Soft Computing Based Framework for Risk Assessment in Global Software Engineering

Hufsa Malik
2010-NUST-MS PhD-CSE(E)-19

Supervisor
Dr. Aasia Khanum

DEPARTMENT OF COMPUTER ENGINEERING
COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING
NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY (NUST)
ISLAMABAD

JUNE, 2014

Soft Computing Based Framework for Risk Assessment in Global Software Engineering

Hufsa Malik
2010-NUST-MS PhD-CSE(E)-19

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Computer Software Engineering

Thesis Supervisor

Dr. Aasia Khanum

Thesis Supervisor's Signature:____

DEPARTMENT OF COMPUTER ENGINEERING
COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING
NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY (NUST)
ISLAMABAD

JUNE, 2014

THE COMMITTEE

Soft Computing Based Framework for Risk Assessment in Global Software Engineering

A dissertation Presented by

Hufsa Malik

2010-NUST-MS PhD-CSE-19

Approved as to style and content by:

Dr. Aasia Khanam, *Supervisor*

Col Dr. Farooque Azam, *Member*

Dr. Saad Rehman , *Member*

Dr. Shahzad Khalid, *Member*

Dr. Aasia Khanam Asst
Prof, Computer Engineering

Declaration

I certify that this research work titled "*Soft Computing Based Framework for Risk Assessment in Global Software Engineering"* is my own work. The work has not been presented elsewhere for assessment. The material that has been used from other sources has been properly acknowledged / referred.

Hufsa Malik
2010-NUST-MS PhD-CSE-19

Language Correctness Certificate

This thesis has been read by an English expert and is free of typing, syntax, semantic, grammatical and spelling mistakes. Thesis is also according to the format given by the university.

Hufsa Malik

2010-NUST-MS PhD-CSE-19

Signature of Supervisor

Copyright Statement

Acknowledgements

I am thankful to my Creator Allah Subhana-Watala, for bestowing my life with utmost blessings and mercy, and to have guided me throughout this work at every step, paving the way for my progress, enlightening my mind with revelation, finally rendering my thoughts capable of deciding the way forward for my research and implement it.

I am profusely thankful to my beloved parents who are always there for me and are great support in my progress and in all aspects of my life.

I would also like to express special thanks to my supervisor Dr. Aasia Khanum for her help, guidance and support, throughout my thesis.

I would also like to thank Col Dr. Farooq e Azam, Dr. Saad Rehman, Dr. Shahzad Khalid and Dr. Arsalan Shoukat for being there for my thesis guidance.

I would also like to thank all those software professionals who gave their valuable and experiences based contribution during survey and were always available for series of discussions.

Finally, I would like to express my gratitude to all the individuals who have rendered valuable assistance to my study including my aunt Ms. Shumaila for her support, Lt Col Asif Saeed ur Rehman for his valuable input and guidance and my class fellows specially Mr. Muhammad Ali, Ms. Rabia Bashir and Ms. Asma Altaf, who have always been great help and support whenever I got stuck during my work.

Dedicated to my beloved Grandmother Bilqees Begum (Late), my exceptional parents Mr. Muhammad Ali Malik, Mrs. Tehmina Malik and adored Aunt Ms. Shumaila, whose tremendous support and cooperation led me to this wonderful accomplishment.

Abstract

Global Software Engineering (GSE), despite its practical importance, is still an immature field, with palpable shortage of systematic guidelines and best practices in various contexts. This thesis presents an approach which uses Soft Computing paradigm for evaluating risks in GSE. The research work is presented in five parts. First part describes concept evolution, encompassing the introduction of GSE, its importance as an emerging trend, its benefits, challenges and specifically risks involved in adopting GSE, leading to the fruition in the form of Risk Hierarchy. Second part presents the technical drive being followed, in order to assess the categorized risks, which comprises introduction to Fuzzy Logic, Neural Networks, Adaptive Neuro-Fuzzy Inference System (ANFIS), and finally the building blocks, architecture, design and structure of the proposed system emanating from the adoption of above mentioned technologies, to evaluate risks involved in specific GSE project. However, since the system is training based, in connection to that, the process of empirical data collection is discussed along with the mention of training and testing of the system, heading to the results phase, which comprises evaluation of acquired results using error measures and graphical presentation of accuracy of the model. Contingent to an attained accuracy level, the conclusion is drawn that the proposed framework is reliable enough to be adopted by project managers working on globally distributed projects.

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1 Background, Scope and Motivation

Since the number of outsourcing ventures to carry out software development by using globally distributed and temporally or culturally diverse human resource increased heavily in last two decades, Global Software Engineering (GSE) has been highlighted as a specific research area. However, studies show that still many relevant areas and their details are left uncovered by the current research. In this regard, management is the major point of concern for researchers. If we narrow down various aspects of management we would definitely come to the conclusion that Risk Management in GSE is still a field that needs much more exploration than paid before and needs to be addressed and sorted out on priority basis keeping in view the number of failed projects and consequently the ultimate financial loss to the software industry. This thesis contributes towards risk assessment in the field of GSE and helps to assess the level of risk using Soft Computing paradigm. Along with an overall risk assessment framework, this dissertation shows the categorization of risks involved in GSE and their hierarchy. While accomplishing the said task, this research work encompasses concepts like Risks, Risks taxonomy, Soft Computing, Fuzzy Inference System (FIS), Neural Networks (NN), Adaptive Neuro Fuzzy Inference System (ANFIS) and various risk assessment models from other disciplines, leading to form the basis and providing dimensions for proposed Soft Computing based Model for Risk assessment in GSE.

1.2 Global Software Engineering (GSE)

Software Engineering itself is relatively newer discipline among various types of engineering disciplines and occupies a significant situate among them, due to its peculiar characteristics concerned with its intangible nature giving rise to uncertainty, consequently rendering managerial tasks and assessments tough and time taking. This industry/ field is however getting popularity and is growing in importance at exponential rate with every passing day as this era is marked as the era of IT. The versatile nature of Software Engineering has necessitated division of this discipline into various sub disciplines according to the rising needs, concepts and advancements. One such evolved form of Software Engineering is Global Software Engineering (GSE).

GSE involves the conduct of software development life cycle at distributed level, outsourcing the employees of same firm or as a collaborative venture of different organizations, requiring innovation and effectiveness of followed practices and techniques for achieving the desired quality of the resulting product.

With the emergence of era of globalization, many new trends flourished in the software industry along with all other disciplines. GSE is a newly adopted but rapidly growing phenomenon of Software Engineering, which is not only gaining acceptance by IT professionals, but also achieving a vital attention of researchers due to its huge impact on global IT environment. During the last decade a remarkable shift has been observed in the approaches and practices followed by various software developers/ firms, regarding development of software projects. However, as agreed upon by various researchers, it is, like software development itself, a complex practice area.

1.3 Benefits of Global Software Engineering

Outsourcing with global IT services and software development has been ranked as one of the top business ideas of the past 100 years according to Harvard Business Manager [1]. It is perhaps due to the reason that in spite of being highly complicated and risk involving approach, GSE brings many benefits to its practitioners, providing them with reasons to be adopted and undoubtedly serves as a key to competitive environment. Some of the core benefits of GSE include:

1.3.1 Versatility of Expertise

Since various geographical areas have different professional and technical practices followed, it makes their expertise differ from people residing at other locations. Thus, hiring human resource from more than one location makes variety of specialties available to the team and helps them in achieving excellence in the competitive development environment. For example, country A has got more universities offering GIS as specialized subject or better opportunities / channels to polish GIS skills of its resources, it is likely to develop more skilled manpower at GIS as compared to other countries. Now, another country B has produced more opportunities for its resources, at information security domain and country C has got best potential for web development, simulation and graphics but has got less expertise in GIS and Information Security. Now, if an organization wants to acquire software which involves all of above mentioned components, it can hire relevant resources from countries A, B and C and create a global team, which can start work at their own locations through mutual collaboration and online communication, to develop the desired application. In this case GSE enables a complete package for the stakeholders and they do not need to get the team physically present at one place, which in turn offers another set of advantages.

1.3.2 Better Intellectuality and Novelty of Ideas

GSE involves hiring of resources from more than one location. When minds from various locations/ different backgrounds are involved, it broadens the scope of intellect and vision, consequently becoming a great help to add creativity to the product. Scope of ideas for different projects may vary from location to location subject to the advancement and background of locality. Naturally, every area has its own historical/ cultural impacts, which molds the thinking patterns of its manpower and their approach toward problem solving. Thus the combination of such diverse intellect available can prove helpful for the purpose.

1.3.3 Better Time Management

Due to dispersed human resources, management gets margin of availability of resources at all slots of time by invoking round the clock working of the team in case of their locations in different time zones. For example, a company in North America hires few resources from Canada and others from Middle East courtiers, better time management can be achieved, since the time gap between locations will be of several hours, when one team will finish day work, second will be starting it off and if their work is interdependent on each other’s input, then the span of completed one day will be cut off, since, when one team will start off next day, second team had already forwarded their out put, while in case of same location, both team had either to work in day/ night shifts or wait for another day to seek other’s output.

1.3.4 Better Hardware Resource Management

Better hardware resource management refers to round-the-clock usage of hardware resources by various members, favoring the idea of cloud computing. Due to the rising concept of cloud computing, hardware resources present at one location can be used by teams of distant locations as well. This concept, when incorporated in GSE prevents GSE vendors/ developers from purchasing high amount of hardware devices and making them available to the other team members.

1.3.5 Lesser Requirement of In House Arrangements / Equipment

GSE provides an opportunity to run business without huge buildings along with furniture requirements, / IT equipment requirements, utility bills and many small, unnoticed expenditure, saving owners from huge mess and financial strains. Therefore, almost 40% of one time investment in arranging hardware equipment and furniture, offices, stationary can be avoided using this kind of development methodology.

1.3.6 Cost Effectiveness

Besides other financial benefits, a lot of amount to be spent as wages are cut off due to the difference of labor cost and currency differences at various regions/ countries. For example, if an organization in Australia, hires team members from South Asian countries, there will be huge currency difference, consequently encouraging human resource due to better currency rates, which when converted to their own will be much better.

1.3.7 Huge Opportunities for IT Professionals

Besides its contributions for development firms, GSE provides promising opportunities to IT individuals all over the world for their career boost and financial strength. This, in turn assists the young IT resources to flourish their skills and gain vide experience along with world wide exposure to latest development and technology trends.

1.4 Challenges of Global Software Engineering

Although, GSE has earned huge acceptance through its distinguishing benefits, the approach brings challenges to its followers. Arising through various areas are certain incongruities which are considered potential enough to halt the flow/ consistency of GSE processes. Some vital aspects which serve as challenge in this process are mentioned below:

1.4.1 Technical Challenges

Technical challenges arise due to technological advancements and their uncertainties. IT is the fastest growing and evolving industry of current era; however, due to novelty of technologies that emerge day to day, many concepts/ tools and techniques in the same domain become obsolete every year. Still, statistics of obsolete and adopted technologies are quite different on implementation grounds at various locations.

Technical advancements vary from state to state, therefore, their software potential in those technologies also vary, based on level of their progression speed. Some countries, which possess high rating among technologically well-established nations, may declare certain technology as obsolete, but same may still be in use in other regions with moderate technical progress. This affects the difference of approach towards selection of various technical parameters during project planning and development, consequently, giving rise to difference of opinion/ understanding or adoption of approach towards developmental stages of the project. Besides inducing the disagreement among stakeholders, this may also affect the level of expertise in arising technologies. For example, engineers from a country with more advanced technical profile will have better expertise in modern developmental aspects/ techniques or tools than the others.

Other than these aspects, the idea of adopting newly invented technology brings a reasonable amount of risk with itself, as we are not sure of its future prospects and what edges it can provide, since practical pros and cons of products are revealed with the passage of time only and is also contingent upon the extent of its usage.

Proceeding further, several technical challenges take place in GSE from implementation point of view as well. Due to distance among team members, software once developed may present challenge at the time of integration of its different modules. Similarly, compatibility issues may arise at the time of system configurations and deployment.

Foregoing, it can be stated that up-to-date technology itself becomes a challenge in the field of GSE.

1.4.2 Organizational Challenges

Project management of any project brings challenges, however, software being intangible and unique sort of product is extra demanding as far as management is concerned. Even among software projects, globally undertaken projects are most vulnerable to organizational and managerial hazards.

Organizational challenges refer to the human resource management, time management, managing the hierarchy, controlling the team from remote location and maintaining time and budget constraints. Preparing project plans and analyzing the progress presents a huge challenge to project managers and other stake holders since various components of the system being developed are distributed among different locations.

1.4.3 Interactional Challenges

Interaction among stakeholders can be easily managed due to technological advancements of IT industry and availability of fast and modern tools of communication like Skype, Team Viewer etc. However, this distant interaction, provided by latest software tools can not substitute natural aptitude of human being towards live conversation in the same workplace and same environment. Remote conversations and video conferencing based discussions can pass information into minds but still, they have no comparison to the understanding developed among the people working under same circumstances. Productivity and approach towards development is much influenced by the working patterns and conventions of the surroundings which in turn vary from region to region. Thus, there is always possibility of communication gap in GSE projects, which, often has the potential to bring financial and time loss to the organizations.

Other than distance, language is extremely critical area in this context. Language difference, even accent difference can bring difference of 180 degrees to the concepts and perception. While better understanding of requirements and clarity of concepts/ way forward at higher and lower level is the key to success in software projects, such lack of perception may bring disastrous impacts to the project.

1.4.4 Location Impact

Besides language and face to face conversational hurdles, regional differences also possess certain risks which can prove themselves as challenging tasks to overcome. They arise due to social, cultural, political and economic situations of states, since contrast of all these factors actually lead and depict the behaviors of local residents. Project managers must be aware of these characteristics of the community from where they are going to hire manpower or include them in stakeholders list since only then they will be able to estimate and predict their working capabilities along with the weaknesses. This assessment will actually aid them in creating effective resource allocation and time management plans.

1.5 Problem Statement

Each of the challenges categorized in this chapter serve as a root cause to various evolving risks; however, ratio of occurrence of these risks differ from project to project depending upon the nature of project. Though many risks are known to occur in GSE, still there is lack of detailed work on risk taxonomy for GSE in modern research work. Moreover, there are no systematic tools for risk assessment in GSE.

The purpose of this work is "to present a comprehensive framework for risk assessment in GSE" which involves the following sub-tasks:

i. To present a taxonomy of major risks in GSE
ii. To present a software approach for risk assessment in GSE based on the proposed risk taxonomy
iii. To propose a prototype tool for GSE risk assessment.

1.6 Summary

Global Software Engineering is an emerging and vigorously adopted discipline of Software Engineering. In the context of current Globalization, and economic constraints worldwide, pressure has been built up on various software vendors/ firms in meeting their set goals. Therefore, due to its ease in espousal, GSE as earned vast acceptance globally. Like every other field of work, GSE has its own benefits and challenges. Being a newly adopted and complex field, GSE claims massive challenges but its benefits can easily outnumber the risks in the contemporary business situation and economic crises all over the world. The need is to have proper homework for projects and careful risk assessment study before starting the projects.

CHAPTER 2: LITERATURE REVIEW

During literature review of our research work, it was revealed that in the past few years, many studies appeared to envisage a layout for solid solutions towards issues pertaining to GSE, however, problem still exists and especially the area of Risk assessment in GSE is yet to be addressed with more efforts. Study of relevant papers was extended to provide way forward for this research work and include reviews and concepts by other researchers on the discussed issues of this dissertation. Various research papers, books, essays and reviews in any other documented form were studied encompassing concepts like GSE - its benefits and challenges, Current State of Research in GSE, Risks involved in GSE, Risks Assessment in GSE, Fuzzy Logic, Neural Networks and ANFIS (Adaptive Neuro-Fuzzy Inference System) etc. Some of the important papers studied during this process are one by one discussed below:

2.1 Research on State of GSE

Darja Smite et al, in their paper "Empirical Evidence in Global Software Engineering: A Systematic Review" evaluated recent state of the art in GSE by systematically reviewing empirical research related to GSE projects. Researchers concluded their results by declaring GSE as an immature field with lack of empirical evaluation of techniques and methods in industrial context and emphasized on the urge of in depth analysis of the same [6].

2.2 Challenges and Benefits of GSE

Muhammad Ali Babar and Mansooreh Zahedi, in their report "Global Software Development: A Review of the State-Of-The-Art (2007 – 2011)", presented a review aimed to identify the state of the art in research pertaining to GSE challenges, benefits and related issues, along with identifying the areas being focused and existing gaps in moderns literature regarding the same [7]

Rafael Prikladnicki et al, in their paper "Global software development in practice lessons learned", discussed the factors urging software business personnel to move towards GSD and lessons learned from GSD practices [5]

Muhammad Ali Babar and Christian Lescher in their editorial for ICGSE, "Global

Software Engineering: Identifying Challenges is Important and Providing Solutions is Even Better", described their categorization of major challenged involved in GSE and proposed mitigation schemes for them as well [3]

Bilal Raza, Stephen G. MacDonell and Tony Clear; in their paper "Topics and treatments in global software engineering research - a systematic snapshot" analyzed the issues currently being addressed in research papers regarding GSE. The researchers used new Systematic Snapshot Mapping (SSM) technique for this purpose and proposed their conclusion that managerial and infrastructure matters are currently being focused in literature addressing the topic, instead of distance or human related factors. [8]

James D. Herbsleb; in his paper, "Global Software Engineering: The Future of Socio-technical Coordination", describes distant coordination as key phenomena in GSD and emphasized on the need to develop better understanding of the required kinds of coordination [13]

Paivi Parviainen and Maarit Tihinen in their paper "Knowledge related challenges and solutions in GSD: Expert systems", analyzed knowledge related challenges of GSD and presented knowledge engineering as a key player in GSD. Authors have worked on 50 case studies from cognitive perspective and have presented example solutions as well to mitigate GSD challenges [9].

2.3 Risks involved in GSE

Indira Nurdiani et al, in their survey paper "Risk Identification and Risk Mitigation Instruments for Global Software Development: Systematic Review and Survey Results" presented a checklist for risks involved in GSD and their mitigation through systematic literature review [10]

Qadeem Khan and Shahbaz Ghayyur; in their paper "Software Risks and Mitigation in Global Software Development"; presented a taxonomy of GSD risks and also suggested guidelines for mitigating these risks [12]

. Christof Ebert et al, in their paper "Managing Risks in Global Software Engineering: Principles and Practices"; discuss empirical study of various risks and mitigation practices in GSE [4]

Darja Šmite and Juris Borzovsin their paper "Managing Uncertainty in Globally

Distributed Software Development Projects" devised a probabilistic framework comprising a knowledge base and a risk barometer for calculating risk in GSE, especially for managers inexperienced in global development. [11]

Sarah Beecham et al in their paper "A Decision Support System for Global Software Development" introduced the Global Teaming Decision Support System - GTDSS for GSD for managers [14]

Christer Magnusson et al in their paper "Risk and Compliance Management Framework for Outsourced Global Software Development", presented a GSD risk and compliance management system by combining COSO-ERM, ISO 20000 and ISO 27001 standards for ERP systems [15].

2.4 Approaches for Risk Management

Ansgar Lamersdorf et al in their paper "A Rule-based Model for Customized Risk Identification in Distributed Software Development Projects", presented a rule-based model for early assessment of risks in GSD. The system can adapt itself in light of experience to work appropriately for different project characteristics [16]

Ronald Jabangwey et al in their paper "Risk Identification and Risk Mitigation Instruments for Global Software Development: Systematic Review and Survey Results" presented a survey on prominent risk management approaches in GSD [17].

2.5 Soft-Computing based Decision Making Systems

Marta Takacs in his paper, "Soft Computing-Based Risk Management - Fuzzy, Hierarchical Structured Decision-Making System", presented soft computing based decision making system principle using Mamdani approach. [18]

Literature review, encompassing several worth reading research publications, books, editorials, few of which are stated above, clearly ratifies the significant requirement to dig out the risk assessment in GSE, since there is very little work found worldwide in the field of GSE, compared to other domains, especially in case of risk assessment in GSE.

2.6 Summary

In the past few decades research work has been conducted encompassing various aspects of GSE, however the field, in spite of its extensive usage is still lacking any solid proposals/ techniques/ writings encompassing various contours of GSE especially in context of Risk Assessment. However, few researchers have made efforts in the field by wring on important factors, threats, managerial issues, benefits and challenges of global software engineering, still the work is in the form of chunks and not very organized, or concrete models in these aspects are proposed yet.

CHAPTER 3: THEORETICAL BACKGROUND

The contemporary research state in the field of GSE, presented in chapter 2, clearly notifies about the lack of risk assessment models and tools for GSE projects. This research work is aimed for the same purpose. In spite of few efforts made on the lines of desired achievement, Soft Computing based framework for risk assessment in GSE will be an addition to the dimensions being worked out in industry. This chapter specifically introduces fundamental concepts and pivotal paradigms of Soft Computing technology in a sequence of their technical evolution for providing an understanding for the proposed model.

3.1 Fuzzy Inference System (FIS)

Fuzzy Logic is an approximate reasoning mechanism which supports variable values ranging between 0 and 1 unlike typical crisp digital output, therefore it offers acceptance for linguistic variables as well. The concept of Fuzzy Logic was presented by Lotfi Zadeh, a professor at the University of California at Berkley back in 1965. Idea was based on the theory of Fuzzy Sets and it got popularity in the decade of 1980s in the fields of control systems [15]. The concept enumerates that more precise results can be generated from control systems if they are not bound to produce output in true or false values.

Fuzzy Logic is commonly used in various dimensions including Artificial Intelligence, embedded/ real time applications, security decision systems, variety of electronic appliances used for both commercial as well as domestic purposes and analytical projects. However, significance of FIS is not limited to control systems, rather another aspect of its popularity lies in its utilization as expert systems with the capability of analyzing the information and produce results/ suggestions simulating the human reasoning.

There are two approaches used to comply above mentioned capabilities of the FIS; Mamdani and Takagi-Sugeno (TS) inference systems. In Mamdani model, both inputs and outputs are fuzzy, while TS model integrates fuzzy inputs with crisp outputs.

The calculations in FIS comprise 3 main steps which include fuzzification, inference and defuzzification. Fuzzification refers to the conversion of real variables into linguistic variables.

Inference refers to the processing of the system according to the set behavior on the basis of predefined rules. Usually, generalized modus ponens is used as inference mechanism.

Finally, defuzzification refers to the conversion of found results into output variables.

3.1.1 Implication of FIS in Risk Assessment

Besides its contribution to enormous number of systems, Fuzzy Logic systems have now acquired place in the field of risk management where algorithm based decision making is difficult due to the versatility of the projects undertaken.

This variance becomes unpredictable when it comes to the Software based projects due to intangible and novel nature of the software. However as discussed earlier about GSE, it is the field in which risk assessment is still a question mark due to its newly emerging dimensions.

Specially, fuzzy systems are still not proposed for risk assessment in GSE and more research is needed to be conducted in this area.

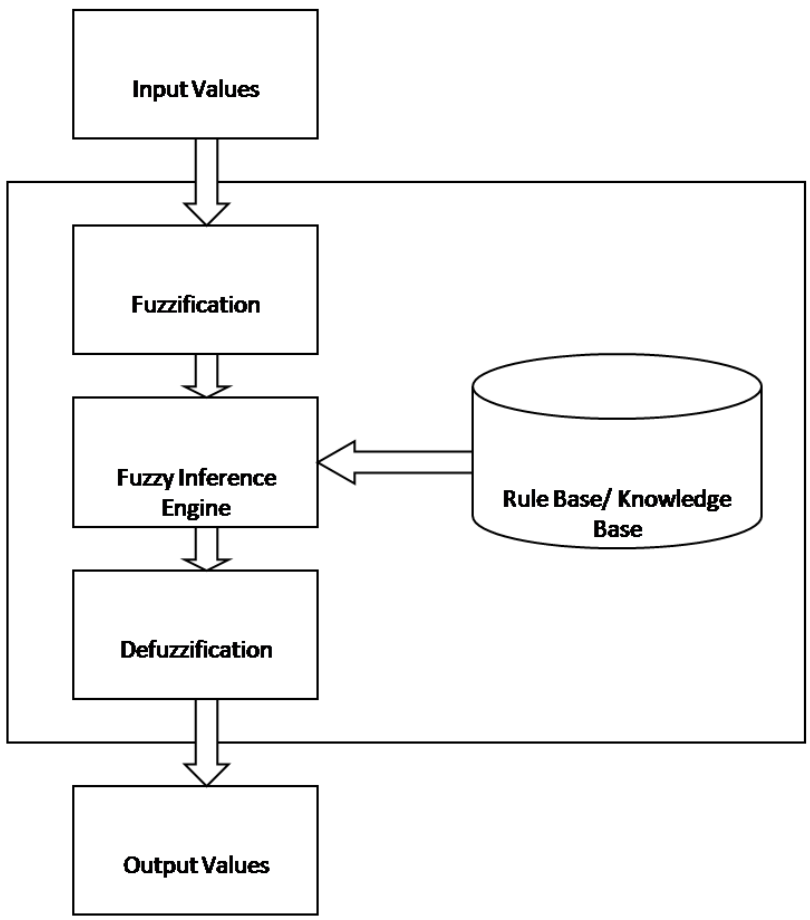

Figure 3.1: Working Mechanism of FIS

3.2 Artificial Neural Network (ANN)

Artificial neural networks are the mimicry of biological neurons and are devised for intelligent or Expert Systems to simulate human decision making process on the basis experience in the form of previous data to train the system instead of using any fixed algorithm. This makes neural networks much more flexible and near to the human decision making patterns in the real world.

Networks of neurons may be feedforward or feedback. In case of feedforwad network, signals in the network or input values travel only in the forward direction making such type of networks more useful for pattern recognition

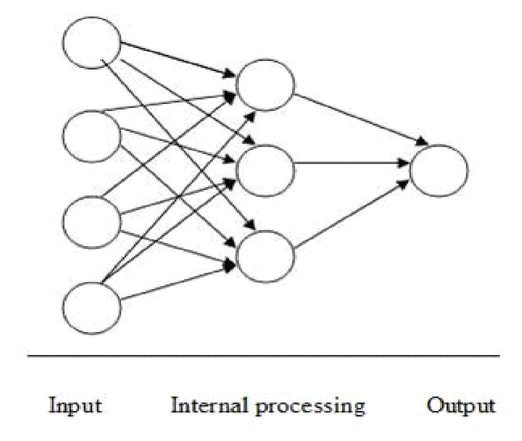

Figure 3.2: FeedForward Network

Signals may move in backward or any direction before delivering the output in case of feedback networks introducing loops in the network (Nunes and Marques, 2012) [21].

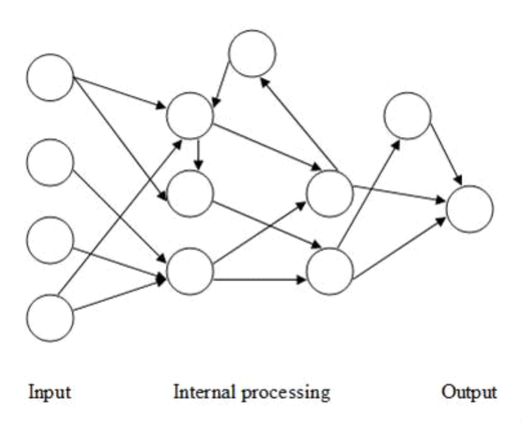

Figure 3.3: Feedback Network

Neural networks simply take information in the form of input variable values, process it and provides output in the form of output variable values. However, mechanism involved in this processing may vary depending upon the requirements of the system.

3.2.1 Implication of Neural Networks in Risk Assessment

Since Neural Networks provide flexibility to achieve results on the basis of trained data, it helps a lot in risk assessment as we can gain much visibility of the risks and their effects from historical data of similar projects.

NN have been used for risk assessment in business domain for analysis and risk calculation however not adequate approaches are yet proposed in the field of Software Engineering and specially the GSE.

3.3 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS integrates both approaches, FIS and ANN; therefore it enjoys the potential to attain the benefits of both above mentioned approaches in a single framework. The FIS provides facilitation in knowledge elicitation from experts and specifying it in an easily interpretable form. However, generally the structure of an FIS is chosen arbitrarily depending on intuition of the expert. An ANN provides capabilities to optimize the FIS structure on the basis of actual data. An ANFIS is created by combining a Takagi-Sugeno (TS) model with an ANN to optimally combine learning and knowledge.

3.3.1 Implication of ANFIS in Risk Assessment

Since risk assessments have more to do with experience than calculations especially in case of software projects, therefore implication of ANFIS to propose a risk assessment framework sounds a productive approach with promising outputs.

3.4 Summary

FIS and Neural Networks constitute the fundamental concepts and give rise to pivotal paradigms of Soft Computing technology. Mamdani and Takagi-Sugeno are two approaches of FIS while Neural Networks can be feed-forward or feed-backward. FIS and Neural Networks as well as ANFIS have their implications in decision making systems, which can be beneficial in risk assessment as well.

CHAPTER 4: PROPOSED FRAMEWORK

In the light of overall discussion carried out above, an urge for risk assessment model for GSE projects was felt necessary in order to fulfill the challenges in the field and achieve maximum benefits from the GSE approach. Keeping in view the complexity of GSE paradigms and uncertainties/ overlapping of input parameters of the GSE and significant utilization and optimistic result of Soft Computing approaches used in modern time, a Soft-Computing based model is devised and proposed for Risk Assessment in GSE. Detailed description of the model along with envisaged concepts is provided below:

4.1 Elements of Proposed Framework

Proposed framework for risk assessment in GSE includes several elements which at a broad level may be categorized into Risk Taxonomy and Risk Assessment. The Risk Taxonomy is composed of Risk Factors and Risk Categories whereas Risk Assessment encompasses Probability of Risk occurrence, Severity if they occur, and Risk Impact, all depicted in the form of Risk Matrix. The relationship between these framework elements is depicted in Figure 4.1.

For identification of these elements in GSE context, we have resorted to the technique of Grounded Theorizing, wherein we have collected qualitative data from various sources including surveys, questionnaires, interviews, and research literature to conceptualize and categorize important risk factors and risk categories along with their inter-relationship. Below we describe each element in detail:

4.2 Risk Taxonomy

The taxonomy of GSE risks is arranged in the form of Risk Factors and Risk Categories where each category is influenced by several factors of risks as described below:

4.2.1 Risk Factors

Risk factors refer to the attributes of the project which can suffer from various kinds of risks or threats. These factors are meant for the classification of the project so that it may be evaluated as to how much risk is involved in a specific type of project with the given criteria.

Risk Factors in proposed model are identified as Industry domain, Product Context, Stakeholders Context, Constraints, Work Distribution and Work Done. Each of these is described next:

Industry Domain

Industry domain refers to the business aspect classification of the project. It defines the specific industry where the developed application will be deployed. Industry domains are identified from broader perspective of system deployment and were identified through brain storming and empirical data available on Internet, irrespective of GSE or or local software development. Later, step-wise refinement approach was followed in a way that GSE-relevant domains were retained and non-relevant domains were discarded.

Product Context

Product context refers to the application domain from software development point of view. This was again identified through brain storming and prior knowledge. Product context classifies the project from the perspective of nature of the software being developed from implementation/ development viewpoint.

Stakeholders Context

Stakeholders' context refers to the number of stakeholders involved in the project. Naturally, number of stakeholders always affects the productivity since more number of people is directly proportional to more viewpoint diversity consequently leading to more difficulty in the achieving consensus.

Constraints

Irrespective of GSE project or not, constraints are always there on projects. These constraints/ limitations may be of several types, however most appropriate and most influential constraints in GSE projects are categorized under this title.

Work Distribution

Distribution of work/ tasks is the key factor in any sort of project management; however, in case of GSE it is likely to be distributed in number of locations. The number of locations involved directly affects the affectivity and efficiency of work. Under this heading, work distribution is referred to as the number of locations in which the tasks/ resources are distributed.

Work Done

Work done refers to the amount of work already completed when the systems risks are being assessed. This attribute can obviously be helpful if the project plan has truly been made and calculations regarding work packages/ deliverables and phase vise distribution of tasks have been carried out. This attribute indicates that proposed models provides projects managers the flexibility of risk assessment at any stage of project development.

4.2.2 Risk Categories

Risk Categories refer to the classification of risks pertaining to the GSE projects. Depending on the conclusions drawn from various research papers, brainstorming and consultancy with domain experts a separate hierarchy for risks classification has been proposed as explained below:

A. Internal Risks

Internal risks represent the risks emerging from internal factors of the development like immediate stakeholders, team members, contractors / purchasers or technology related issues. Internal risks are further divided into Technical and Managerial threads.

Technical Risks

Technical aspects give rise to three types of risks namely bleeding edge/ Cutting edge technology risks, compatibility risks and integration risks.

Bleeding Edge Technology Risk- Bleeding Edge Technology Risk refers to the criticality of using emerging technology since new technologies does not possess surety of success until experienced. And IT technology offers the fastest evolution rate proffering more and more challenges to the developers.

Compatibility Issues Risks-Compatibility issues can arise in context of development in more than one location due to versions of the software / hardware used.

Integration Issues Risk- Integration issues are progeny of modular development and can take drastic shape when development is done in geographically separate locations.

Managerial Risks

Managerial risks include Financial Risks, Poor Supplier Service risks, Human resource Risks, Inadequate IPR management risks, project deliver risks and inadaptability with overly high change rate.

Financial Risks- Financial Risks arise due to insufficient budget planning and overlooking inflation rates in accordance with the timelines of the project and can be controlled if careful estimation and trend analysis of market rates is done. These are further classified into wage inflation cost inflation and lock in risks.

Wage Inflation Risks- Wage inflation refers to increase in wages with the passage of time which is specially risen if timelines are not followed.

Cost Inflation Risks- Cost inflation refers to cost of equipment / software used and lock in risks provide the seized amounts idea.

Poor Supplier Services Risks- Poor Supplier Services risks have high probability of occurrence in projects where third party is involved for provision of any kind of services/ hardware/software supplies as there is somehow a bit issue in monitoring if third party is located some geographically separate location.

Human Resource Risk- Human Resource risk is one of the major and most influential risks in such kind of projects. However, same can become the strongest point in success of any project if managed carefully leading to winning combination as a consequence of teamwork. Human resource aspect is further categorized into two areas pointing staff turnover and insufficient competencies.

Staff Turnover- Staff turnover refers to the rate of change of staff which can prove itself to be highly critical if occurs during the course of project as it takes time for new members to fully cope with the details and requirements and acquire understating of the ongoing work. This usually happens due to wages issues and thus two risks are interrelated with each other.

Insufficient Competencies- The other human resource issue is usually insufficient competencies. Competencies of resources is a challenge in present era of ever new evolving technology since GSE involve resources from more than one location so their level of understating and their competencies may vary and if so, consequences can take the form of lack of understanding thus creating loop holes in the activities.

Inadequate IPR Management- Inadequate IPR Management is a huge issue since Intellectual Property Rights (IPR) management becomes very complex in GSE.

Project Delivery Risks- Project Delivery Risks deal with quality and deadline risks of the product delivery. As discussed earlier in cost and wage inflation that they are directly proportional to the time delays therefore time is very critical in project delivery and same is the case with quality which is sometimes unintentionally compromised by the project managers in order to meet the deadlines. However, poor quality has long term effects on the business in the form of reputation of the firm.

Instability in Change Rate- Instability with overly high change rate encompasses the matters related to the changes proposed by stakeholders time to time. This risk fall under the aspects of change control management. Project managers should carefully mechanize the process of incorporating changes and monitor them, since uncontrolled change incorporation may bring expansion in the scope of project consequently disturbing the whole budget and timeline analysis and planning.

B. External Risks

External Risks are those which are caused by the external agents like global market trends, government policies, political stability, temporal differences etc. These risks are extremely influential in case of GSE since it involves more than one government, society and may be different time zones. Therefore, risks falling under the umbrella of external risks category are more general than internal risks but proffer very thin margin to deal with. Best approach is to conduct careful analysis and taking into account the up to date situation on external factors mentioned below before starting the project or awarding the contract. External risks involve Political, Economy related, distance related and legal issues risks details of which are mentioned below:-

Political Risks- Political Risks involves the political situation of the locations in which developers / team members / contractor or purchasers are located.

Economy Risks- Economy Risks arise due to global market trends and economic situation of the involved states.

Distance Related Risks- Distance Related Risks comprise the most typical risks of GSE which include Temporal Risks, Geological Risks, Socio Cultural Risks and Communication Risks. All of these risks are crucial and least likely to be eliminated in case of GSE.

Temporal Risks- refer to the time differences of the working hour due to different time zones especially when this difference is of 12 hours or more. However same is considered a strong point as well since this difference ensures the round the clock continuous work on the project, but still it may cause inconvenience of communication. Geological and socio cultural risks are progeny of language, values and cultural differences of various different geographical locations due to which working environment and work practices also differ thus giving way to communication gap.

Legal Issues Risks- Legal Issues Risks involve the risks related to contracts and governments policies of the involved states.

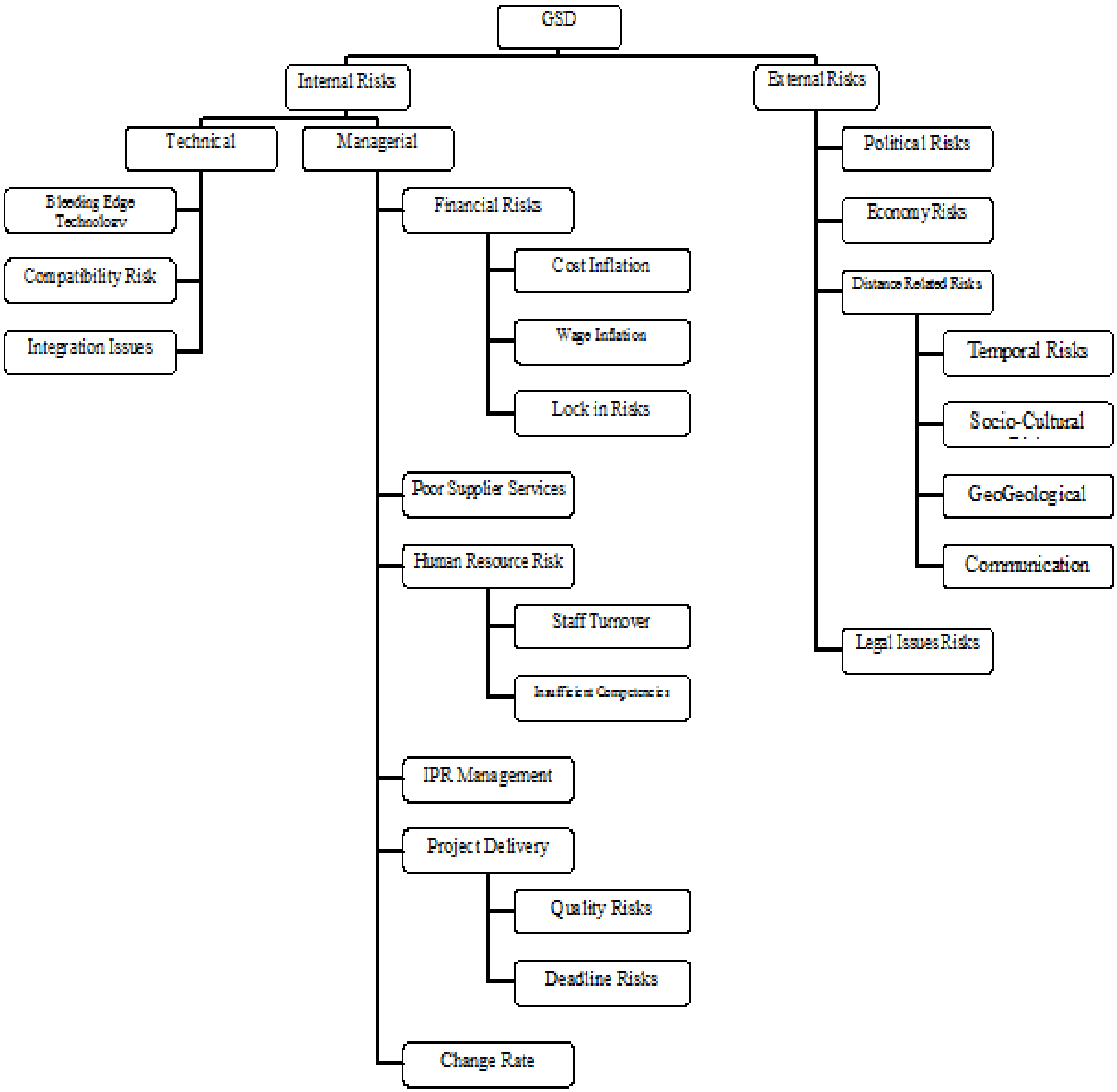

Figure 4.1: Risks Taxonomy for GSE

4.3 Risk Assessment

The risk assessment workflow is illustrated in Figure. Each element of the workflow is explained below:

4.3.1 Risk Matrix

In order to prioritize risk and assess its chronic state a risk matrix is defined consisting three elements Risk Probability, Severity, and Impact.

Risk Probability- By Risk probability we mean the chances of occurrence of risk in the specific project. It will be measured in context of the term set {Nil, Unlikely, Even, Likely, Highly Possible}.

Risk Severity- Severity of Risk will be evaluated in context of the termset {Very Low, Low, Medium, High, Very High}.

Risk Impact- Risk impact will be calculated on the basis of risk probability and severity in each type of risk for a specific project. Term set for Risk Impact is {Negligible, Low, Medium, High, Catastrophic}.

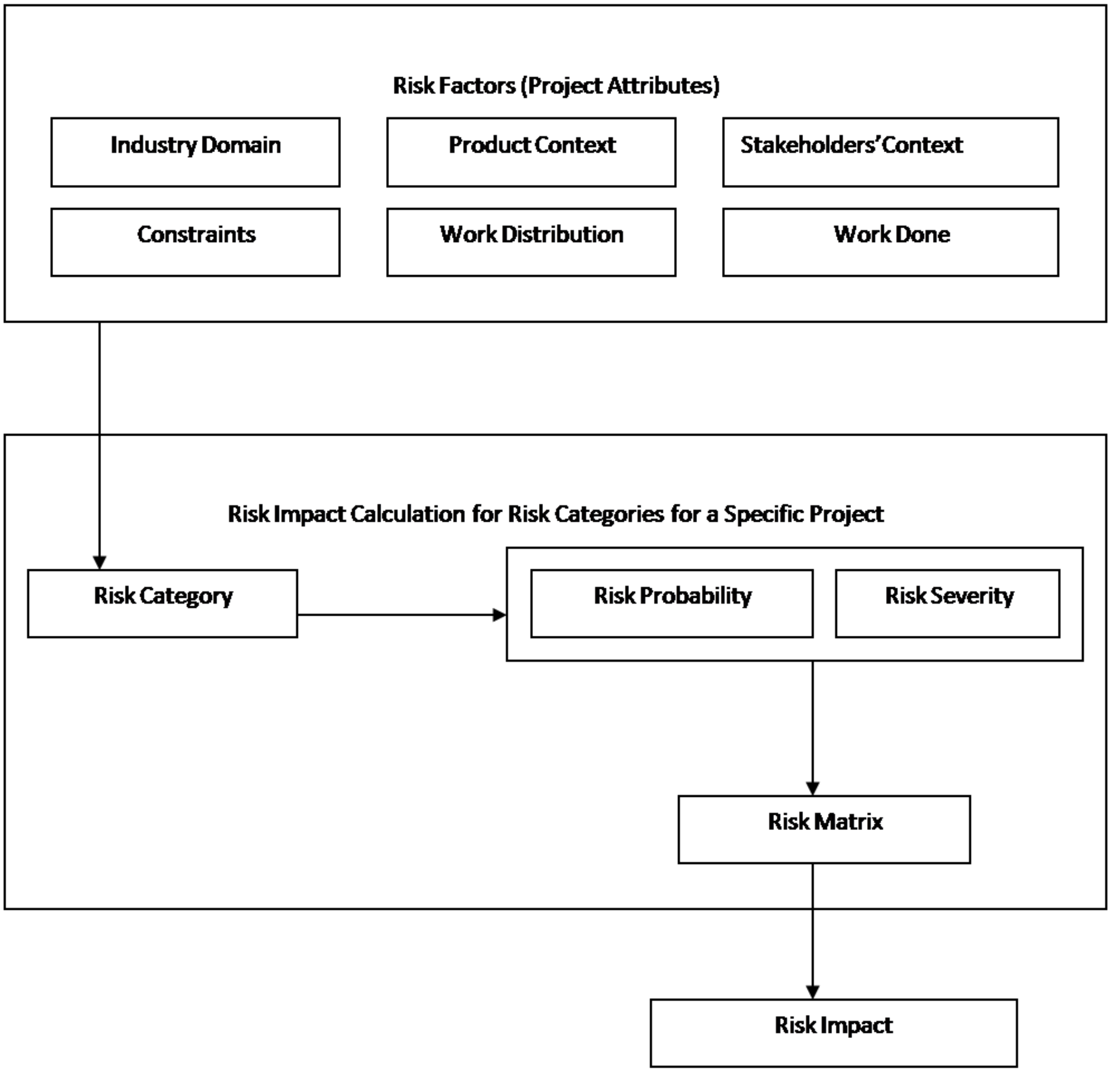

Figure4.2: Relationship of Risk Factors with Risk Category and Calculations

Risk probability and severity are combined in PMBOK-defined Risk Matrix to specify overall risk impact.

Table 4.1: Risk Matrix

	Very Low	**Low**	**Medium**	**High**	**Very High**
Nil	Negligible	Negligible	Low	Low	Medium
Unlikely	Negligible	Low	Low	Medium	High
Even	Low	Low	Medium	High	High

Likely	Low	Medium	High	High	Catastrophic
Highly Possible	Medium	High	High	Catastrophic	Catastrophic

4.4 Risk Assessment with Soft Computing Technique

The proposed framework is based on ANFIS. It uses a set of training data to learn an optimized FIS for assessing risk in a given GSE project. The trained system takes risk factors as input. It then works on this factor information to estimate probability and severity of various risk categories. On the basis of these probabilities and severities it generates an estimate of overall risk impact for the given project.

4.4.1 ANFIS Structure

ANFIS structure is shown in Figure. The ANFIS has a five layer structure, as described below:

Layer 0 (Input)-

Each of the risk factors will be input of the system. Possible values of each input are defined as fuzzy sets mentioned below which will become membership functions in the system.

Table 4.2: Fuzzification of Input Variables

Risk Factor	Fuzzy Linguistic Categories	Fuzzification
Industry Domain	{Business, Education, Health, Defence, Research, Game, Tool}	Singleton
Product Context	{ Web, Desktop App., Mobile App., ERP system, Real Time System}	Singleton
Stakeholders Context	{ Large, Medium, Small}	Trapezoidal
Constraints	{Budget, Time}	Singleton
Work Distribution	{2-3 Locations, 3-4 Locations,	Gaussian

	4-5 Locations}	
Work Done	{0-20%, 20-40%, 40-60%, 60-80%, 80-100%}	Triangular

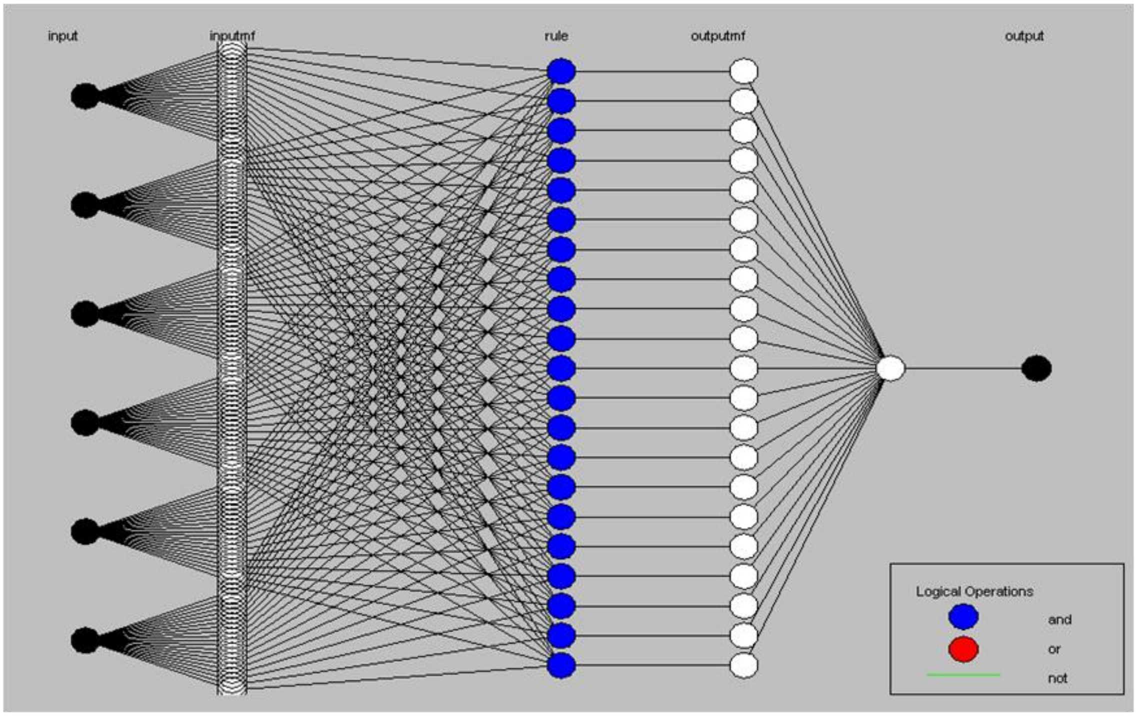

Figure 4.3: ANFIS Structure

Layer 1 (Fuzzification)

Fuzzification of the input values is the first step which is elaborated in table 2 below in which context type refers to the input variables and linguistic categories to the membership functions of those input variables. These Membership Functions are depicted in the Figures 5 – 10 as well.

Layer 2 (Rule IF part matching)

Triangular-shaped built-in membership function trimf was used by Matlab to manipulate the input values with following formula:-

$$f(x,a,b,c) = \max\left(\min\left(\frac{x-a}{b-a},\frac{c-x}{c-b}\right),o\right) \qquad \text{Equation 4.1}$$

where a, b and c are parameters while f is the function of vector x.

Each input in layer 2 provides strength to the rules by means of multiplication. This uses and operation and can be depicted in formula as (Senvar, 2013)[22]:

$$Wi = \mu_{Ai} \times \mu_{Bi} \qquad \text{Equation 4.2}$$

where Wi is the strength of the rule while Ai, Bi refer to input membership functions.

Layer 3 (Normalization)

This layer normalizes the strength of rules according to the formula (Senvar, 2013)[22]:-

$$\overline{Wi = \frac{Wi}{\sum_{j=1}^{R} W_j}} \qquad \text{Equation 4.3}$$

where i presents the rule no. and Wi depicts its strength.

Layer 4 (Rule Aggregation)

The aggregation of the rules is rules is done using the following formula (Senvar, 2013)[22]:-

$$Wifi = Wi(P_0 X_0 + P_1 X_1 ... P_n X_n) \qquad \text{Equation 4.4}$$

where p refers to input parameters.

Layer 5 (Defuzzification)

In the end fuzzy output is transformed into a crisp output which is called defuzzification and formula used for this purpose is:-

$$FinalOutput = \frac{\sum_{i=1}^{N} WiZi}{\sum_{i=1}^{N} Wi} \qquad \text{Equation 4.5}$$

where i presents the rule , Wi presents its strength and Zi is the output level of each rule.

4.5 Summary

Taxonomy of risks involved in GSE has been formulated categorizing them into internal, external, technical and managerial groups. Soft Computing based model is devised for calculating the risks proposed in taxonomy comprising elements of ANFIS.

CHAPTER 5: TESTING AND RESULTS

5.1. Test Setup

The proposed system was implemented in MATLAB version 7.0 environment. The type of network used in ANFIS is Feedforward Neural Network. Since it is a knowledge based evaluation, field survey was conducted in which software professionals working in the concerned field (GSE) were interviewed and questionnaires were asked to be filled by them in the light of their experience. These software professionals were working in the locations including Islamabad, Rawalpindi, Faisalabad, Peshawar, Wah Cantt, Bahawalpur and Dubai. Total 30 samples were acquired. Their input was used to train and test data. 70% of the data collected was used to train the system while 30% of data was used to test the generated ANFIS. The system was trained using hybrid method and 10 epochs. Error tolerance was set to be zero.

For evaluation purpose, three measures were selected, including RMSE (Root Mean Square Error), MAPE (Mean Absolute Percentage Error) and R (Correlation Coefficient). Datasets were both trained and tested for variety of membership functions against each risk category; however, it was observed that type of membership functions did not have any effect on the results. As seen in Table 5.1 convergence of test and trained data was attained. This supported the accuracy and likeliness of success of the proposed model.

The results of Table 5.1 are discussed in following sections:

5.2. RMSE

Root Mean Square Error (RMSE) is a measure of the differences between values predicted by a model or an estimator and the values actually observed from the thing being modeled or estimated. Since the RMSE is a good measure of accuracy, therefore, it is ideal if it has a small value. It is calculated as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(X_{obs,i} - X_{model,i} \right)^2}{n}} \qquad \text{Equation 5.1}$$

In this formula, n refers to the total number of data points, while Xs refer to observed and model (predicted) values.

After applying this formula, RMSE for various risk categories was calculated for both training and testing datasets. Different values were found out among which, the lowest error was found to be for communication risk and it was 0.202. However, the highest value of 0.713 was observed for compatibility risk. Values for others lied between these two.

In case of testing data the system produced much better and closer to real estimates with the minimum error value as 0.196 for communication risk and highest error value as 0.582 for compatibility risk.

5.3 MAPE

Mean Absolute Percentage Error (MAPE) refers to the measure of accuracy for continuous output. It calculates the percentage error for the difference of actual and predicted values of the model.

Out of various formulae available for MAPE, the one used here is mentioned below:-

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - \overline{x_i}}{x_i}\right| \times 100\%$$ Equation 5.2

where n refers to the total number of data points and i and j refer to observed and predicted values respectively.

MAPE for various risk categories was calculated for both training and testing. For training, the lowest error was found to be for bleeding edge technology risk and it was 0.119. However the highest values 0.713 was observed for compatibility risk. Values for others lied between these two extremes. In case of testing data the system produced much better and closer to real estimates with the minimum error value as 0.148 for communication risk and highest error value as 0.502 for integration issues risk.

5.4 Correlation Coefficient R

Correlation Coefficient (R) is a degree to indicate how much two separate series of values are related to. Value of R varies from -1 to +1, however, if R value near to 1 indicates strong positive correlation between the values reflecting the consistency and closer relationship.

Formula for correlation coefficient is mentioned below:

$$R = \frac{\sum_{i=1}^{n}\left(x_{obs,i} - x_{obs}\right).\left(x_{\mathrm{model},i} - x_{\mathrm{model}}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_{obs,i} - x_{obs}\right)^2 . \sum_{i-1}^{n}\left(x_{\mathrm{model},i} - x_{\mathrm{model}}\right)^2}}$$ Equation 5.3

Correlation coefficient r for the model was calculated after the completion of training phase. Since the values of r were found to be 0.925 and 0.985 for training and testing respectively, it shows the strong positive correlation revealing the consistency of results.

Table 5.1: Results acquired after completion of training and testing of the proposed model

RISKS TYPES	MFs	RMSE		MAPE %	
		Training	**Testing**	**Training**	**Testing**
Communication Risk	Trimf	0.202	0.196	0.169	0.148
	Guass	0.202	0.196	0.169	0.148
	Trapez	0.202	0.196	0.169	0.148
Compatibility Risk	Trimf	0.7137	0.582	0.7131	0.492
	Guass	0.7137	0.582	0.7131	0.492
	Trapez	0.7137	0.582	0.7131	0.492
Bleeding Edge Tech Risk	Trimf	0.384	0.291	0.119	0.201
	Guass	0.384	0.291	0.119	0.201
	Trapez	0.384	0.291	0.119	0.201
Integration Issues	Trimf	0.699	0.519	0.664	0.502
	Guass	0.699	0.519	0.664	0.502
	Trapez	0.699	0.519	0.664	0.502
Wage Inflation	Trimf	0.338	0.216	0.323	0.194
	Guass	0.338	0.216	0.323	0.194
	Trapez	0.338	0.216	0.323	0.194
Staff Turnover	Trimf	0.567	0.551	0.521	0.541
	Guass	0.567	0.551	0.521	0.541
	Trapez	0.567	0.551	0.521	0.541
Insufficient Competencies	Trimf	0.431	0.388	0.399	0.379
	Guass	0.431	0.388	0.399	0.379
	Trapez	0.431	0.388	0.399	0.379
Temporal Risks	Trimf	0.526	0.401	0.499	0.381
	Guass	0.526	0.401	0.499	0.381
	Trapez	0.526	0.401	0.499	0.381

Average Errors

Average error for each of the mentioned risk categories are calculated below and are further depicted in the Graph as well.

$$AverageError = \frac{Trimf + Guass + Trapez}{3}$$ Equation 5.4

Putting values for all Risks (including both MAPE and RMSE errors) in the above formula we can acquire average error and fol graphs can be acquired by plotting average error on y-axix and Risk Categories on x-axis:-

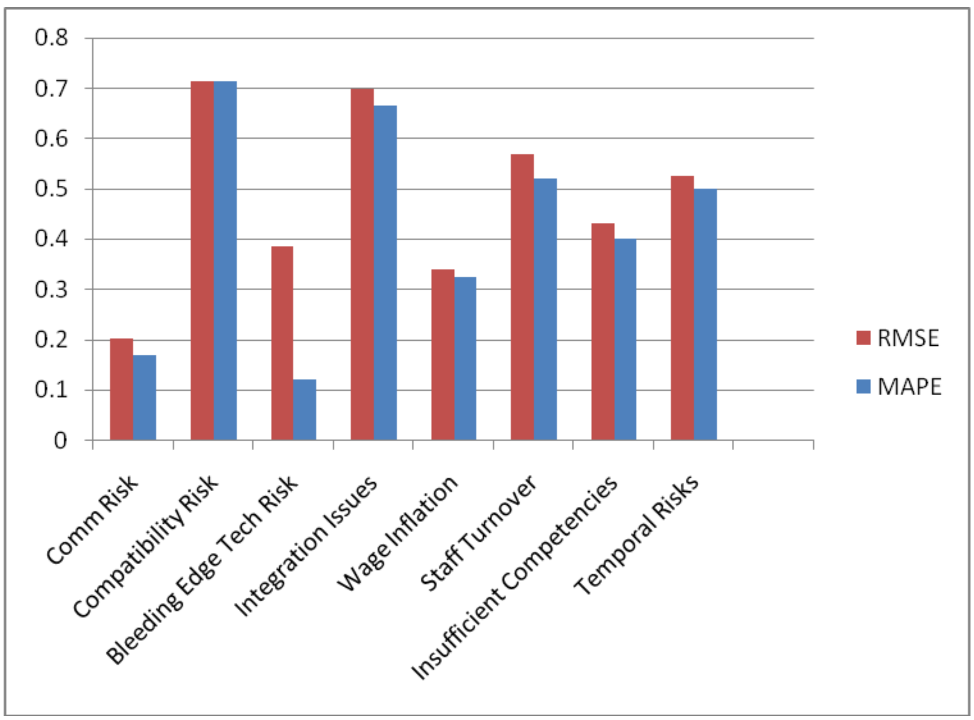

Figure 5.1: Graphical Representation for RMSE and MAPE for Training

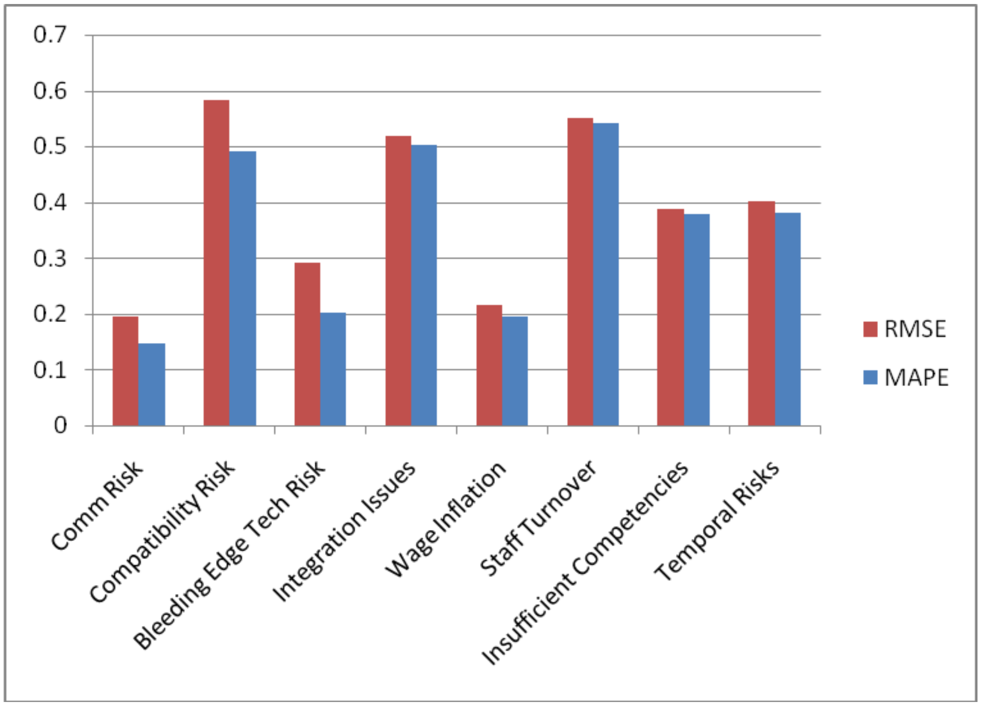

Figure 5.2: Graphical Representation for RMSE and MAPE for Testing

The results show that Soft Computing based framework for Risk assessment possesses the desired potential for evaluation of various categories of Risks in GSE and it can be concluded that system produces much better and closer to real estimates. Keeping in view the results obtained after completion of training and testing phase, it can be finally concluded that model seems to be promising and successful to be adopted and can be productive for project managers to calculate the risk they want to be predicted for their project. Therefore, approach can be adopted and programmed accordingly in order to fulfill the needs of project managers of GSE projects.

5.5 Comparison of Results with Existing Systems

Since, the proposed system itself is the Novel in its nature and addition in the field of Risk Assessment in GSE; therefore no existing systems are currently present which utilize soft computing platform for acquiring the same purpose. However, Marta Takacs has proposed a fuzzy based approach for risk calculations in the article: "Soft Computing-Based Risk Management - Fuzzy, Hierarchical Structured Decision-Making System" [23].

A comparison has been drawn on the basis of results acquires from implementing Fuzzy Inference System (FIS) for the subject purpose and results obtained from proposed ANFIS framework implementation, which is elaborated below in the table 5.3:-

Table 5.3: Results acquired after comparison of the proposed ANFIS model with FIS based model

RISKS TYPES	MFs	ANFIS based Proposed Approach				FIS Approach	
		RMSE		MAPE %			
		Training	Testing	Training	Testing	RMSE	MAPE %
Communication Risk	Trimf	0.202	0.196	0.169	0.148	2.443	2.310
	Guass	0.202	0.196	0.169	0.148	2.443	2.310
	Trapez	0.202	0.196	0.169	0.148	2.443	2.310
Compatibility Risk	Trimf	0.7137	0.582	0.7131	0.492	3.798	3.548
	Guass	0.7137	0.582	0.7131	0.492	3.798	3.548
	Trapez	0.7137	0.582	0.7131	0.492	3.798	3.548
Bleeding Edge Tech Risk	Trimf	0.384	0.291	0.119	0.201	2.884	2.423
	Guass	0.384	0.291	0.119	0.201	2.884	2.423
	Trapez	0.384	0.291	0.119	0.201	2.884	2.423
Integration Issues	Trimf	0.699	0.519	0.664	0.502	3.100	2.864
	Guass	0.699	0.519	0.664	0.502	3.100	2.864
	Trapez	0.699	0.519	0.664	0.502	3.100	2.864
Wage Inflation	Trimf	0.338	0.216	0.323	0.194	2.445	2.296
	Guass	0.338	0.216	0.323	0.194	2.445	2.296
	Trapez	0.338	0.216	0.323	0.194	2.445	2.296
Staff Turnover	Trimf	0.567	0.551	0.521	0.541	2.994	2.697
	Guass	0.567	0.551	0.521	0.541	2.994	2.697
	Trapez	0.567	0.551	0.521	0.541	2.994	2.697
Insufficient Competencies	Trimf	0.431	0.388	0.399	0.379	2.751	2.601
	Guass	0.431	0.388	0.399	0.379	2.751	2.601
	Trapez	0.431	0.388	0.399	0.379	2.751	2.601
Temporal Risks	Trimf	0.526	0.401	0.499	0.381	2.928	2.800
	Guass	0.526	0.401	0.499	0.381	2.928	2.800
	Trapez	0.526	0.401	0.499	0.381	2.928	2.800

Table 5.3 indicates the remarkable difference between accuracy of both risk assessment frameworks; however it can be clearly drawn from comparing the results columns that proposed ANFIS approach possesses marginal advantage over the other one. Therefore, these results are a

clear evidence for accuracy of the proposed system. Graphical Representation of the results is shown below:-

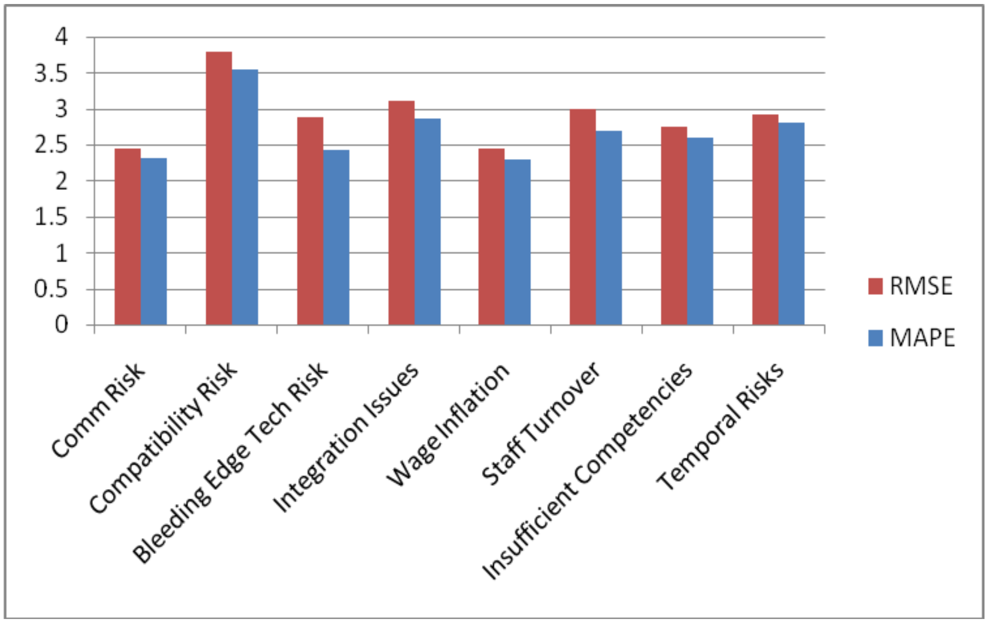

Figure 5.3:Graphical Representation for RMSE and MAPE for FIS

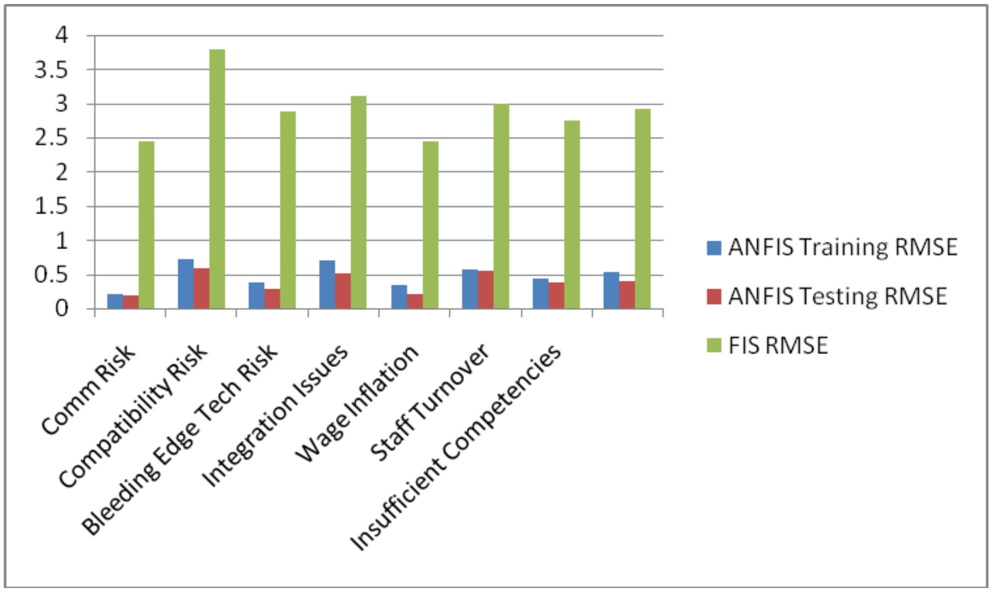

Figure 5.4: Graphical Representation for RMSE ANFIS and FIS

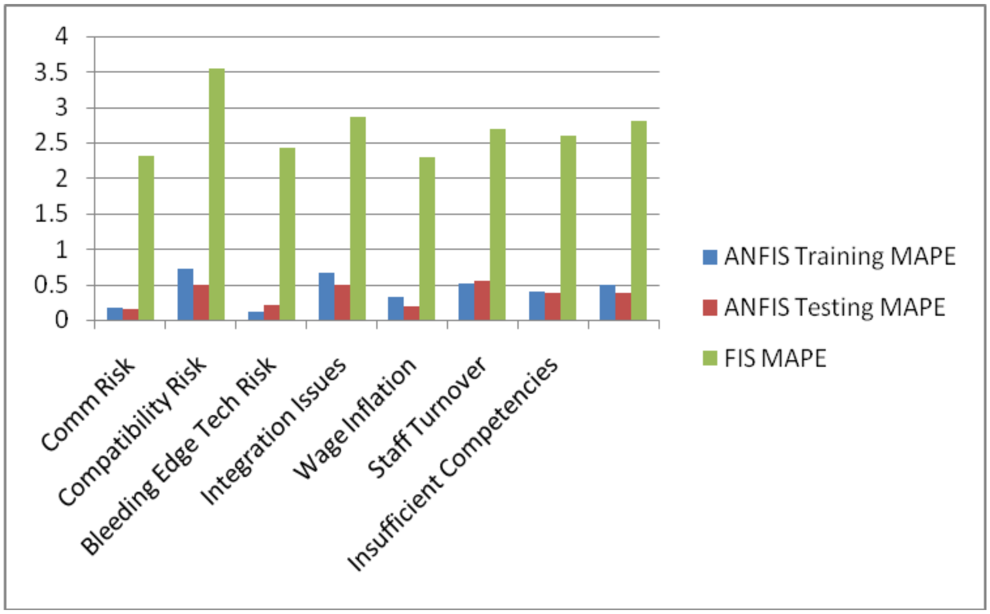

Figure5.5: Graphical Representation for MAPE for ANFIS and FIS

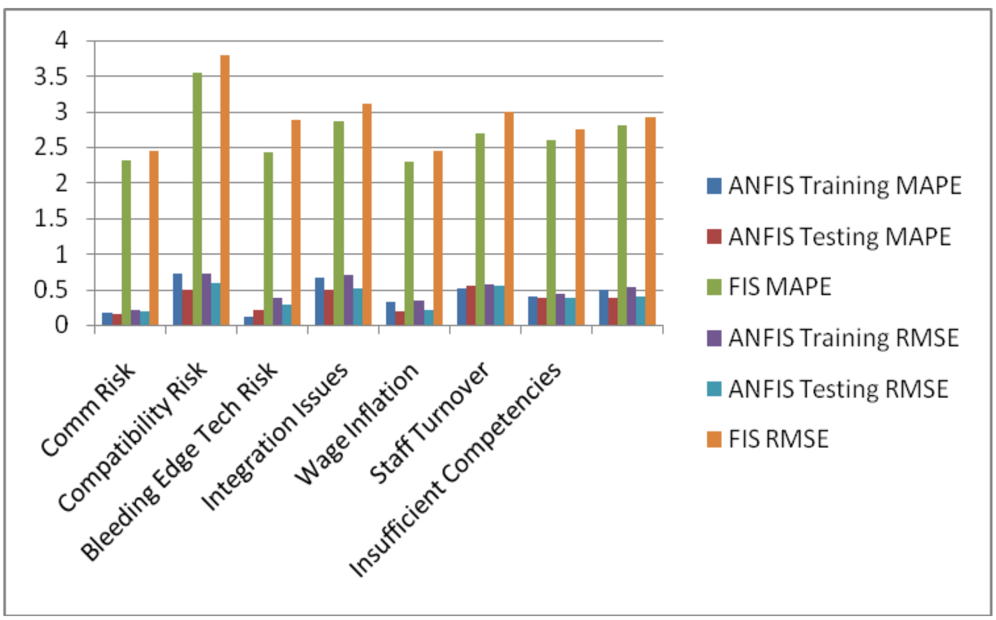

Figure 5.6: Graphical Representation for RMSE and MAPE for ANFIS and FIS

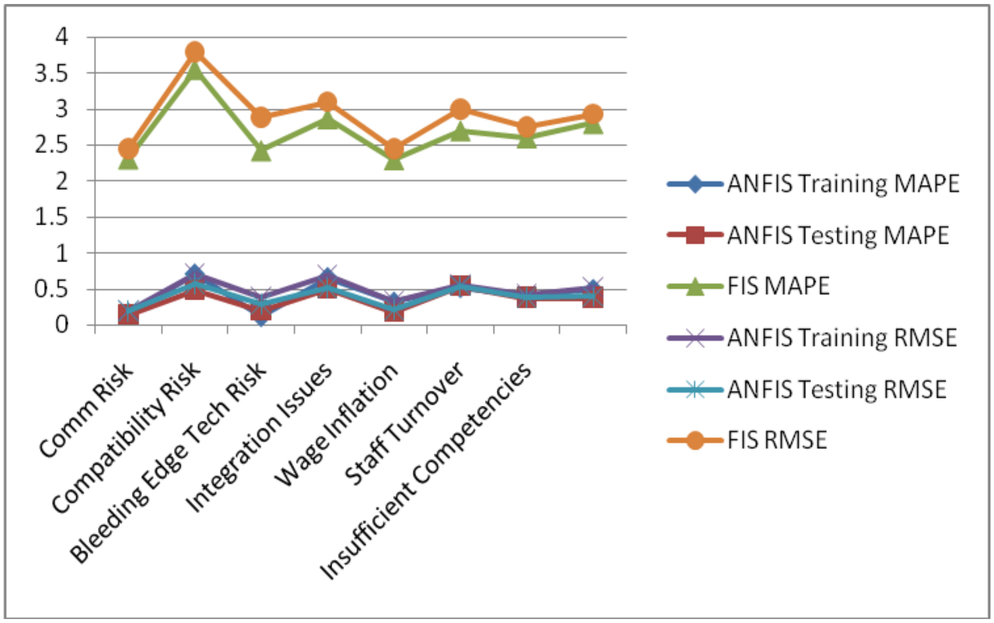

Figure 5.7: Graphical Representation for RMSE and MAPE for ANFIS and FIS in form of Line Graph

5.6 Summary

Proposed framework is evaluated on the bases of trained data and acquired results are found to be satisfactory. Errors are calculated using RMSE and MAPE. Obtained results are then compared with FIS system and it is observed that ANFIS produces much better results than FIS using same parameters.

CHAPTER 6: CONCLUSIONS AND FUTURE WORK

In this thesis we presented a comprehensive framework for risk assessment in GSE. Since it was concluded from literature review of various research papers published in IEEE and many other journals, conference papers and proceeding of many conferences, books published on subject matter, experts reviews and related resources that Global Software Engineering, which is not only an emerging trend but paradigm for expansion of IT/ Software Industry, is still far behind as a researched element. Though, many researchers have put up their efforts to bring forward the static and dynamic aspects of Global Software Engineering, still there is lot more to explore. Even project management aspects of the Software Engineering itself are unknown. Foregoing, Project Management in specifically Global Software Engineering become as rare area being researched over. Consequently giving rise to the scarcity of research in domain of risk assessment in Global Software Engineering. Therefore, research work was kept focused on risk assessment in Global Software Engineering. Risk assessment is a rarely researched but critical domain in various other disciplines as well but it can be disastrous if not handled in time and properly due to the peculiar nature of software and tough procedure of its management especially when it comes to global software engineering. In this research a framework is proposed for risk assessment in global software engineering which is itself a novel proposed framework since no detailed framework is found already in the field. The framework includes a complete taxonomy of prominent GSE risks. It also incorporates a Soft Computing approach applying ANFIS for assessment of GSE in context of any given project. Experimental results show that Soft Computing based framework for Risk assessment possesses the desired potential for evaluation of various categories of Risks in GSE produced much better and closer to real estimates. The model seems to be promising and successful to be adopted and can be productive for project managers to calculate the risk they want to be predicted for their project. Therefore, the approach can fulfill the needs of project managers of GSE projects.

Evolution of novel ideas is always welcomed and required in the field of Software Engineering. Since GSE is currently serving as an overwhelming discipline in software development; room is open for further research in the same field in the light of unending challenges faced by the project managers. Proposed framework may be enhanced by various dimensions and few are discussed in this chapter. By incorporating all these suggestions a complete package for risk management in GSE can be generated.

6.1. Cause and Effect Calculation

Cause and Effect Calculations to provide details of interdependencies of the risks and input values so that a scenario can be sketched to predict importance of each risk with relevance to others in for certain set of input parameters can be thought of for implementation. However, in such case, survey has to be conducted at huge level for collection of data. Then, test cases need to be developed and verified against collected data for establishing cause and effect relationship between various risks and input parameters.

6.2. Interdependencies Summary

Summarizing interdependencies of the risks for providing the criticality of the risk in the form of an interdependencies chart can be worked out. In such chart, interdependencies arising from cause and effect calculation can be summarized providing basis for analyzing interdependencies of risks within one another and risks arising from other risks.

6.3. Risk Mitigation Strategies

As a next step to above mentioned research work, mitigation strategies can be incorporated in application to overcome the risks. However, extremely detailed survey as well as team level analysis is required to be done in order to implement that factor.

GLOSSARY

ANFIS	Adaptive Neuro Fuzzy Inference System
FIS	Fuzzy Inference System
GSD	Global Software Development
GSE	Global Software Engineering
MAPE	Mean Absolute Percentage Error
NN	Neural Networks
RMSE	Root Mean Square Error

REFERENCES

[1] Christof Ebert; 2012, "Global Software and IT: A Guide to Distributed Development, Projects, and Outsourcing"

[2] Muhammad Ali Babar and Christian Lescher; 2013, "Global Software Engineering: Identifying Challenges is Important and Providing Solutions is Even Better" International Conference on Global Software Engineering (ICGSE), IEEE, Guest Editorial.

[3] Sami ul Haq, M. Naeem Ahmed Khan and Muhammad Tariq; 2011, "The Context of Global Software Development: Challenges, Best Practices and Benefits"; Information Management and Business Review, Vol. 3, No. 4, pp. 193-197, ISSN 2220-3796

[4] Christof Ebert, Bvs Krishna Murthy, Namo Narayan Jha; 2008 , "Managing Risks in Global Software Engineering: Principles and Practices"; IEEE International Conference on Global Software Engineering

[5] Rafael Prikladnicki, Jorge Luis Nicolas Audy and Roberto Evaristo;2004, "Global software development in practice lessons learned"

[6] Darja Smite, Claes Wohlin, Tony Gorschek, Robert Feldt; 2010, "Empirical Evidence in Global Software Engineering: A Systematic Review"; Journal of Empirical Software Engineering

[7] Muhammad Ali Babar and Anssoreh Zahedi; 2012, "Global Software Development: A Review of the State-Of-The-Art (2007 – 2011)"; 2012 IT University Technical Report Series

[8] Bilal Raza, Stephen G. MacDonell and Tony Clear; 2013, "Topics and treatments in global software engineering research - a systematic snapshot"; 8th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE2013), pp.85 – 96

[9] Paivi Parviainen, Maarit Tihinen; 2011, "Knowledge related challenges and solutions in GSD: Expert systems"; The Journal of Knowledge Engineering

[10] Indira Nurdiani, Ronald Jabangwey, Darja Smitey, Daniela Damian; 2011, "Risk Identification and Risk Mitigation Instruments for Global Software Development: Systematic Review and Survey Results"; 6th IEEE International Conference on Global Software Engineering

[11] Darja Šmite, Juris Borzovs; 2008, "Managing Uncertainty in Globally Distributed Software Development Projects"

[12] Qadeem Khan, Shahbaz Ghayyur; 2010, "Software Risks and Mitigation in Global Software Development"; Journal of Theoretical and Applied Information Technology, Pp. 58-69

[13] James D. Herbsleb; 2007, "Global Software Engineering: The Future of Socio-technical Coordination"; 29th Int. Conference on Software Engineering

[14] Sarah Beecham, John Noll, Ita Richardson Lero Deepak Dhungana; 2011, "A Decision Support System for Global Software Development"; 6th IEEE International Conference on Global Software Engineering

[15] Christer Magnusson, Sung-Chun Chou; 2010, "Risk and Compliance Management Framework for Outsourced Global Software Development", International Conference on Global Software Engineering

[16] Ansgar Lamersdorf, Alicia Fernández-del Viso Torre, Markus Heinz, Jürgen Münch, Carlos Rebate Sánchez and Dieter Rombach; 2010, "A Rule-based Model for Customized Risk Identification in Distributed Software Development Projects"; International Conference on Global Software Engineering, pp.209-218

[17] Ronald Jabangwey, Darja Smitey and Daniela Damian; 2011, "Risk Identification and Risk Mitigation Instruments for Global Software Development: Systematic Review and Survey Results"

[18] Marta Takacs; 2011, "Soft Computing-Based Risk Management - Fuzzy, Hierarchical Structured Decision-Making System"; Risk Management Trends, Prof. Giancarlo Nota (Ed.), ISBN: 978-953-307-314-9

[19] Lofti A. Zadeh; 1965, "Fuzzy Sets"; Information and Control

[20] Lotfi A. Zadeh; 1994, "Fuzzy Logic Neural Networks and Soft Computing"; Communications of the ACM, Vol 37. No.3

[21] Isabel L. Nunes and Mário Simões-Marques; 2012, "Applications of Fuzzy Logic in Risk Assessment - The RA_X Case 2012 Fuzzy Inference System - Theory and Applications", 1st Ed

[22] O. Senvar, G. Ulutagey and E. Taranolu Bekar; 2013, "Forecasting Oil Consumption of Turkey using Adaptive Neuro Fuzzy Inference System"; Intelligent System and Decision Making for Risk Analysis and Crisis Response, ISBN: 978-1-138-00019-3

[23] Marta Takacs (2011). Soft Computing-Based Risk Management - Fuzzy, Hierarchical Structured Decision-Making System, Professor Glancario Nota (Ed.), ISBN: 978-953-307-314-9, InTech.