "A comparison of warehouse slotting measures," in Annual Meeting of the Decision Sciences Institute, 2003.

[7] J. R. Mantel, P. C. Schuur and S. S. Heragu, "Order oriented slotting strategy for warehouses," in IIE annual conference and Expo 2007, 2007.

[8] E. Jones and T. Battieste, "Warehouse picking productivity by slotting inventory in the golden zone," in IIE Annual Conference and Expo, 2004.

[9] J. ,. Wu, T. D. Qin, J. Chen, H. P. Si, K. Y. Lin, Q. Zhou and C. B. Zhang, "Slotting optimization of the stereo warehouse," Advanced Materials Research, Vols. 756-759, pp. 1371-1376, 2013.

[10] Q. Zu, "Slotting optimization of warehouse based on hybrid genetic algorithm," 2011.

[11] B. S. Kim and J. S. Smith, "Slotting methodology using correlated improvement for a zone-based carton picking distribution system," Computers and Industrial Engineering , vol. 62, no. 1, pp. 286-295, 2012.

[12] M. Kofler, A. Beham, S. Wagner and M. Affenzeller, "Affinity Based Slotting in Warehouses with Dynamic Order Patterns," Advanced Methods and Applications of Computational Intelligence, vol. 6, pp. 123-143, 2014.

[13] S. Lozano, F. Guerrero, L. Onieva and J. Larrañeta, "Kohonen maps for solving a class of location allocation problems," European Journal of Operational Research, vol. 108, pp. 106-117, 1998.

[14] T. Kohonen, "Self-organized formation of topologically correct feature maps," Biological Cybernetics, vol. 43, no. 1, pp. 59-69, 1982.

[15] K.-H. Hsieh and F.-C. Tien, "Self-organizing feature maps for solving location-allocation problems with rectilinear distances," Computers and Operations Research, vol. 31, no. 7, pp. 1017-1031, 2004.

[16] H. L. R. Rushmeier and G. Almasi, "Visualizing customer segmentations produced by self-organzing maps," in IEEE Visualization Conference Proceedings, 1997.

[17] J. Vesanto and E. Alhoniemi, "Clustering of the Self-Organizing Map," IEEE Transactions of Neural Networks, vol. 11, no. 3, pp. 586-600, 2000.

[18] E. b. Orkun, E. Aydara, A. Gourgaudb, H. Artunerc and H. Bayhana, "Clustering of volcanic ash arising from different fragmentation mechanisms using Kohonen self-organizing maps," Computers and Geosciences, vol. 33, no. 6, pp. 821-828, 2007.

[19] M. M. Mostafa, "Clustering the ecological footprint of nations using Kohonen's self-organizing maps," Expert Systems with Applications, vol. 37, no. 4, pp. 2747-2755, 2010.

[20] M. M. Mostafa, "A psycho-cognitive segementation of organ donors in Egypt using Kohonen's self-organizing maps," Expert Systems with Applications , vol. 38, no. 6,6906-6915, 2011.

[21] A. N. Ahmad, C. Sekhar and A. Yashkar, "Music genre classification using music information retrieval and self-organizing maps," in 3rd International Conference on Soft Computing for Problem Solving , 2014.

[22] A. Stambuk, N. Stambuk and P. Konjevoda, "Application of Kohonen Self-Organizing Maps (SOM) Based Clustering for the Assessment of Religous Motivation," in 29th International Conference on Information Technology Interfaces, 2007.

[23] M. Chattopadhyaya, P. K. Danb and S. Mazumdarc, "Application of visual clustering properties of self-organizing map in machine-part cell formation," Applied Soft Computing, vol. 12, no. 2, pp. 600-610, February 2012.

[24] Y. Liu, R. H. Weisberg and C. N. Mooers, "Performance evaluation of the self-organizing map for feature extraction," Journal of Geophysical Research, vol. 111, 2006.

[25] S. Gabrielsson and S. Gabrielsson, "The use of Self-Organizing Maps in Recommender Systems," 2006.

[26] P. Georg, "Survey and Comparison of Quality Measures for Self-Organizing Maps," in Proceedings of the Fifth Workshop on Data Analysis, 2004.

[27] S. Shivam, T. N. Singh, S. V. K. and A. K. Verma, "Epoch determination for neural network by self-organized map (SOM)," Computational Geosciences, vol. 14, no. 1,199-206, 2012.

[28] R. Nuzzo, "http://www.nature.com/nature/index.html," Spring Nature, 12 February 2014. [Online]. Available: http://www.nature.com/news/scientific-method-statistical-errors-1.14700.

[29] AT&T Laboratories Cambridge, The database of faces at http://www.cl.cam.ac.uk/research/dtg/attarchive/facesataglance.html.

[30] Face Recognition using Self-Organizing Map and Principal Component Analysis, Dinesh Kumar, C.S. Rai and Shakti Kumar, 0-7803-9422-4/05, 2005 IEEE.

[31] Face Recognition: Convolutional Neural Network Approach, Steve Lawrence, C. Lee Giles, Ah Chung Tsoi, Andrew, IEEE Transactions on Neural Networks, Vol. 8, No. 1, January 1997.

[32] S. Albawi, O. Bayat, S. Al-Azawi, and O. N. Ucan, "Social Touch Gesture Recognition Using Convolutional Neural Network," Computational Intelligence and Neuroscience, 2018. [Online]. Available: https://www.hindawi.com/journals/cin/2018/6973103/abs/. [Accessed: 02-Apr-2019].

[33] T. A. Mohammed, A. alazzawi, O. N. Uçan, and O. Bayat, "Neural Network Behavior Analysis Based on Transfer Functions MLP & RB in Face Recognition," in Proceedings of the First International Conference on Data Science, E-learning and Information Systems, New York, NY, USA, 2018, pp. 15:1–15:6.

[34] O. M. Ahmed, O. Bayat, O. N. UÇAN, *PATTERN RECOGNITION USING NEURAL NETWORKS*, in AURUM Journal of Engineering Systems and Architecture, Istanbul, Turkey.