

# A Knowledge Acquisition Framework in Trace-Based Reasoning for Valuing Knowledge

نظام استقصاء المعرفة مستند على الاستنتاج التتبعي لتقييم المعرفة

By

**Dareen Sayed Khattab** 

Supervisor

Dr. Hussein H. Owaied

A Thesis Submitted in Partial Fulfillment of the Requirements for the Master Degree

in Computer Science

Faculty of Information Technology

Middle East University

Amman, Jordan

June, 2012

### **AUTHORIZATION FORM**

إقرار تفويض

أنا دارين سيد خطاب أفوض جامعة الشرق الأوسط بتزويد نسخ من رسالتي للمكتبات أو المؤسسات أو الهيئات أو الأفراد عند طلبها.

التوقيع: عنه ا

التاریخ: > / ۹ / ۱ . . .

### **AUTHORIZATION FORM**

I, Dareen Sayed Khattab, authorize Middle East University to supply a copy of my thesis to libraries, establishments or individuals upon their request.

Signature: 1 2/9/2012

### **EXAMINATION DECESION**

This is to certify that the thesis entitled "A Knowledge Acquisition Framework in Trace-Based Reasoning for Valuing Knowledge" was successfully defined and approved on June/2102.

### **Examination Committee Members**

Signature

Dr. Hussein H. Owaied Department of Computer Science Middle East University

Dr.Ala Abu Samaha Department of Computer Information System Middle East University

(2/2/10)

Prof. Musbah J. Aqel Department of Computer Science Al Zarqa'a Private University

Must

#### **DECLARATION**

I do hereby declare the present research work has been carried out by me under the supervision of Dr. Hussein Owaied, and this work has not been submitted elsewhere for any other degree, fellowship or any other similar title.

Dareen Sayed Khattab Department of Computer Science Faculty of Information Technology Middle East University

### **DEDICATION**

# (يَرْفُع اللَّهُ الَّذِينَ آمَنُوا مِنكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ)

Almighty Allah says "Allah will raise up, to (suitable) ranks (and degrees), those of you who believe and who have been granted knowledge". May Allah raise us ever after to the highest degree in Iman and knowledge.

I dedicate this work to my great parents, my two brothers, my relatives, my friends, and all those who helped, supported, and taught me.

### **ACKNOWLEDGMENTS**

I would like to thank my father and my mother for their continuous support during my study.

I also would like to thank my great supervisor Dr. Hussein Owaied for his support, encouragement, proofreading of the thesis drafts, and for helping me throughout my studies, putting me in the right step of scientific research. I would like to thank the Information Technology Faculty members at Middle East University. I would also like to thank all of my family members specially my brothers Esam & Adel, Aunt Nawal and her children, and all of my friends specially Heba, Umayia Murad, Shima'a, Safa'a, Zainab, and Alaa.

### **Abstract**

The purpose of this thesis is to build A Knowledge Acquisition Framework in Trace-Based Reasoning for Valuing Knowledge. The knowledge Acquisition Framework consisting of context information retrieval from proposed algorithm in the first stage, then an adaptive neuro-fuzzy model for the second stage, which can be trained to detect the value of knowledge used. The training has been based on gathered surveyed data. After training the model with proper data, a clear target-oriented towards the best usage of knowledge will be available. Final stage will be implicitly processed via a back propagation feature exists in the neuro-fuzzy model mentioned above. Trace Based Reasoning is used in this framework instead of Case Based Reasoning which had been used for solving problems previously, due to the problem of lacking to context information in Case Based Reasoning .In this study six models have been developed for the second stage with different types of input/output membership functions and trained an input array. The models are compared based on their ability to train with

lowest error values. The Gaussian member function input with either constant or linear Sugeno output member function was the best choice for the proposed framework to be adopted in its second stage which is Task Analysis Module. This framework can be utilized in firms, societies or even in individuals' life events.

### الملخص

الهدف من هذا البحث هو بناء اطار متكامل لاستقصاء المعرفة مستندا على الاستنتاج التتبعي لتقييم المعرفة و يمر هذا الاطار بثلاث مراحل :المرحلة الأولى مرحلة استرجاع المعلومات من السياق المتوفر، المرحلة الثانية باستخدام النظام المتكيف من المنطق الضبابي و الشبكات العصبونية بحيث يتكيف النظام مع التمرين لتقييم المعرفة المستخدمة في الشركات و المجتمعات و حتى في مواقف حياة الأفراد . وقد استخدمت بيانات مسحية مجمعة لتدريب النموذج، و بعد تدريب النموذج يصبح بالامكان الكشف عن الاستخدام الموجه للهدف المنشود من المعرفة المستخدمة .

و المرحلة الثالثة و التي يتم معالجتها ضمنيا في مرحلة المنطق الضبابي مع الشبكات العصبونية كصفة موجودة مسبقا ضمن ازدواجيتهما وهي التغنية الراجعة.وقد تم الاستناد بهذا البحث على الاستنتاج التتبعي بدلا من الاستنتاج المبني على الحالات المبنية من مواقف سابقة،وذلك لفقر الحالات السابقة التي يبنى عليها الاستنتاج الى وجود المعرفة المتوفرة بسياق المشكلة وانما وجود فقط مواقف قديمة مماثلة.

لقد استخدم نظام MATLAB ANFIS لبناء ستة نماذج مختلفة في المرحلة الثانية للاطار المقترح، و يكمن اختلاف هذه النماذج باختلاف انواع الاقترانات المستخدمة لتمثيل المدخلات و المخرجات الفعلية لهذه النماذج عن طريق مقدرتها على التعلم ثم مقارنة مقدار الخطأ الناتج من التمرين من كل اقتران.و قد حصل النموذج ذو الاقتران الغاوسي للبيانات المدخلة و الاقتران الخطي أو الثابت للبيانات المخرجة على أقل قيم للخطأ لذلك سيكون الخيار الأفضل لاستخدامه بالمرحلة الثانية من الاطار المقترح. ويمكن استخدام هذا الاطار و تطبيقه بالشركات والمجتمعات وحتى بمواقف الحياة لدى الافراد.

# **List of Figures**

Figure 1. 1 The Knowledge Maturing Process (Schmidt,2005)
Figure 3. 1 (R&D) and Patenting Time Series Relationship (Hall,2004)1
Figure 3. 2 Trend in total patent applications across the world through years
(1985 – 1998) (WIPO indicators,2010)
Figure 3. 3 The evolution of the patents and patent growth in semiconductor
technology space (Dibiago & Nasiriya 2008)
Figure 3. 4 The intellectual capital types effect on four sample firms (Miller
et al.,1999)
Figure 4. 1 The main process of the proposed framework
Figure 4. 2 The input membership function for Axioms
Figure 4. 3 The input membership function for Network Effects
Figure 4. 4 The input membership function for Cumulativity
Figure 4. 5 The input membership function for Sources of Knowledge
Figure 4. 6 The input membership function for Context of Knowledge
Figure 4. 7 The input membership function for Six Challenges of Knowledge
Management
Figure 5. 1 Fuzzy model of the 6 input mfs
Figure 5. 2 The Structure of the j48 rules based neuro/fuzzy model
Figure 5. 3 Fuzzy Inference System (Shing and Jang, 1993)

Figure 5. 4 The confusion matrix using Cross Validation from WEKA
program(snapshot)
Figure 5. 5 The confusion matrix using Percentage Split from WEKA
program (snapshot)
Figure 5. 6 The Structure of the j48 classification tree
Figure 5. 7 Axioms Generalized Bell/Constant model after training
Figure 5. 8 Network Effects Generalized Bell /Constant model after training 1
Figure 5. 9 Context of Knowledge Generalized Bell/Constant model after
training
Figure 5. 10 Context of Knowledge Generalized Bell/Linear model after
training
Figure 5. 11 Context of Knowledge Gaussian/Constant/Linear model before
training
Figure 5. 12 Context of Knowledge Gaussian/Constant/Linear model after
training
Figure 5. 13 Context of Knowledge Gaussian2/Constant/Linear model before
training
Figure 5. 14 Context of Knowledge Gaussian2/Constant/Linear model after
training

# **List of Tables**

Table 3. 1 Revision results for OWL DL Axiom Ontology (Nikitina et al.,
2011)
Table 3. 2 Evolution of cumulativeness of technological advance (Dibiago &
Nasiriyar,2008)
Table 3. 3 The questionnaire items pertaining to the survey (Detlor et al.,
2006)
Table 3. 4 Matrix of loadings and cross-loadings of the survey's items (Detlor
et al.,2006)
Table 3. 5 Construct reliability and variance extracted for survey's items
(Theriou et al.,2010)
Table 5. 1 Model Specifications
Table 5. 2 Training Input Array
Table 5. 3 Axioms Generalized Bell/Constant Model
Table 5. 4 Network Effects Generalized Bell/Constant Model 56
Table 5. 5 Context of Knowledge Generalized Bell/Constant Model 57
Table 5. 6 Context of Knowledge Generalized Bell/Linear Model 59
Table 5. 7 Context of Knowledge Gaussian/Constant or Linear Model 61
Table 5. 8 Context of Knowledge Gaussian 2 Constant/Linear Model 63
Table 5. 9 Error values after testing the models for constant/Linear models 66

XIV

### **List of Abbreviations**

AHighI Axioms High Impact

AI Artificial Intelligence

ALowI Axioms Low Impact

ANFIS Adaptive Neuro Fuzzy Inference System

ANN Artificial Neural Netwrok

B2C Business to Customer

CBR Case-Based Reasoning

CHighE Cumulativeness High Effect

CLowE Cumulativeness Low Effect

CVKM Construct Validity of Knowledge Management

ECapitalE External Capital Effect

FET Future and Emerging Technologies

FL Fuzzy Logic

GUI Graphical User Interface

HCapitalE Human Capital Effect

KM Knowledge Management

KME Knowledge Management Environment

KM1 Knowledge Management 1

KM2 Knowledge Management 2

KM3 Knowledge Management 3

KM4 Knowledge Management 4

Mf Member Function

NHighE Network High Effect

NLowE Network Low Effect

OIB Organizational Information Behavior

OWL Web Ontology Language

PIB Personal Information Behavior

R&D Research and Development

RMSE Root Mean Square Error

SCapitalE Structural Capital Effect

TBR Trace Based Reasoning

UN United Nations

VEKM Variance Extracted of Knowledge Management

WIPO World Intellectual Property Organization

# **Table of Contents**

AUTHOR	RIZATION FORM (in Arabic)II
AUTHOR	RIZATION FORM (in English) III
EXAMIN	IATION DECISIONIV
DECLAR	ATIONV
DEDICA'	TIONVI
ACKNOV	WLEDGMENTSVII
Abstract i	n EnglishVIII
Abstract i	n Arabic X
List of Fig	guresXI
List of Ta	iblesXIII
List of Ab	obreviationsXIV
Table of C	ContentsXVI
Chapter C	One: Introduction1
1.1.	Overview
1.2.	Motivation
1.3.	Problem Definition
1.4.	Objectives4
1.5.	Thesis Outline

Cł	napter Tv	wo: Literature Review and Related Studies	6
2.1		Overview	6
2.2	•	Context Information Retrieval	6
2.3	•	Adaptive Neuro-Fuzzy Inference Systems (ANFIS)	11
2.4	•	Valuing Knowledge	15
Cł	napter Tl	nree: Valuing Knowledge Mnagement	18
3.1		Overview	18
3.2		Valuing Knowledge	18
	3.2.1.	Axioms	19
	3.2.2.	Netwrok Effects	20
	3.2.3.	Cumulativity	22
	3.2.4.	Sources of Knowledge	24
	3.2.5.	Context of Knowledge	25
	3.2.6.	Six Challenges of Knowledge Management	27
	C	hapter Four: Knowledge Acquisition Framework Design &	
Implemen	ntations	3	30
4.1		Introduction	30
4.2		Framework Architecture	31
4	4.2.1.	Inference Network	32
4	1.2.2.	Task Analysis Module	34

4.2.2.1. Membership Functions of Input Factors	34
Axioms	34
Network Effects	35
Cumulativity	36
Sources of Knowledge	37
Context of Knowledge	38
Six Challenges of Knowledge Management	40
4.2.2.2. Rules and Output Membership Functions	42
4.2.3. Relevance Feedback	42
Chapter Five: Experimental Study	43
5.1. Overview	43
5.2. Neuro-fuzzy Models	43
5.3. Training	47
5.4. Results	50
5.4.1. J48 Classification Results	50
5.4.2. The impact of training array	53
5.4.3. Performance	65
Chapter Six: Conclusion and Future Work	67
6.1. Conclusion	67
6.2. Future Work	68

References 60	ferences 6	9
---------------	------------	---

# **Chapter One**

### Introduction

#### 1.1. Overview

Solving problems is one of many tasks that are strongly related to the survival of human being. There are many methods for solving problems, and there are many differences between these methods used from different perspectives and factors such as the kind of the problem, the domain and the problem space. Considering the problem space representation, most of the problem solving methods are relying on the problem space representation and depends even if slightly on similar problem solved or observed in past experiences (Owaied,2010).

Case-based reasoning (CBR) is one of the methods in solving problems that all reasoning is based on past cases personally experienced. But depending only on the past experience is not enough to solve some problems, what makes a main problem of the case-based reasoning to appear is the lack of relevant context information in the problem space to be considered in solving new problems (Cordier, 2008)

A Macro model presented by (Schmidt, 2005), states that how important the context-aware systems are in supporting learning processes. An example of such systems is the "The Knowledge Maturing Process" with its five stages shown below in Figure 1.1.

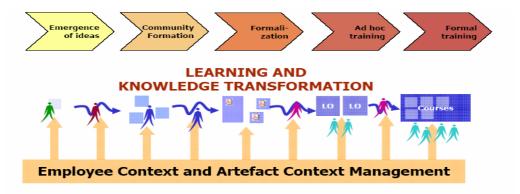


Figure 1. 1 The Knowledge Maturing Process (Schmidt, 2005)

The main important conclusion obtained from this model is that by determining the considered context and relevant artifacts, the system can help the learner in making best use of existing information. Therefore, the proposed framework in this research will focus on how to identify the relevant context information and how to use it efficiently which means to extend the (CBR) and use (TRB) instead when solving problems.

#### 1.2. Motivation

Using the available information in the domain of any problem, can be very useful in finding the suitable solution(s) for that problem and in an appropriate time. However, ignoring these information will lead to inaccurate results in the solution(s) of the problem, and will waste the time to obtain the exact solution. The aim of the proposed framework is to help human in any situation in life to exploit each available data and information within the problem domain in order to get more accurate, efficient, and exact solution(s) to his problem, and this framework will be utilized in this research work for the purpose of valuing

knowledge of an organization as per knowledge is considered as intangible assets to any firm or organization, and has its valuable effect in enhancing its operations, but its existence in firms and organizations is not an explicit one, therefore, via this research the knowledge valuation will be presented in numbers to having it explicitly existed in the firms and organizations.

Another benefit from this proposed model is to use these gathered information in an efficient way, by adapting the whole integrated stages within this proposed system (Artificial Neural network (ANN), Fuzzy Logic, Trace Based Reasoning (TBR) and Relevance Feedback) in the knowledge acquisition and adaptation process.

### 1.3. Problem Definition

According to the definition of (CBR), solving problems is based on the solutions of similar past problems. From the system's point of view, this might be true, but from the user's point of view, identical problems may need different solutions. This is due to that (CBR) suffers from the "frame problem": in some situations, the context information is missing.

Moving from the Case-Based Reasoning to Trace-Based Reasoning (TBR) is the solution of this problem. Trace-Based Reasoning is an extension of the Case-Based Reasoning, allowing the context to be included in the reasoning, but this actually will lead to many different problems to be identified as following:

- 1. How to identify relevant context information in traces.
- 2. How to make sure all the elements we need are in the trace and then use them by an efficient model to solve the faced problems.
- 3. How to utilize this proposed framework in valuing knowledge in a firm or in an organization, by transferring the intangible factors that are needed to valuate knowledge in an organization into numbers, in order to help understanding how an organization's knowledge adds value to its operations and thus enabling informed management of its knowledge assets.

### 1.4. Objectives

The main objective of this research work is to build a knowledgeacquisition framework, which is capable to achieve the following:

- Interacting with each element in the environment for a specific task to be fulfilled.
- Identifying the context information in traces for each problem faced during the adaptation process.
- 3) Using these traces in knowledge valuation process, which includes transferring the factors affecting knowledge valuation into numbers, by using an integrated framework of Artificial Neural Network(ANN), Fuzzy Logic(FL), Trace Based Reasoning (TBR) in tracing records of activities and a Relevance Feedback algorithm.

### 1.5. Thesis Outline

The rest of the thesis is organized as follows: Literature review is presented in Chapter 2.In Chapter 3, is an explanation of the methodology used in this study; particularly, various factors which impact the knowledge valuation process. In Chapter 4, the three stages of the proposed framework is presented including the ANFIS model. In Chapter 5, an experimental study and results will be presented obtained after applying the proposed model. Conclusions and future work are presented in Chapter 6.

## **Chapter Two**

### **Literature Review and Related Studies**

#### 2.1. Overview

The proposed framewrok includes several areas of study including Context Information Retrieval, Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Knowledge Valuation. Following is a brief literature review of the areas covered in this thesis.

#### 2.2. Context Information Retrieval

Context retrieval information as a first stage in the proposed framework, is presented in order to extend CBR to TRB in solving problems methodologies, and this is done by including the context of information in the problem domain in the process of solving problems, many researchers have been concentrating via their works on the field of context information retrieval by many different methods and using different techniques. Following is a brief of the related works to this thesis content.

Salton & Buckely (1990) declared that relevance feedback is an automatic process, introduced over 20 years ago, and designed to produce improved query formulations following an initial retrieval operation. The principal relevance feedback methods described over the years are examined briefly, and evaluation data are included to demonstrate the effectiveness of the various

methods. Prescriptions are given for conducting text re-trieval operations iteratively using relevance feedback.

Budzik et al. (2001) claimed that user interactions with productivity applications (e.g., word processors, Web browsers, etc.) provide rich contextual information that can be leveraged to support just-in-time access to task-relevant information. As evidence for their claim, they presented Watson, a system which gathers contextual information in the form of the text of the document the user is manipulating, in order to proactively retrieve documents from distributed information repositories related to task at hand, as well as process explicit requests in the context of this task. They described the results of several experiments with Watson, which consistently has provided useful information to its users.

Other researchers have addressed how useful the contextually retrieved information in search queries as Sieg et al. (2005) stated that one of the key factors for accurate and effective information access is the user context. The critical elements that make up a user's information context include the semantic knowledge about the domain being investigated, the short-term information need as might be expressed in a query, and the user profiles that reveal long-term interests. Sieg et al. (2005) propose a framework for contextualized information access that seamlessly combines these elements in order to effectively locate and provide the most appropriate result for users' information needs. In particular, they focused on integrating a user's query with semantic knowledge from an existing concept hierarchy to assist the user in information retrieval. In their framework, the user's "context" is captured via nodes in a concept lattice induced from the original

ontology and is updated incrementally based on user's interactions with the concepts in the ontology. Their experimental results showed that utilizing the user context improves the effectiveness of the search queries, especially in the typical case of Web users who tend to use very short queries. A term-vector based representation is used for concepts. To generate a term-vector representation, the content of all the associated relations with the concept are combined to yield a single term-vector. A weighted term-vector its symbol is  $\mathbf{n}_i$  for each concept i. Each concept contains a collection of relations  $R_i$ , and a set of sub-concepts  $S_i$ .

Thus the user context is represented as a pair of elements:  $ci = \{P,N\}$ , where P is a term-vector of positive evidence (min operation) :  $P = \min(n1,n2)$ , and N is a term-vector of negative evidence (max operation) :  $N = \max(n1,n2)$ . The min and max operations could be extended to more logical operations intersection and union operations, respectively. Thus, the positive evidence will be represented as  $P = n1 \cap n2 \cap n3 \cap .... \cap nk$  and the negative evidence will be represented as  $P = n1 \cap n2 \cap n3 \cap .... \cap nk$ . Each time the user interacts in the specific domain seeking more information, the user's short term interest as a context ci, which is a pair of positive and negative evidence. In order to represent the user\s long-term context, i.e. the user profile as a set of contexts:  $P = \{c0, c1, c2, ...., cn\}$ . Depending on user behavior, a specific context in the user profile can be updated or a new context can be added.

Hardian et al. (2006) stated that application autonomy can reduce interactions with users, ease the use of the system, and decrease user distraction. On the other hand, users may feel loss of control over their applications. A further

problem is that autonomous applications may not always behave in the way desired by the user. To mitigate these problems, autonomous context-aware systems must provide mechanisms to strike a suitable balance between user control and software autonomy. Hardian et al. (2006) presented a survey of research on balancing user control and system autonomy in context-aware systems. They addressed various issues that are related to the control-autonomy trade-off, including issues in context modeling, programming models and tools, and user interface design.

Soules (2006) described that personal data is growing at ever increasing rates, fueled by a growing market for personal computing solutions and dramatic growth of available storage space on these platforms. Users, no longer limited in what they can store, are now faced with the problem of organizing their data such that they can find it again later. Unfortunately, as data sets grow the complexity of organizing these sets also grows. This problem had driven a sudden growth in search tools aimed at the personal computing space, designed to assist users in locating data within their disorganized file space. Despite the sudden growth in this area, local file search tools are often inaccurate. These inaccuracies have been a long-standing problem for file data, as evidenced by the downfall of attribute-based naming systems that often relied on content analysis to provide meaningful attributes to files for automated organization. While file search tools have lagged behind, search tools designed for the World Wide Web have found wide-spread acclaim. Interestingly, despite significant increases in non-textual data on the web (e.g., images, movies), web search tools continue to be effective. This is because

the web contains key information that is currently unavailable within file systems: context.

Continuous developments of mobile technologies and their use in everyday life increase our need to be continuously connected to others and to the Internet, anywhere and at any time. However, in mobile, pervasive environments user connectivity is mainly affected by wireless-communications constraints and user mobility. These boundary conditions do not allow us to design communication environments based on unique and fully connected networks or assume a stable path between each pair of users wishing to communicate. Delmastro et al. (2010) introduced opportunistic networking which has emerged as a new communication paradigm to cope with these problems. It exploits user mobility to establish communications and content exchange between mobile devices in pervasive, mobile computing environments. Content sharing (of either information available on the Internet or user-generated resources) through, for example, YouTube or Flickr currently represents one of the most popular services. Thus, users are becoming the principal actors of the network, particularly in mobile environments. Efficient development of this kind of service in opportunistic networks imposes mobility support, requiring knowledge of user context and social behavior. Therefore, information about the network's users and their habits, interests and social interactions plays a fundamental role, allowing the system to generate routes on the fly to correctly deliver messages to the intended recipients.

The social- and context-aware content-sharing service that Delmastro et al. (2010) designed and developed in the framework of the European Commission's

Information Society Technologies/Future and Emerging Technologies (FET) Haggle project exploits a context definition designed for opportunistic networks. The main idea is that each user who wants to participate in the service can declare information about the contents she wants to share, as well as a certain amount of personal information that enables the system to trace her social interactions and mobility patterns.

Fritz (2011) introduced in his thesis that a software developer must continuously search for the small portions of information pertinent to his work within the flood of project information. He added "today's artifact-centered development environments make finding the needed information tedious or infeasible". In his research, he introduced two models, the degree-of-knowledge model and the information fragments model. These two models showed that it is possible to add developer-centric models to a development environment and ease a developer's access to the information relevant to work-at-hand addressing the developer's individual information needs.

### 2.3. Adaptive Neuro-Fuzzy Inference Systems (ANFIS)

Combining the ANN features and Fuzzy Logic rules, the Hybrid ANFIS system was presented and have been used frequently in modeling and solving problems in computer science and other related fields, past few decades have seen a resurgent trend towards establishment of intelligent manufacturing systems which are capable of using advanced knowledge-bases and intelligence techniques in aiding critical operational procedures in manufacturing.

Khosravi and Lu (2006) developed a new method to model occurred faults in different parts of nonlinear systems. Using an Adaptive NeuroFuzzy Inference System (ANFIS) they built a model for faultless plant which is used in the procedure of fault modeling. The considered model for fault is again an ANFIS system and its parameters are adjusted in an indirect way using difference between actual output and output of plant model. Simulation results on a nonlinear system were shown in their work and they clearly demonstrated the capability of the proposed method for fault modeling. Multiple inputs single output models were developed to predict radial expansion ratio, unit density, bulk compressibility and spring index of the nanocomposite foams. An individual ANFIS model was developed by Lee et al. (2008) each mechanical property using clay content, temperature, pressure and torque as input parameters.

Increasing demands on productivity and quality with the increase in global competitiveness have necessitated development of sound predictive models and optimization strategies. Sivarao et al. (2009) presented the modeling technique and prediction of surface roughness for Manganese Molybdenum pressure vessel plate by Hybrid Intelligence, namely, adaptive neuro-fuzzy inference system (ANFIS). Back propagation optimization method has been employed to optimize the epoch number and training of data sets. To compare the accuracy of the ANFIS model, the errors were calculated through Root Mean Square Error (RMSE) which yielded 0.3 and below. On the other hand, the prediction accuracy by the finalized ANFIS model had yielded up to 90% and above proving the prediction stability. The uniqueness of this modeling technique is that, all modeling, variable selection,

model validation, prediction, etc. was done using a graphical user interface (GUI) developed using Matlab. The non-traditional laser machining, was used in the modeling investigation as this machining process requires controlling of more than seven critical parameters and to date, no researchers has used ANFIS to model this exact phenomenon. The modeling technique has been successfully developed to predict the cut edge quality with excellent degree of accuracy and strongly belief that ANFIS could be the best hybrid AI tool with the capability of data training and rule setting which has to be further explored with critical consideration in producing precise part of any material in the field of precision manufacturing. The RMSE values were compared with various training variables to develop the best predictive model yielding 0.3 and below. The model was then used to predict the surface roughness and the prediction accuracy obtained was above 90% proving the optimizing technique and methods were accurate in producing excellent ANFIS model.

With the rapid development of Internet, the number of online customers is growing fast. This growth is supported by spreading of Internet usage around the globe. However, the question of security and trust within e-commerce has always been in doubt. It was Nilashi et al. (2011)'s study specifically gave an overview to understand different factors about security and trust between companies and their consumers. In order to Three e-stores and their websites were examined based on the model proposed. Nilashi et al.'s study also mentioned that security and trust work parallel and close to each other. If a consumer feels that an online deal is secured and they can trust the seller, it leads to a confident e-commerce's trade. The

main focus of this study is to find out a suitable way to resolve security and trust issues that make e-commerce an uncertain market place for all parties.

As a result of Nilashi et al.' work the character of security is regarded as the most important to building trust of B2C websites. The proposed model applied Adaptive Neuro-Fuzzy model to get the desired results. Two questionnaires were used in this study. The first questionnaire was developed for e-commerce experts, and the second one was designed for the customers of commercial websites. Also, Expert Choice is used to determine the priority of factors in the first questionnaire, and MATLAB and Excel are used for developing the Fuzzy rules. Finally, the Fuzzy logical kit was used to analyze the generated factors in the model.

Chaudhari et al. (2012) contributed to compare the results of decision making of maximizing profit in farm cultivation namely rice using ANFIS model and Multi Objective Linear Programming Problem by optimization method. Data is uploaded and tasted for training. ANFIS rule base is auto generated for determining the better performance of the model. The performance of the ANFIS model is evaluated in terms of training performance and classification accuracies and the results confirmed that the proposed ANFIS model is useful tool in decision making. The farmer can take decision about the expenditure to be made on various heads in farm cultivation considering uncertainties up to maximum extent and get maximum yield in order to maximize the profit. This model will help the farmer to choose the appropriate quantity of input variables and make the necessary arrangements of farm cultivation to decide the quantity purchase and expenses to be made in advance.

#### 2.4. Valuing Knowledge

Considering Knowledge as main assets in companies, organizations and even in our life as indivisuals, it needs to be managed effectively and this is done by expressing it explicitly, many researchers had tried figuring this idea out as listed below.

Ongoing transition of United Nations Member States to knowledge-based economies is a watershed event in the evolution of the global knowledge economies. This transition marks a paradigmatic shift from energy-based economies with traditional factors of production to information based economies based upon knowledge assets and intellectual capital. As envisioned in the UN Millennium Declaration, development of national knowledge societies should encompass social, cultural, and human development besides economic growth. Accordingly, one objective of Malhotra et al. (2003)'s study is to develop the theoretical and pragmatic foundations for management and measurement of knowledge assets to facilitate this vision of holistic growth and development. Based upon a review of theory, research, practices, and national policies, they critically analyzed and contrasted the most popular models available for measurement of national knowledge assets. Their review includes knowledge modeling and measurement frameworks and their applications by reputed developmental organizations and national governments. There are two other key outcomes of the above review and analysis. First, to build the capacity of the public sector for measuring and managing knowledge assets, they proposed, developed, and defined specific frameworks, methodologies, models and indicators with illustrative real world applications. Second, they made specific recommendations for necessary improvements needed in knowledge assets management and measurement models and indicators. Prudent and effective policy directives depend upon pragmatic but theoretically and psychometrically valid measurement for their success. They recommended that the future development of such models be based upon better understanding of human capital and social capital as well as their synthesis with existing intellectual capital frameworks and models. The findings and recommendations of this study will provide the cornerstone for measuring and managing national knowledge assets for United Nations Member States toward holistic socio-economic development.

Carlucci et al. (2004)'s theoretical paper explored the fundamental issue of how knowledge management initiatives impact business performance. Reflecting on the management literature in the fields of knowledge management and performance management enabled the deduction of four basic assumptions, representing the links of a conceptual cause-and-effect framework – the knowledge value chain. Drawing on the resource-based view and the competence-based view of the firm, the paper identified strategic, managerial, and operational dimensions of knowledge management. The review of performance management frameworks discussed the role of knowledge management in those models. These reflections allow linking knowledge management with core competencies, strategic processes, business performance, and finally, with value creation.

Piller and Christian (2009) stated that "the fact that we ought to prefer what is comparatively more likely to be good, I argue, does, contrary to consequentialism, not rest on any evaluative facts. It is, in this sense, a deontological requirement. As such it is the basis of our valuing those things which are in accordance with it. We value acting (and believing) well, i.e. we value acting (and believing) as we ought to act (and to believe). In this way, despite the fact that our interest in justification depends on our interest in truth, we value believing with justification on non-instrumental grounds. A deontological understanding of justification, thus, solves th Value of Knowledge Problem". To survive and flourish in a changing and unpredictable world, organizations and people must maintain strategic power over necessary resources - often in the face of competition.

Knowledge is constructed, used and evaluated via cyclically-iterated processes. Hall et al. (2011) introduced nine time-based frames of reference based in this Popperian autopoietic paradigm to explore the relationships between time and a utility-based valuation of knowledge as it is constructed and applied. They believe this framework and associated paradigmatically consistent vocabulary provide useful tools for analyzing organizational knowledge management needs.

# **Chapter Three**

## Valuing Knowledge Management

#### 3.1. Overview

The proposed framework is based on managing the knowledge valuation via different factors which in turn they affect getting the desired value of the available knowledge. In this chapter an overview of these factors will be viewed. In addition a view of how important knowledge management is in firms and business relations will be listed.

## 3.2. Valuing Knowledge

Current economic crisis is leading all the companies and organizations to have functional units that should do the management of information and knowledge related activities as basic standards and the highest priorities in business (Malhotra, 2003).

The aim of knowledge valuation ontology is allowing the users to express factors relevant to valuing a particular piece of knowledge (O'Hara & Shadbolt, 2001). Since an artificial neural network will be used to allow the system to adapt various inputs of the factors will be illustrated below, figures will be used and results from surveys and questionnaires for each factor, in order to express each effect in a digital data processing step in the proposed system.

Following is a brief description of the various factors and their potential impact on valuing knowledge.

#### **3.2.1. Axioms**

As per (O'Hara & Sadbolt,2001)'s comment cited from Fox and Gruninger ,1999,p.111 that retrieval of information not directly stored in the data base does not require wider search characteristic if ontologies stored the means for relatively straightforward deductions within themselves, i.e. by using axioms.

There are five kinds of components used to specify knowledge in ontologies: concepts, relations, functions, axioms and instances. Axioms are model sentences that are always true. Their existence in an ontology is to constrain its information, verify its correctness or deduce new information (Gruber, 1993).

Table 3.1 illustrates one of the developed methodology which is an ontology-supported literature search for is specified in the Web Ontology Language OWL DL (OWL Working Group,2009). Tools have been employed for automated textual analysis to produce a set of document annotations, which was then manually evaluated. Six distinct annotation sets  $S_1$  to  $S_6$  using different annotation methods for 2,289 logical axioms.

The results of this methodology was that the decision space (means keeping tracking of the dependencies between axioms) saved about 75% of reasoned calls and the appropriate choice of axioms leads to a better performance (Nikitina et al.,2011).

	S	(54, 9	4%)	$S_2$ (60, 100%)			
$impact^+$	69%	4,677	36,773	83%	2,584	18,702	
guaranteed	48%	11,860	51,677	65%	8,190	55,273	
$impact^-$	9%	17,828	46,461	12%	20,739	67,625	
upper bound	74%	4,110	11,399	83%	2,645	27,850	
random	45%	-	1,291	60%	-	1,090	
	S	3 (40, 4	5%)	.5	$S_4$ (35, 4	18%)	
$impact^+$	20%	3,137	26,759	29%	2,198	15,601	
guaranteed	43%	3,914	27,629	43%	3,137	18,367	
$impact^-$	28%	9,947	46,461	31%	7,309	10,217	
upper bound	48%	3,509	13,202		2,177	7,002	
random	31%	-	764	31%	-	534	
	S	$f_5$ (26, 20	6%)	5	$S_6$ (72, 1	2%)	
$impact^+$	8%	1,778	11,443	13%	9,352	212,041	
guaranteed	39%	1,290	6,647	54%	8,166	99,586	
$impact^-$	54%	954	1,438	76%	6,797	16,922	
upper bound	54%	801	1,989	76%	5,219	19,861	
random	41%	-	212	57%	-	1,065	

Table 3. 1 Revision results for OWL DL Axiom Ontology (Nikitina et al. 2011)

#### 3.2.2. Network Effects

Network effects are characteristic of advanced technology and information based sectors of the economy. The more a piece of knowledge is used, the more valuable it is(O'Hara and Shadbolt,2001).

The added value in every incident of networking lies in its contributions to the knowledge of the participants and to the enhancement of its value to them (Choucri, 2007).

(R&D) is one of a corporate activity, as a mutually beneficial formal relationship between two or more parties, i.e. via network activities for increasing the stock of knowledge (Wikipedia, 2012).

Figure 3.1 shows the strong correlation between patents and (R&D) (Hall, 2004).

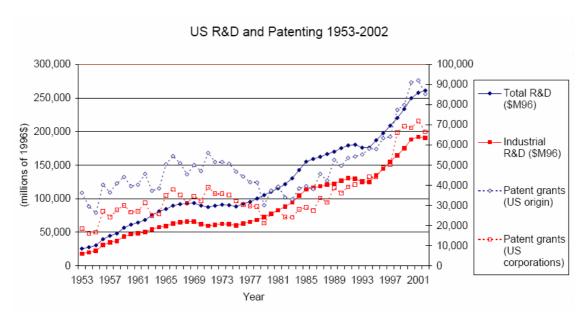


Figure 3. 1 (R&D) and Patenting Time Series Relationship (Hall,2004)

Referring to WIPO (World Intellectual Property Indicators, 2010) and as shown in Figure 3.2 below, overviews the direct proportional relationship between patent applications across the world versus years (1985 - 2008).

The overall percentage growth rate was positive through years excluding some slowdown periods had been occurred due to the global economic decline in that time which was in 2008.

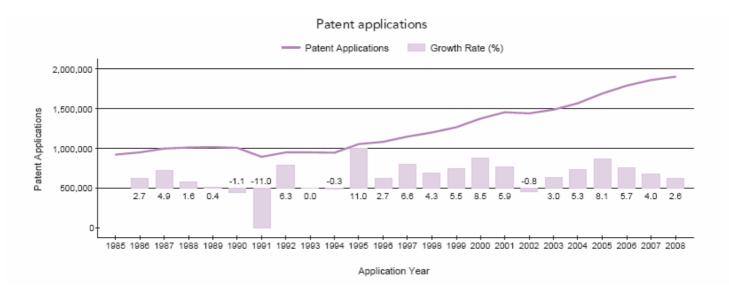


Figure 3. 2 Trend in total patent applications across the world through years (1985 – 1998) (WIPO indicators, 2010)

## 3.2.3. Cumulativity

To understand and acquire a piece of knowledge is strongly influenced by other pieces of knowledge that are related to it (O'Hara and Shadbolt, 2001).

Jeffrey et al. (2006) mentioned that the cumulative nature of the knowledge is recognized as central to economic growth. Using the cumulative nature of innovation development in the semiconductor industry, an analysis was achieved indicating how much new innovative outputs (patents) are based on already existing technological knowledge. Table 3.2 shows the correlation coefficient for each year which was calculated at first by calculating the intensity of each technological combination, and then correlating the combination vector of each year with the observations of the previous year (Dibiago and Nasiriyar 2008).

Year	Cumulativeness	Year	Cumulativeness
1963	.5288	1980	.9195
1964	.7443	1981	.8862
1965	.8288	1982	.8971
1966	.8532	1983	.8802
1967	.8581	1984	.894
1968	.8621	1985	.9245
1969	.8713	1986	.9359
1970	.8731	1987	.9439
1971	.8873	1988	.9353
1972	.8777	1989	.9442
1973	.8538	1990	.9575
1974	.8716	1991	.9684
1975	.8917	1992	.9746
1976	.9057	1993	.9723
1975	.8917	1994	.9737
1976	.9057	1995	.9793
1977	.9006	1996	.9822
1978	.9074	1997	.9899
1979	.8982	1998	.9816

Table 3. 2 Evolution of cumulativeness of technological advance (Dibiago & Nasiriyar, 2008)

The ranges of the high or low effects of the cumulativity factor which will be figured out later in the next chapter were depending on the number of patents and patent growth in semiconductor technology space as shown in Figure 3.3 for each year mentioned in Table 3.2.

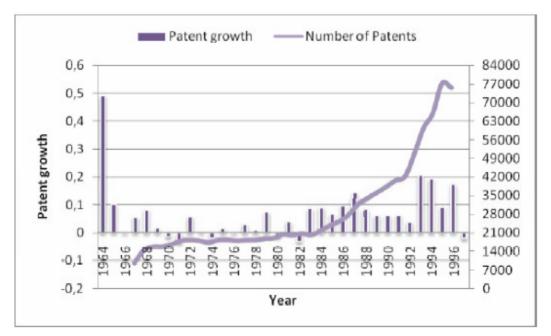


Figure 3. 3 The evolution of the patents and patent growth in semiconductor technology space (Dibiago & Nasiriya,2008)

## 3.2.4. Sources of Knowledge

Sources of knowledge are the fourth factor affecting the valuation process of the knowledge. Referring to intellectual capital Stewart's definition mentioned in (Malhotra, 2003): "the intellectual material – knowledge, information, intellectual property, experience – that can be put to use create wealth".

According to the intellectual capital, there are three sources of knowledge assets: External Capital, Human Capital and Structural Capital (O'Hara and Shadbolt, 2001). A questionnaire obtained by a research team in Amsterdam 1999 from four companies: Institution of Higher Education, High-Tech Firm, Petroleum Exploration & Production Firm and Energy Delivery, has resulted in the shown below chart in Figure 3.4 for indicating the usefulness of each (Human, Structural and External (Customer)) capital in each of the four samples of companies(Miller et al., 1999).

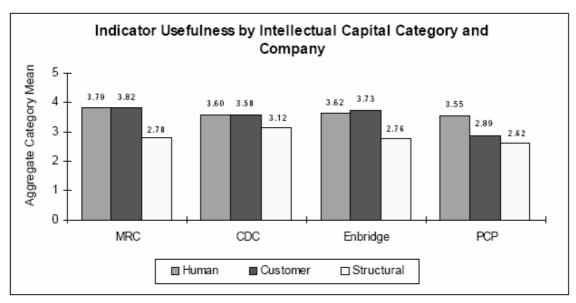


Figure 3. 4 The intellectual capital types effect on four sample firms (Miller et al.,1999)

## 3.2.5. Context of Knowledge

Knowledge's context refers to circumstances or events that form the environment within which something exists or takes place. Because of this relation between knowledge and these circumstances, they have their effects on improving and valuating knowledge (Young and Letch, 2003).

By referring to a questionnaire had been adopted for the purposes of an organization's information management practices, information behavior and values, and information uses. Table 3.2 shows the questionnaire items for this survey.

Construct	Item	Statement
	KME1	My organization has a culture intended to promote knowledge and information sharing.
KME	KME2	Knowledge and information in my organization is available and organized to make it easy to find what I need.
NIVIE	KME3	Information about good work practices, lessons learned, and knowledgeable persons is easy to find in my organization.
	KME4	My organization makes use of information technology to facilitate knowledge and information sharing.
	OIB1	The people I work with regularly share information on errors or failures openly.
OIB	OIB2	The people I work with regularly use information on failures or errors to address problems constructively.
	OIB3 (Reversed)	Among the people I work with regularly, it is normal for individuals to keep information to themselves.
	PIB1	I often exchange information with the people with whom I work regularly.
PIB	PIB2	I often exchange information with people outside of my regular work unit but within my organization.
	PIB3	I often exchange information with citizens, customers, or clients outside my organization.
	PIB4	I often exchange information with partner organizations.

Table 3. 3 The questionnaire items pertaining to the survey (Detlor et al., 2006)

For the purposes of studying the effects of the context of knowledge on knowledge valuation process, KME (Knowledge Management Environment) items are the only to be focused on as listed in the table above (KME1,KME2,KME3 and KME4),and analyzing their impact after observing the values of convergent validities of the four previously mentioned items with both OIB (Organizational Information Behavior) and PIB (Personal Information Behavior) as viewed in Table 3.4 below.

	KME	PIB	OIB	Age	Sex	Job Category	Sex*Job Category	Years in Org.
KME1	0.797	0.161	0.341	-0.052	-0.062	-0.103	-0.041	-0.045
KME2	0.832	0.151	0.267	0.000	-0.121	-0.089	-0.097	-0.080
KME3	0.838	0.175	0.398	0.008	-0.128	-0.121	-0.067	-0.075
KME4	0.806	0.209	0.297	-0.002	-0.070	-0.067	-0.047	-0.020
PIB1	0.259	0.680	0.342	0.020	0.027	0.096	0.058	-0.027
PIB2	0.089	0.720	0.116	0.194	0.085	0.139	0.082	0.152
PIB3	0.121	0.782	0.055	0.227	0.213	0.310	0.282	0.129
PIB4	0.142	0.758	0.132	0.163	0.082	0.129	0.079	0.176
OIB1	0.323	0.151	0.871	-0.018	-0.132	-0.125	-0.132	-0.166
OIB2	0.363	0.241	0.879	-0.050	-0.032	-0.074	-0.049	-0.118
OIB3	0.343	0.165	0.757	-0.087	-0.017	0.101	0.018	-0.057
Age	-0.014	0.209	-0.061	1.000	0.009	0.005	-0.002	0.534
Sex	-0.119	0.150	-0.072	0.009	1.000	0.478	0.884	0.152
JobCat	-0.120	0.243	-0.041	0.005	0.478	1.000	0.602	0.235
SexJobCat	-0.077	0.187	-0.065	-0.002	0.884	0.602	1.000	0.202
YearsInOrg	-0.068	0.140	-0.135	0.533	0.152	0.235	0.202	1.000

Table 3. 4 Matrix of loadings and cross-loadings of the survey's items (Detlor et al.,2006)

## 3.2.6. Six Challenges of Knowledge Management

Knowledge Management has the following challenges:

- Knowledge acquisition
- Knowledge modelling
- Knowledge retrieval
- Knowledge reuse
- Knowledge publishing
- Knowledge maintenance

(O'Hara and Shadbolt, 2001).

The effect of knowledge acquisition challenge will be used in terms of Knowledge Management effect on knowledge valuation process. A survey achieving this purpose had been undertaken consisting of 930 Greek companies;

this study identified and discussed the critical success factors or enablers that determine the KM effectiveness within organizations, which in turn influence the total performance of the firm (Theriou et al., 2010).

Table 3.5 shows the construct validity and variance extracted for each of the factors listed to obtaining the survey's purposes mentioned above. In this table the last item which is Knowledge Management effectiveness was adopted for this thesis for valuing knowledge.

The calculation of the construct reliability of each factor leads the researcher to conclude whether or not the various items of a construct as a set are reliable, in the sense of producing similar construct metrics every time is used by different researchers for similar contexts(Theriou et al.,2010).

Items	$\lambda_{ii}$	$\epsilon_{ii}$	$\lambda_{ii}^{2}$
Leadership			
LED1 -	0.79	0.38	0.6241
LED2	0.49	0.76	0.2401
LED3	0.53	0.72	0.2809
LED4	0.73	0.47	0.5329
	2.54	2.33	1.6780
0.73 construct reliability			
0.49 variance extracted			
Culture			
CUL1	0.32	0.90	0.1024
CUL2	0.83	0.31	0.6889
CUL3	0.86	0.25	0.7396
CUL4	0.89	0.20	0.7921
CUL5	0.80	0.36	0.6400
	3.70	2.02	2.9630
0.87 construct reliability			
0.59 variance extracted			
Strategy			
STR1	0.87	0.24	0.7569
STR2	0.92	0.16	0.8464
STR3	0.85	0.27	0.7225
3113	2.64	0.67	2.3258
0.91 construct reliability	2.04	0.07	2.3236
0.78 variance extracted			
Technology	0.75		0.5505
TEC1	0.75	0.44	0.5625
TEC2	0.45	0.80	0.2025
TEC3	0.57	0.68	0.3249
TEC4	0.69	0.53	0.4761
TEC5	0.71	0.50	0.5041
TEC6	0.51	0.73	0.2601
	3.68	3.68	2.3302
0.79 construct reliability			
0.39 variance extracted			
People			
PEP1	0.41	0.83	0.1681
PEP2	0.41	0.83	0.1681
PEP3	0.84	0.30	0.7056
PEP4	0.93	0.13	0.8649
ILIT	2.59	2.09	1.9067
0.76 construct reliability	2.39	2.09	1.5007
0.48 variance extracted			
KM effectiveness			
	0.00	0.04	0.0004
KM1	0.98	0.04	0.9604
KM2	0.89	0.21	0.7921
KM3	0.80	0.36	0.6400
KM4	0.68	0.54	0.4624
KM5	0.69	0.52	0.4721
KM6	0.85	0.27	0.7225
	4.89	1.94	4.0535
0.92 construct reliability			
0.68 variance extracted			
Firm performance			
KMF1	0.85	0.27	0.7225
KMF2	0.93	0.13	0.8649
	1.78	0.4	1.5874
0.89 construct reliability	1.70	0.4	1.50/4
0.80 variance extracted			
0.00 variance extracted			

Table 3. 5 Construct reliability and variance extracted for survey's items (Theriou et al.,2010)

# **Chapter Four**

# **Knowledge Acquisition Framework Design and**

# **Implementation**

#### 4.1. Introduction

There are many factors expressed for the purpose of knowledge valuation ontology. As per (O'Hara & Shadbolt, 2001), six factors will be used to valuing knowledge. These factors will be utilized an Adaptive Neural Fuzzy Inference System (ANFIS) that covers the second and the third stages of the proposed framewrok.

The first stage will include context-aware retrieval information algorithm. All the above three stages will be used for valuing knowledge.

In this chapter an illustration of the model structure is presented. It contains a full description of all the input membership functions and the rules for these inputs will be listed and more details about how these rules have been structured will be presented in the next chapter.

#### 4.2. Framework Architecture

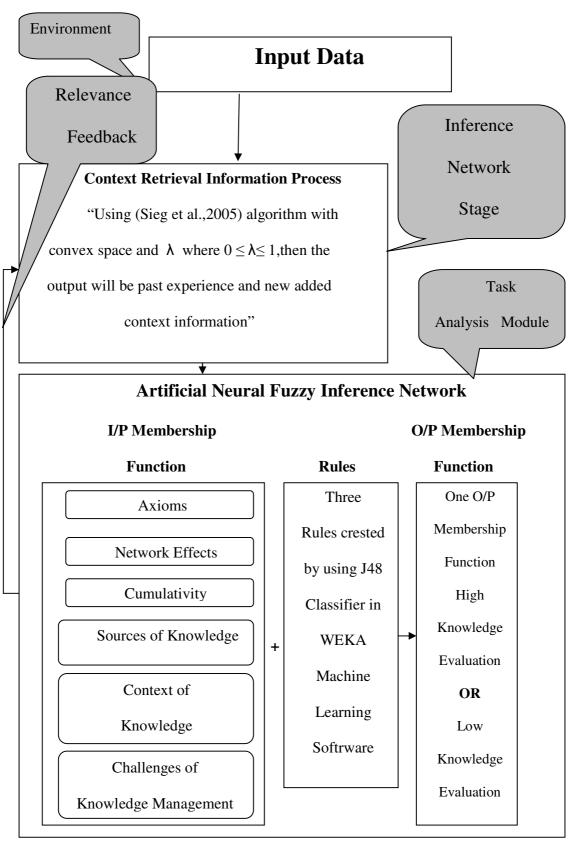


Table 4. 1 The main process of the proposed frame work

#### **4.2.1.** Inference Network

In this stage a context retrieval information algorithm will be used that integrates the essential elements of user's information context. This algorithm has been submitted in this stage other than presented methods in previous works due to the need for equations and numbers to be stated clearly in this stage and then to be fed later to the second stage, but other works were mechanisms and theories without numbers to be used in this work efficiently. In this algorithm the user's context is represented taking into account the user's short-term and long-term profiles, as well as relevant concepts from a pre-existing ontology (Sieg et al.,2005). In their framework, the user's "context" is captured via nodes in a concept lattice induced from the original ontology and is updated incrementally based on user's interactions with the concepts in the ontology. Their experimental results showed that utilizing the user context improves the effectiveness of the search queries, especially in the typical case of Web users who tend to use very short queries. A term-vector based representation is used for concepts. To generate a term-vector representation, the content of all the associated relations with the concept are combined to yield a single term-vector. To convert the problem space from ordinary space to convex space  $\lambda$  will be used here, in addition to generalize the normal spaces. A weighted term-vector its symbol is n<sub>i</sub> for each concept i. Each concept contains a collection of relations R<sub>i</sub>, and a set of sub-concepts S<sub>i</sub> .To compute n<sub>i</sub>, first we compute a term-vector  $n_R$  for each element  $r \in R_i$ . Then  $n_i$  is computed as the following:

$$n_i = (1 - \lambda) \sum_{r \in R} n_r + \sum_{s \in S} n_s$$

where each  $n_s$  is a term-vector for each sub-concept  $s \in S$  and  $0 \le \lambda \le 1$ .

Let  $n_1 = \{w^1_1, w^1_2, w^1_3 \dots w^1_k\}$  and  $n_2 = \{w^2_1, w^2_2, w^2_3 \dots w^2_k\}$  be two nodes in the problem space. Then  $n_1 \le n_2$  if and only if  $\forall j \ wj1 \le wj2$ , where wji is the weight of a term j in the term vector for ni. The operations on these nodes are summarized in the selection and deselection of these nodes, depending on the user query or on the stored profile for the user. Selection and deselection operations are translated to vector operations min and max operation, respectively as per the following:

$$\begin{aligned} & \min(n_1, n_2) = \{ \min(w^1_1, w^2_1), \dots, \min(w^1_k, w^1_k) \} \text{ and} \\ & \max(n_1, n_2) = \{ \min(w^1_1, w^2_1), \dots, \max(w^1_k, w^1_k) \} \end{aligned}$$

When  $\lambda=1$  then the sub-concept will be the main content for the term vector  $n_i$  and when  $\lambda=0$  both relations and sub-concepts will be included in each  $n_i$ . Thus the user context is represented as a pair of elements:

ci = {P,N},where P is a term-vector of positive evidence (min operation) : P =  $\min(n1,n2)$ ,and N is a term-vector of negative evidence (max operation) : N =  $\max(n1,n2)$ . The min and max operations could be extended to more logical operations intersection and union operations, respectively. Thus, the positive evidence will be represented as  $P = n1 \cap n2 \cap n3 \cap .... \cap nk$  and the negative evidence will be represented as  $N = n1 \cup n2 \cup n3 \cup .... \cup nk$ . Each time the user interacts in the specific domain seeking more information, the user's short term interest as a context ci, which is a pair of positive and negative evidence. In order to represent the user's long-term context, i.e. the user profile as a set of

contexts:  $pr = \{c0, c1, c2, ..., cn\}$ . Depending on user behavior, a specific context in the user profile can be updated or a new context can be added.

Via this algorithm, solving the faced problems has been transferred from the Case Based Reasoning approach to Trace Based reasoning approach, which in turn achieving one of the aims of this research.

As a conclusion of this stage, the user's context information represented by user's short term and long term profiles, in addition to the past pre-existing ontology, are fed as inputs for the next stage of the model which is Task Analysis Module which is illustrated below.

## 4.2.2. Task Analysis Module

In this stage, there will be an implementation of an ANFIS model (Artificial Neural Fuzzy Inference System) via using linguistic variables represented by member function (mf) indicating the degree and the status of each factor on the process of valuing knowledge.

The six factors and their member functions are listed as follows:

## 4.2.2.1. Membership Functions of Input factors

#### **Axioms**

The data set was adopted from an evaluation used in NanOn ontology which is specified in the Web Ontology Language OWL DL (OWL Working group). This ontology comprises 2,289 logical axioms.

In order to figure out the effects of Axioms factor in knowledge valuation, two linguistic variables are created to implement the impact of axioms in valuing knowledge namely Axioms high impact (HighI) and axioms low impact (LowI).AHighI values range from 954 to 20,739 reasoner calls and ALowI values range from 1,438 to 212,041 reasoner calls, please refer to Table 3.1. The Generalized Bell member functions for AHighI and ALowI are shown in Figure 4.2.

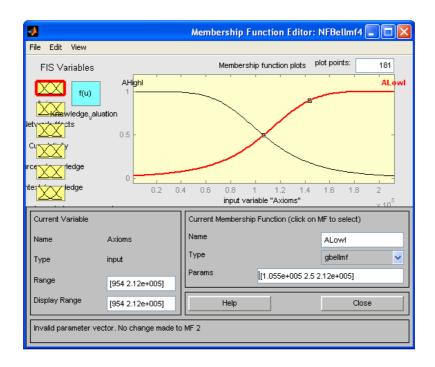


Figure 4. 2 The input membership function for Axioms

#### **Network Effects**

Patents which is strongly related to (R&D) across firms, please refer to Figure 3.1 are considered here as the inputs theme for measuring the value of knowledge and referring to (R&D) activities within a firm by network of relations. The data set adopted from the World Intellectual Property Indicators 2010, please see Figure 3.2.

The Network Effects factor has been converted into numbers, so two linguistic variables are created to implement the impact of network effects

namely network high effect (NHighE) and network low effect (NLowE). NHighE values range from 1.6 to 11 and NLowE values range from 0.3 to 11 The Generalized Bell member functions for NHighE and NLowE are shown in Figure 4.3.

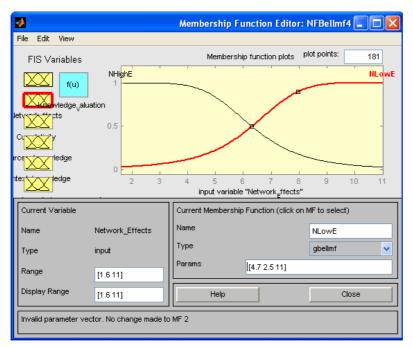


Figure 4. 3 The input membership function for Network Effects

#### **Cumulativity**

In the semiconductor industry, an analytical study revealing the cumulative nature of innovation development, explained how much new innovative outputs (patents) are based on existing technological knowledge. The data set adopted here from the correlation coefficients of cumulativeness, please refer to Table 3.2. This is the main contribution here by converting the intangible cumulativity effects into numbers.

Two linguistic variables are created to implement the impact of cumulativity namely cumulativeness high effect (CHighE) and cumulativeness

low effect (CLowE). CHighE values range from 0.5288 to 0.9822 and CLowE values range from 0.8716 to 0.9074, please see Figure 3.3. The Generalized Bell member functions for CHighE and CLowE are shown in Figure 4.4.

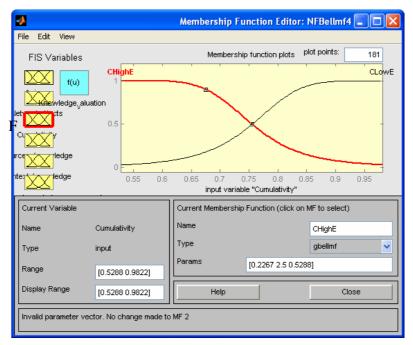


Figure 4. 4 The input membership function for Cumulativity

## **Sources of Knowledge**

According to the intellectual capital, there are three sources of knowledge assets: External Capital, Human Capital and Structural Capital. The data set used here was from a questionnaire obtained by a research team in Amsterdam 1999 from four companies (O'Hara and Shadbolt,2001). These intangible sources have been converted to numbers as illustrated below.

Two linguistic variables are created here to implement the impact of the above three mentioned sources namely external capital effect (ECapitalE), human capital effect (HCapitalE) and structural capital effect (SCapitalE). ECapitalE values range from 2.89 to 3.82 , HCapitalE values range from 3.55 to

3.79 and SCapitalE values range from 2.62 to 3.12, please refer to Figure 3.4. The Generalized Bell member functions for ECapitalE, HCapitalE and SCapitalE a're shown in Figure 4.5.

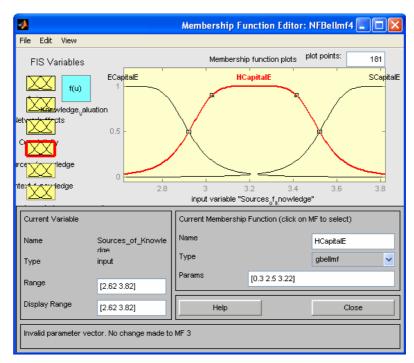


Figure 4. 5 The input membership function for Sources of Knowledge

#### **Context of Knowledge**

The data set to illustrate the context of knowledge effects, have been recorded from a questionnaire had been done for many kinds of Knowledge Management Environments statuses, please refer to Table 3.3, these items affect both Organizational Information Behaviors (OIB) and Personal Information Behaviors (PIB). Accordingly, these data have been used to view the context of knowledge effect practically by using numbers for this intangible factor.

Four linguistic variables are created to implement knowledge's context effect on the valuation process of knowledge namely knowledge management 1 (KM1),knowledge management 2 (KM2),knowledge management 3 (KM3) and

knowledge management 4 (KM4).KM1 values range from 0.161 to 0.341,KM2 values range from 0.151 to 0.267,KM3 values range from 0.175 to 0.398 and KM4 values range from 0.209 to 0.297,please refer to Table 3.4. The Generalized Bell member functions for KM1, KM2, KM3 and KM4 are shown in Figure 4.6.

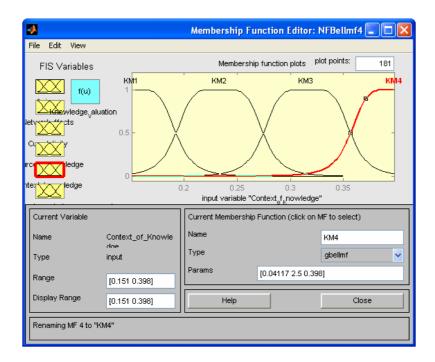


Figure 4. 6 The input membership function for Context of Knowledge

## Six Challenges of Knowledge Management

The data set used here for measuring the knowledge management's effect on knowledge valuation process was taken from an empirical research of the Greek medium and large firms (Theriou et al., 2010).

As per the previous five factors of knowledge valuation, the main contribution is converting the intangible challenges of knowledge management effects into numbers, two linguistic variables are created to implement the

impact of knowledge management namely construct validity of knowledge management (CVKM), and variance extracted of knowledge management (VEKM). CVKM values range from 0.68 to 0.98 , VEKM values range from 0.04 to 0.54, please refer to Table 3.5. The Generalized Bell member functions for CVKM and VEKM are shown in Figure 4.7.

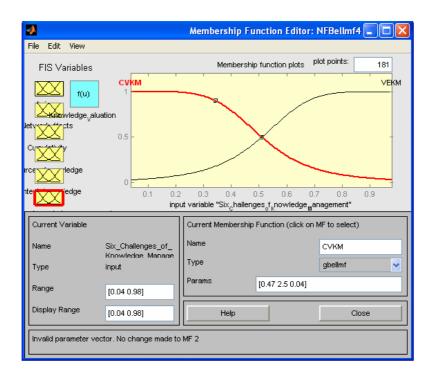


Figure 4. 7 The input membership function for Six Challenges of Knowledge Management

The following Table 4.1 summarizes the input variables and their corresponding fuzzy linguistic variables and ranges.

Table 4. 1 Input factors and their Linguistic Variables and Ranges

Input Factor	Ling.	Ling.	Ling.	Ling.
	Var 1	Var 2	Var3	Var 4
Axioms	AHighI	ALowI		
AXIOIIIS	954 - 20,739	1,438 - 212,041		
Netwrk	NHighE	NLowE		
Effexts	1.6 – 11	0.3 - 11		
	CHighE	CLowE		
Cumulativity	0.5288 -	0.8716 - 0.9074		
	0.9822			
Sources of	ECapital E	HCapitalE	SCapitaE	
Knowledge	2.89 – 3.82	3.55 – 3.79	2.62 – 3.12	
Context of	KM1	KM2	KM3	KM4
Knowledge	0.161 - 0.341	0.151 - 0.267	0.175 - 0.398	0.209 - 0.297
Six	CVKM	VEKM		
Challenges of	0.68 - 0.98	0.04 - 0.54		
Knowledge				
Management				

## **4.2.2.2.** Rules and Output Membership Functions

The output from this model will be two different outputs describing the status of valuing the knowledge process, that will be either good knowledge valuation or poor knowledge valuation affected by the factors listed in the previous section.

The relationship between the input and output variables is done by creating rules, the J48 classifier in WEKA is used for this purpose. WEKA is a machine learning software written in Java, contains a collection of visualization tools and algorithms for data analysis and predicting modeling, with an easy to use graphical user interface. The rules obtained are listed below:

- 1. If (Sources of Knowledge is Low) then (Knowledge Valuation) is Low.
- 2. If (Sources of Knowledge is High) and (Six Challenges of Knowledge Management is High) then (Knowledge Valuation) is High.
- 3. If (Sources of Knowledge is High) and (Six Challenges of Knowledge Management is Low) then (Knowledge Valuation) is High.

## 4.2.3. Relevance Feedback

The last stage is the relevance feedback process which has already been chosen when using the ANFIS editor when we want to train our FIS model. Via this process, the training of the model will be enhanced and the error measure accordingly will be adjusted for better measures and thereafter there will be a close measure for the desired output of the model being processed.

# **Chapter Five**

## **Experimental Study**

#### 5.1. Overview

In this chapter, an overview for the complete neuro-fuzzy models with 6 input factors and 1 output valuation, be presented. In these models, three input membership functions will be used, namely generalized bell (gbellmf), and gaussian (gaussmf) and gaussian2(gauss2mf). The models will use two variations of the Sugeno output, namely the constant and the linear output functions. One training set will be used to test the models.

## 5.2. Neuro-fuzzy Models

The fuzzy model for all the input factors and the output valuation is shown in Figure 5.1. The models are built using MATLAB ANFIS editor with the input member functions of Gaussian Bell.

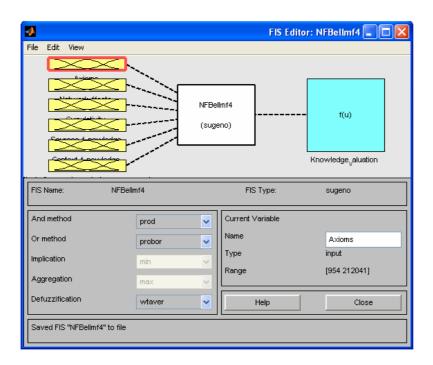


Figure 5. 1 Fuzzy model of the 6 input mfs

MATLAB ANFIS editor supports only the Sugeno type, and the Fuzzy Inference System (FIS) supports two types of output functions, the constant, and the linear function. The rule in Sugeno fuzzy model has the form

#### If (input 1 = x) and (input 2 = y) then output z = ax + by + c.

For the constant Sugeno model, the output level z is constant c, where a = b = 0. The output level  $z_i$  of each rule is weighted by firing strength  $w_i$  of the rule.

Six distinct neuro-fuzzy models are used to demonstrate the correlation and delectability of knowledge valuation using the 6 factors presented earlier. The classification of the models is given in Table 5.1. Each model is

characterized by the type of the input/output membership function and constant or linear output type.

Table 5. 1 Model Specifications

Model Name	Input member function	Output member function
Generalized Bell	Gbellmf	Constant
Generalized Bell	Gbellmf	Linear
Gaussian	Gaussmf	Constant
Gaussian	Gaussmf	Linear
Gaussian2	Gauss2mf	Constant
Gaussian2	Gauss2mf	Linear

For each of the models shown in Table 5.1, the neuro-fuzzy structure was built. The structure of the J48 rules based neuro-fuzzy model is shown in Figure 5.2.

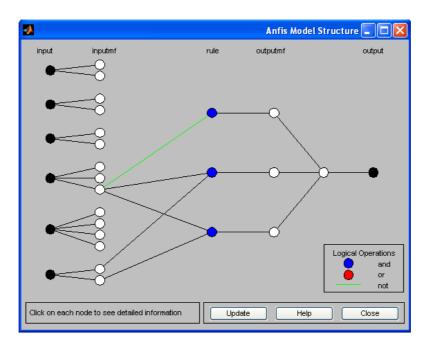


Figure 5. 2 The Structure of the j48 rules based neuro/fuzzy model

A particular architecture of neuro-fuzzy systems is that of the Adaptive Neuro Fuzzy Inference System (ANFIS) introduced by (Shing and Jang 1993). Figure 5.3 shows the fuzzy inference system used in ANFIS and it is composed of four functional blocks.

The knowledge base block contains database and rule base. Database defines the membership functions and rule base consists of fuzzy if-then rules. A fuzzification interface which transforms the crisp inputs into degrees of match with linguistic values; a defuzzification interface which transforms the fuzzy results of the inference into a crisp output. The fuzzy rules used in ANFIS are of Takagi-Sugeno type. This type of fuzzy rule has fuzzy sets involved only in the premise part; the consequent part is described by a non-fuzzy equation of the input variables.

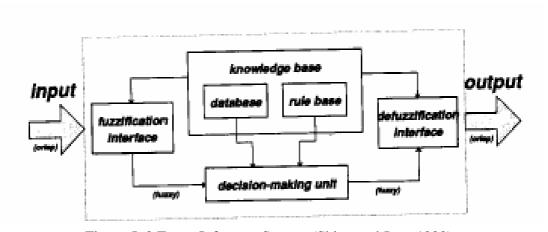


Figure 5. 3 Fuzzy Inference System (Shing and Jang, 1993)

Each of the models is characterized by an input membership function (Generalized Bell, Gaussian or Gaussian2) and an output membership function. Initial parameters have to be chosen; for each input membership function.

For the above mentioned membership functions, the generalized bell function depends on three parameters a, b, and c as given by

$$f(x,a,b,c) = 1/1 + |x-c/a|^{2b}$$

For Gaussian membership function, the Gaussian curve is given by

$$f(x) = \exp\{-0.5(x-c)^2/\sigma^2\}$$

where c is the mean and  $\sigma$  is the variance. The output values are selected in a manner similar to the method described in the previous chapter. The Gaussian2 MF block implements a membership function based on a combination of two Gaussians. The two Gaussian functions are given by

$$f_k(x) = \exp\{-0.5(x-c_k)^2/|\sigma_k|^2\}$$

## 5.3. Training

The purpose of the training is to adjust the model parameters, particularly the input membership function parameters, and the corresponding output values. The adjustment and tuning depend on the accuracy of the training data, as will be shown later.

Training needs two kinds of arrays, the first is the training array and the other one is the testing array. A training array is a two dimensional array  $[m \times n]$ , where (m) is the number of rows containing input values, and (n) is the number of input factors plus one for the output column.; in our model, n=7 since there are 6 distinct input variables and one output variable. Each row of the array contains some of the possible values for each input corresponding to the first n-1 columns representing the 6 variables, and the last column holds the desired output values. The testing array holds the data in the same way as the training

array, but the data in this array is more accurate than the data of the training array.

The possible combinations for 6 inputs variables and 2 output values. Each input factor has on the average two linguistic variables, thus making the total combinations =  $2^6$ , which equal to 64. The training array used is shown in Table 5.2, where m = 64 and n= 7.

Table 5. 2 Training Input Array

954	0.3	0.5308	2.62	0.297	0.98	1
955	0.4	0.5309	2.63	0.298	0.05	0
956	0.5	0.531	2.64	0.299	0.09	0
957	0.6	0.5311	2.65	0.3	0.54	1
958	0.7	0.5312	3.22	0.301	0.6	0
959	0.8	0.5313	2.66	0.302	0.81	1
960	0.9	0.5314	3.52	0.303	0.55	0
961	1	0.5315	2.67	0.304	0.68	1
962	1.1	0.5316	3.24	0.305	0.77	0
963	1.2	0.5317	2.68	0.306	0.89	0
964	1.3	0.5318	2.69	0.307	0.4	0
965	1.4	0.5319	3.33	0.308	0.11	0
966	1.5	0.532	2.71	0.309	0.56	1
967	1.6	0.5321	2.72	0.31	0.43	0
968	0.35	0.5322	3.82	0.311	0.19	0
969	0.36	0.5323	2.73	0.312	0.38	0
1016	0.37	0.5324	2.74	0.313	0.93	1
1017	0.38	0.5325	3.55	0.151	0.66	0
1018	0.39	0.5326	3.79	0.152	0.21	0
1019	0.4	0.5327	2.75	0.153	0.217	0
1020	0.41	0.5328	2.76	0.154	0.3	0
1021	0.42	0.9161	2.77	0.155	0.83	1
1022	0.43	0.9162	2.78	0.156	0.88	1
1023	0.44	0.9163	2.78	0.157	0.73	1
1024	0.45	0.9164	2.79	0.158	0.97	1
1025	0.46	0.9165	2.81	0.159	0.93	1
1026	0.47	0.9166	2.82	0.16	0.33	0
1027	0.48	0.9167	2.83	0.161	0.4	0
1028	0.49	0.9168	2.84	0.311	0.45	1
1029	0.5	0.9169	2.85	0.312	0.23	0
1030	0.51	0.917	2.86	0.313	0.16	0
1031	0.52	0.9171	3.14	0.314	0.11	0

20746	0.53	0.9172	3.13	0.315	0.09	0
20747	0.54	0.9173	3.8	0.316	0.1	0
20748	0.55	0.9174	3.17	0.317	0.49	0
20749	0.56	0.9175	3.33	0.318	0.54	0
20750	0.57	0.9176	3.82	0.319	0.43	0
20751	0.58	0.9177	3.32	0.32	0.32	0
20752	0.59	0.9178	3.36	0.321	0.36	0
20753	0.6	0.9179	3.46	0.322	0.45	0
20754	0.61	0.918	3.5	0.323	0.38	0
20755	0.62	0.9181	3.77	0.324	0.28	0
20756	0.63	0.9182	3.81	0.325	0.21	0
20757	0.64	0.9183	2.8	0.326	0.27	0
20758	0.65	0.9188	3.5	0.327	0.24	0
20759	0.66	0.9189	3	0.328	0.13	0
20760	0.67	0.919	3.19	0.329	0.52	0
20761	0.68	0.9191	3.8	0.33	0.07	0
20762	0.69	0.9192	2.81	0.331	0.44	1
30009	0.7	0.9193	2.82	0.332	0.42	0
30010	0.71	0.9194	2.83	0.333	0.51	1
30011	0.72	0.9195	2.84	0.334	0.5	1
30012	0.73	0.9196	2.67	0.335	0.48	1
30013	0.74	0.9197	2.68	0.336	0.06	0
30014	0.75	0.9198	2.69	0.337	0.53	1
30015	0.76	0.9199	2.7	0.338	0.39	0
30016	0.77	0.92	2.71	0.339	0.19	0
30017	0.78	0.9201	2.72	0.34	0.04	0
30018	0.79	0.9202	2.73	0.341	0.46	1
30019	0.8	0.9203	2.74	0.342	0.34	0
30020	0.81	0.9204	2.76	0.343	0.43	0
200000	0.82	0.9205	3.33	0.344	0.85	0
100000	0.83	0.9206	2.85	0.345	0.66	1
50000	0.84	0.9207	2.86	0.346	0.82	1

The above listed data in the training input array is selected randomly, but has been constructed according to the rules obtained via J48 classifier in WEKA.

The testing array will have more accurate data that will be chosen carefully far away from intersection points between the values of the linguistic variables for each factor. For example values for ECapitalE will be adopted which ranges from (2.89) to (3.82) without intersecting the value ranges for HCapitalE for valuing knowledge by the factor sources of knowledge.

#### 5.4. Results

The membership functions in the ANFIS system have 2 stages. In the start the membership functions are at their default shapes. This default shape changes when the ranges are assigned to them. After performing the training the membership functions have a changed shape. The reason for this change is that when an ANFIS undergoes from training process it tunes the membership functions according to the corresponding training data and rules. So membership functions of a trained ANFIS have a different shape as compared to an untrained ANFIS.

Another important thing to remember is that the shapes of only those membership functions are changed which are included in any rules. Thus the following will illustrate these changes of the membership functions due to training process.

#### **5.4.1. J48 Classification Results**

#### **Using Cross Validation (10 folds)**

Here is the confusion matrix of J48 classifier. 60% data was used for training and 40% for testing and the data is selected Randomly Here its show only the 40% of the testing data.

```
=== Confusion Matrix ===
a b <-- classified as
5 4 | a = High
3 7 | b = Low</pre>
```

Figure 5. 4 The confusion matrix using Cross Validation from WEKA program(snapshot)

The confusion matrix shows that 12 instances were correctly classified out of 19 and 7 instances were incorrectly classified. In other words, here 5 High values and 7 Low values are correctly classified and 4 High values and 3 Low values are incorrectly classified. 7 instances are miss classify because the classification is done by applying rules so there is may be an article which is according to the rules in class High but in actual it is in class Zero. So according to our system it is a miss classified article because our system done classification according to the rules. The performance of J48 classifier is 63 %.

### **Using Percentage Split**

The classification was also done by using the percentage split.

```
=== Confusion Matrix ===

a b <-- classified as

0 0 | a = High
2 4 | b = Low
```

Figure 5. 5 The confusion matrix using Percentage Split from WEKA program (snapshot)

The level of performance achieved by using percentage split is a little higher than the cross validation 66.6 % The results shows that 4 instances were correctly classified and 2 instances were wrongly classified.

## J48 Classification Tree

The decision tree shown below is obtained by applying the J48 classifier on the input data. The inputs having the strong influence on the result are included in this tree. In other word it could be said that these are the inputs which influence the classification results.

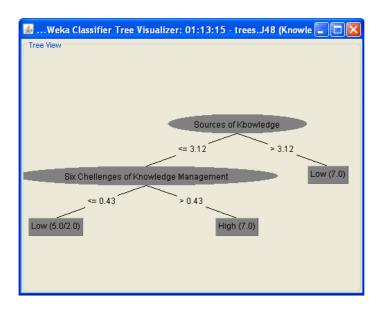


Figure 5. 6 The Structure of the j48 classification tree

## **5.4.2.** The impact of training array

As illustrated at the begging if this chapter, three membership functions namely generalized bell (gbellmf), Gaussian (gaussmf)., and gaussian2 (gauss2mf) will be used. The models will use two variations of the Sugeno output, namely the constant and the linear output functions.

Firstly, by using the Generalized Bell membership functions with the constant output, the results are shown in table 5.3 for the parameters a, b and c before and after the training for Axioms.

Table 5. 3 Axioms Generalized Bell/Constant Model

	a		b		С	
	u .		_		•	
sma	Before	After	Before	After	Before	After
Axioms	Training	Training	Training	Training	Training	Training
AHighI	1.055e+005	1.06e+005	2.5	2.5	954	1510
ALowI	1.055e+005	1.06e+005	2.5	2.5	2.12e+005	2.12e+005

Note that the a values increased for ALowI and AHighI, while b values remain the same for the both AHighI and ALowI and c increased for AHighI. Figure 5.7 shows the Axioms input membership functions after applying the training, which is an adjustment for its membership function before training as seen in Figure 4.2.

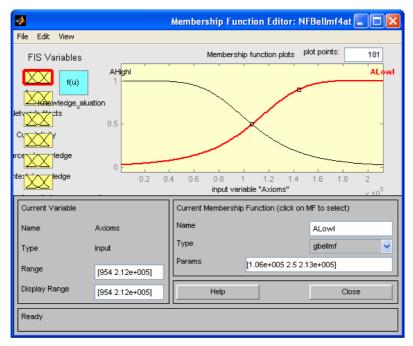


Figure 5. 7 Axioms Generalized Bell/Constant model after training

The following three tables 5.4, 5.5 and 5.6 also show the changes and adjustments in the parameters values after applying training to the six factors, but they are only for Network Effects and Context of Knowledge in addition to Axioms shown above, as per the other remaining 3 factors the values for the parameter of their membership functions resulted without any change before and after training.

Table 5. 4 Network Effects Generalized Bell/Constant Model

		a		b		С	
Netwrk	Effects	Before Training	After Training	Before Training	After Training	Before Training	After Training
NHighE		4.7	4.7	2.5	2.5	1.6	1.575
NLowE		4.7	4.7	2.5	2.5	11	10.98

Note that the c values decreased for NHighE and NLowE, while a and b values remain the same for the both NHighE and NLowE. Figure 5.8 shows the Network Effects input membership functions after applying the training, which is an adjustment for its membership function before training as seen in Figure 4.3.

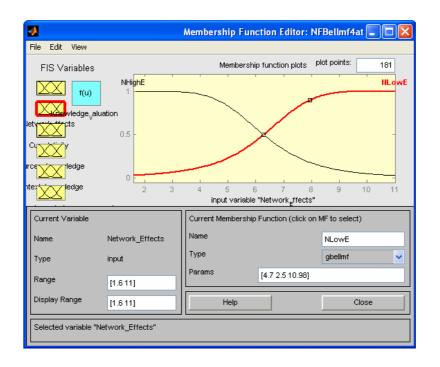


Figure 5. 8 Network Effects Generalized Bell /Constant model after training

Table 5. 5 Context of Knowledge Generalized Bell/Constant Model

		a		b		С	
t of	gps	Before	After	Before	After	Before	After
Context of	Knowledg	Training	Training	Training	Training	Training	Training
KM	-	0.04117	0.04117	2.5	2.5	0.151	0.1503
KM	7	0.04117	0.0412	2.5	2.5	0.2333	0.2307
KM	8	0.04117	0.0412	2.5	2.5	0.3157	0.315
KM	4	0.04117	0.04117	2.5	2.5	0.398	0.3967

Note that the c values decreased for KM1, KM2, KM3 and KM4, and the a values increased for both KM2 and KM3. Figure 5.9 shows the Context of Knowledge Generalized Bell/Constant input membership functions after applying the training, which is an adjustment for its membership function before

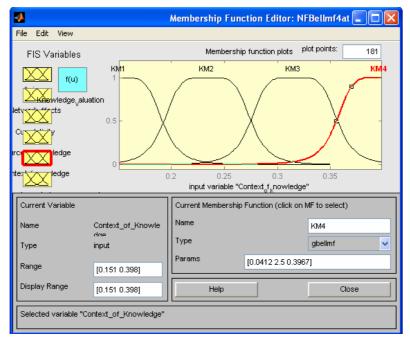


Figure 5. 9 Context of Knowledge Generalized Bell/Constant model after training

By using the Generalized Bell function with linear output, the results are shown in Table 5.6 for the parameters a, b and c before and after the training for Context of Knowledge factor because the other factors remained the same for each parameter of the Generalized Bell function.

Table 5. 6 Context of Knowledge Generalized Bell/Linear Model

	Table 3. 6 Context of Thio wreage Constanted Bell Elinear Model							
		a		b		С		
e								
Context of	Knowledg	Before	After	Before	After	Before	After	
x	le							
te	W(	Training	Training	Training	Training	Training	Training	
	n		Ü			Ü	Ü	
Ü	X							
		0.04115	0.04115	2.5	2.5	0.151	0.151	
_		0.04117	0.04117	2.5	2.5	0.151	0.151	
KM	1							
<u> </u>								
		0.04117	0.04117	2.5	2.5	0.2333	0.2333	
$\mathbf{z}$	->	0.01117	0.01117	2.3	2.3	0.2333	0.2333	
KM	2							
		0.04117	0.04117	2.5	2.5	0.3157	0.317	
KM	3							
<b>×</b>								
		0.04117	0.0412	2.5	2.5	0.398	0.3987	
7		0.0411/	0.0412	2.3	2.3	0.390	0.3907	
KM	4							

Note that the a values increased for KM4, and the c values increased for KM3, and KM4. Figure 5.10 shows the Context of Knowledge Generalized Bell/Linear input membership functions after applying the training.

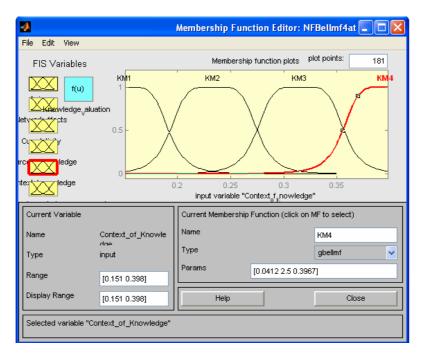


Figure 5. 10 Context of Knowledge Generalized Bell/Linear model after training

Secondly, by using the Gaussian membership functions with either constant or linear output, the parameters  $\sigma$  and C values remain without any changes for any of the six factors affecting the knowledge valuation, sample of the results are shown in Table (5-7) for the parameters  $\sigma$  and C before and after the training for Context of Knowledge.

Table 5. 7 Context of Knowledge Gaussian/Constant or Linear Model

		Tuble 5. 7 Contex	t of Knowledge Gau	351dii/ Collstailt of	Diffedi Wiodei	
<u>.</u>	<b></b>	Σ		С		
t 0	edg	Before	After	Before	After	
Context of	Knowledg	Training	Training	Training	Training	
KM	1	0.03497	0.03497	0.151	0.151	
KM	2	0.03497	0.03497	0.2333	0.2333	
KM	3	0.03497	0.03497	0.3157	0.3157	
KM	4	0.03497	0.03497	0.398	0.398	

Figure 5.11 shows the Context of Knowledge Gaussian Constant/Linear input membership functions before applying the training, followed by Figure 5.12 that shows the Context of Knowledge Gaussian Constant/Linear parameters after applying the training.

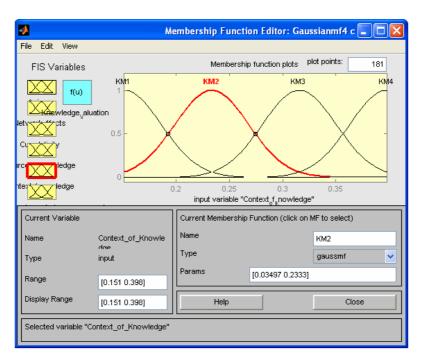


Figure 5. 11 Context of Knowledge Gaussian/Constant/Linear model before training

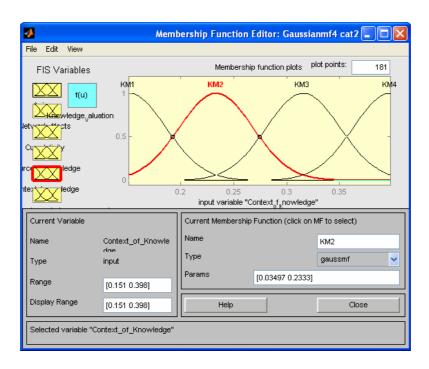


Figure 5. 12 Context of Knowledge Gaussian/Constant/Linear model after training

Finally, by using the Gaussian2 membership functions with either constant or linear output, the parameters  $\sigma_1$ , C1,  $\sigma_2$ , C2 values remain without any change for any of the six factors affecting the knowledge valuation, sample of the results are shown in Table 5.8 for the parameters  $\sigma_1$ , C1,  $\sigma_2$ ,  $C_2$  before and after the training for Context of Knowledge.

Table 5. 8 Context of Knowledge Gaussian 2 Constant/Linear Model

		$\sigma_1$		$\mathbf{C}_1$		$\sigma_2$		$C_2$	
Context of	Knowledg	Before Training	After Training	Before Training	After Training	Before Training	After Training	Before Training	After Training
KM	1	0.01243	0.01243	0.1245	0.1245	0.01243	0.01243	0.1775	0.1775
KM	2	0.01243	0.01243	0.2068	0.2068	0.01243	0.01243	0.2598	0.2598
KM	3	0.01243	0.01243	0.2892	0.2892	0.01243	0.01243	0.3422	0.3422
KM	4	0.01243	0.01243	0.3715	0.3715	0.01243	0.01243	04245	04245

Figure 5.13 shows the Context of Knowledge Gaussian2 Contant/Linear input membership functions before applying the training, followed by Figure 5.14 that shows the Context of Knowledge Gaussian2 Contant/Linear parameters after applying the training.

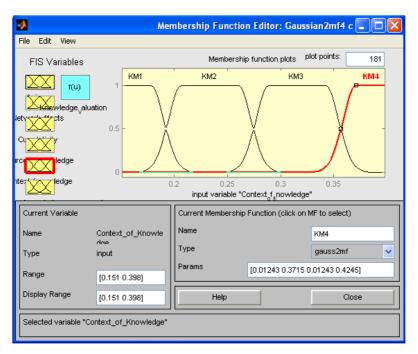


Figure 5. 13 Context of Knowledge Gaussian2/Constant/Linear model before training

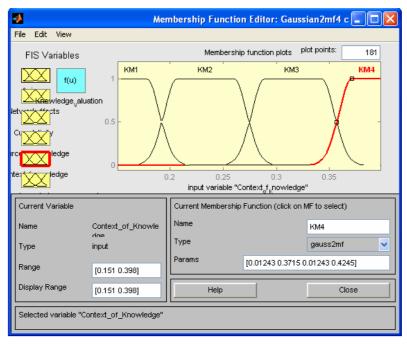


Figure 5. 14 Context of Knowledge Gaussian2/Constant/Linear model after training

It is clear that some of the parameters of the member functions had been changed some stayed the same. Note that no change implies that the initial choice of the parameters is in line with reality of the model. In other words, the established correlation between the input factors and the output is consistent. The adjustment of the parameters takes place whenever the initial choice of the parameter is not proper.

## **5.4.3. Performance**

In this section, the impact of training array on the performance of the models will be measured. In particular, there is an observation of the error rate of the models under same numbers of epochs which is 800 epochs. An epoch in the ANFIS is one full cycle staring from the application of input at layer 1 of the model, until the firing weight of the rule is adjusted. At the end of an epoch, the error, which is defined as the difference between the desired output and the computed output value, is measured. In this section, the models are trained by using testing array its inputs have been chosen carefully.

Table 5.9 shows the different values of the average testing errors for the three used membership functions (generalized bell, Gausian, Gaussian2).

Table 5. 9 Error values after testing the models for constant/Linear models

Model Name	Number of Epoch	The value of Error
Generalized Bell/Constant/Linear	800	0.46043
	<u> </u>	T
Gaussian/Constant/Linear	800	0.45846
Gaussian2/Constant /Linear	800	0.4641

It is clear from the table shown above that the best model to choose is Gaussian Constant or Linear for achieving the least error value in order to obtain the desired output within the framework.

# **Chapter Six**

## **Conclusion and Future Work**

#### 6.1. Conclusion

A Knowledge Acquisition framework for valuing knowledge using MATLAB has been presented in this study, this framework consists of three stages: the first stage has been deduced from a previously existing algorithm which has achieved the purpose of converting (CBR) to (TBR), the second stage has been built by using a neuro-fuzzy model by ANFIS editor (which is a digital data processing in the computer system and needs figures to work and have results) fed by 6 factors, each of these intangible factors have been translated into numbers in terms of membership functions by importing numbers and results of questionnaires and data surveys, reflecting these factors impact on the process of valuing knowledge, the fuzzy rules used in this stage has been deducted by using a machine learning software written in Java called WEKA, then the models were trained and the output parameters representing knowledge valuation have been adjusted using an array of training data. The performance of the model is measured in terms of error value obtained between the expected outputs; the Gaussian function is the most optimal in terms of trainability and producing low error values, and the choice of Sugeno either linear or constant output function is convenient accompanied to the Gaussian function in this proposed framework. this

framework has added to previous studies concerning solving problem methodologies, the importance of TBR adopted in its first stage, that means how important including context of information in any problem domain in order to solve the problem effectively, and after processing the framework, results show the context of knowledge as one of the six factors affecting the knowledge valuation process is the most important factor due to its high changes were more noticeable than others. The results presented in this study show that knowledge can be valued using a neuro-fuzzy model.

## **6.2. Future Work**

The following points could be implemented in the future in order to improve our work:

- 1. Submitting other available membership functions in ANFIS editor in order to have other possible may be less errors of the output such as Psigmoid, zmf, smf, dsigmf and others.
- 2. Working on the framework to obtaining accurate results by implementing an algorithm for adjusting the epochs for any used model through trying better training arrays for the framework to learn accordingly.

## References

Budzik J., Hammond K. & Brinbaum L. (2001), Information Access in Context, Knowledge-Based Systems Journal, 14, (1-2), 37-53.

Carlucci D., Marr B. & Schiuma G. (2004), The knowledge value chain: how intellectual capital impacts on business performance, *Int. J. Technology Management*, Vol. 27, Nos. 6/7.

Chaudhari O.,Khot P.,Deshmukh K. & Bawne N. (2012), ANFIS Based Model in Decision Making to Ootimize the Profit in Farm Cultivation, International Journal of Engineering Science and Technology (IJEST),Vol. 4 No. 02.

Chen Q., Wang J. & Tu Q. (2000), Connectionist Models for Tracing User-task Profiles, Challenges of Information Technology Management in the 21st Century, the proceedings of the Information Resources Management Association International Conference, <a href="http://www.irma-international.org/viewtitle/31538/">http://www.irma-international.org/viewtitle/31538/</a>.

Choucri N. (2007), The Politics of Knowledge Management, MIT, USA and Chair, Scientific Council MOST Programme, UNESCO.

Cordier, A. (2008), *Interactive and Opportunistic Knowledge Acquisition in CBR*, PhD Thesis, Université Claude Bernard Lyon 1.

Cordier A., Mascret B. & Mille A. (2009), Extending Case-Based Reasoning with Traces, *Association for the Advancement of Artificial Intelligence*, Stanford, California.

Delmastro F., Conti M. & Passarella A. (2010), Social-aware content sharing in opportunistic networks, PerAda Magazine.

Detlor B.,Ruhi U,Turel O,Bergeron P,Wei Choo C,Heaton L & Paquette S. (2006),The Effect of Knowledge Management Context on Knowledge Management Practices: an Empirical Investigation, *The Electronic Journal of Knowledge Management*, Volume 4 Issue 2, pp. 117-128, available online at www.ejkm.com

Dibiago L. & Nasiriyar M. (2008), Entrepreneurship and Innovation - rganizations Insstitutions, Systems and Regions, *Druid 25th Celebration Conference 2008*, Copenhagen, CBS, Denmark.

Dieng R. & Corby O. (Eds.) (2000), Knowledge Engineering and Knowledge Management Methods, Models, and Tools, 12th International Conference, EKAW 2000 Juan-les-Pins, France, October 2-6, 2000 Proceedings, pp. (80-96).

Duarte L., Kramer J. & Uchitel S. (2006), Model Extraction Using Context Information, In ACM/IEEE 9th International Conference on Model Driven Engineering Languages and Systems, Genoa, 380-394.

Fritz T. (2011), Developer-Centric Models: Easing Access to Relevant Information in a Software Development Environment (PhD thesis, The University of British Columbia, British Columbia, Canada).

Gruber T. (1993), A Translation Approach to portable Ontology Specifications, *Knowledge Acquisition*, 5,199-220.

Hall B.(2004), Patent Data as Indiactors, University of California at Berkeley, NBER, and IFS London, <a href="http://www.oecd.org/dataoecd/45/23/33835392.pdf">http://www.oecd.org/dataoecd/45/23/33835392.pdf</a> Viewed on 13-5-2012 at 1:07 am.

Hall W., Else S., Martin C. & Philp W. (2011), Time-Based Frameworks for Valuing Knowledge: Maintaining Strategic Knowledge, *Kororit Institute Working Paper* No. 1.

Hardian B., Indulska J. & Henricksen K. (2006), Balancing Autonomy and User Control in Context-Aware Sytems – a Survey, Pervasive Computing and Communications Workshops 2006 PerCom Workshops 2006 Fourth Annual IEEE International Conference.

Jeffrey L., Furman & Stern S. (2006), Climbing Atop The Shoulders of Giants: The Impact of Institutions on Cumulative research, NBER Working Paper no. W12523, National Bureau of Economic Research, Cambridge.

Kebler, C. (2010), *Context-Aware Semantics-Based Information Retrieval*, PhD Thesis, Georgsmarienhütte, Deutschland.

Khosravi, A. & Lu J., Fault Modeling for Nonlinear Systems Using ANFIS, Proceedings of the International Multiconference on Computer Science and Information Technology, pp. 75-83.

Lee S., Hanna Milford & Jones D.(2008), An Adaptive Neuro-Fuzzy

Inference System for Modeling Mechanical Properties of Tapioca StarchPoly(Lactic Acid) Nanocomposite Foams, *Starch/Stärke 60*, pp. 159-164.

Malhotra Y. (2003), Measuring Knowledge Assets of a Nation: Knowledge System for Development, Research paper prepared for the invirted Keynote Presentation delivered at United Nations Advisory Meeting of the Departement for Public Administration and Development Management.

Miller M., DuPont B., Fera V., Jeffery R. Mahon B, Payer B & Starr A.

(1999), Measuring and Reporting Intellectual Capital from a Diver Canadian Industry Perspective: Experiences, Issues and Prospects

Nikitina N., Rudolph S. & Glimm B.(2011), Reasoning-Supported Interactive Revision of Knowledge bases, *In Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011), AAAI Press/IJCAI*, 2011, pp. 1027–1032.

Nilashi M., Fathian M. Gholamian M., Ibrahim O.& Khoshraftar A.(2011), The Identification Level of Security, Usability and Transparency Effects on Trust in B2C Commercial Websites Using Adaptive Neuro Fuzzy Inference System (ANFIS), International Journal of Artificial Intelligence And Expert Systems (IJAE), Vol. (2), Issue (3).

O'Hara K.. & Shadbolt N. (2001), Issues for an Ontology for Knowledge Valuation, *In Proceedings IJCAI Workshop on E-Business and the Intelligent Web, Seattle*.

Owaied, H. (2010), Frame Model for Intelligent Robot as Knowledge-Based System, *Special Issue of Ubiquitous Computing and Communication Journal*.

Piller, Christian (2009), 'Valuing Knowledge: A Deontological Approach'. Ethical Theory and Moral Practice, Vol. 12, No. 4, 2008, 413-428.

Salton, G. & Buckley, C. (1990), Improving Retrieval Performance by Relevance Feedback, *Journal of the American Society for Information Science*, 41, 288-297.

Schmidt, A. (2005), Potential and Challenges of Context-Awareness for Learning Solutions, *In the proceedings of the 13<sup>th</sup> Annual Workshop of the SIG Adaptivity and User Modeling in Interactive Systems (ABIS 05)*, 2005, 63-68.

Shing J. & Jang R. (1993), ANFIS: Adaptive-Network-Based Fuzzy Inference System, *IEEE Transactions on Systems,MAN, and CUBERNETICS*, Vol.23, No.3.

Sieg A., Mobasher B., Burke R., Prabu G. & Lytinen S. (2005), Representing User Information Context with Ontologies, *Proceedings of HCI International* 2005 Conference, Las Vegas, Nevada, 210–217.

Sivarao, Brevern P., El-Tayeb N.S.M.., Vengkatesh V.C. (2009), Representing User Information Context with Ontologies, *International Journal of Intelligent Information Technology Application*,, 191–198.

Soules C. (2006), *Using Context to Assist in Personal File Retrieval* (PhD thesis School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA).

Theriou, N., Maditinos, D., and Theriou, G. (2010). Knowledge Management Enabler Factors and Firm Performance: An empirical research of the Greek medium and large firms. American Journal of Economics and Business Administration, 254a.

Vanrompay Y., Yasar A., Preuveneers D. & Berbers Y. (2010), Context-aware Optimized Information Dissemination in Large Scale Vehicular Networks, *PETRA'10 conference*, June 23 - 25, Samos, Greece.

Wikipedia,<<u>http://en.wikipedia.org/wiki/Neural\_network</u>>, viewed by 24 December 2011.

Wikipedia, <a href="http://en.wikipedia.org/wiki/Research\_and\_developmentviewed\_by">http://en.wikipedia.org/wiki/Research\_and\_developmentviewed\_by</a>
12 May 2012.

WIPO (2010),

http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/pdf/941\_2010 .pdf, viewed on 13-5-2012 at 1:41 am.

Yao J. (2003), Knowledge Based Descriptive Neural Network, http://www2.cs.uregina.ca/~jtyao/Papers/1215.pdf

Young, R. and Letch, N.(2003), Knowledge Contexts: Through the Theoretical Lens of Niklas Luhmann, *7th Pacific Asia Conference on Information Systems*, Adelaide, Australia, University of South Australia.