A Predictive Machine Learning-Driven Approach to Requirements Engineering for Satisfying Stakeholder Needs

Submitted By

MD. NURNABI FAHIM

ID: 221-35-1049

Department of Software Engineering

Daffodil International University

Supervised By

A.H.M Shahariar Parvez

Associate Professor

Department Of Software Engineering

Daffodil International University

Bachelor of Science

DAFFODIL INTERNATIONAL UNIVERSITY

APPROVAL

This thesis titled on "**A Predictive Machine Learning-Driven Approach to Requrements Engineering for Satisfying Stakeholder Needs**", submitted by **MD.NURNABI FAHIM (ID: 221-35-1049)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

Dr. S M Hasan Mahmud — **Chairman**
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

A.H.M Shahariar Parvez — **Internal Examiner 1**
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Tapushe Rabaya Toma — **Internal Examiner 2**
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Khalid Been md. Badruzzaman Biplob — **Internal Examiner 3**
Lecturer (Senior Scale)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Dr. Md Sazzadur Rahman — **External Examiner**
Professor
Institute of Information technology
Jahangirnagar University, Bangladesh

A Predictive Machine Learning-Driven Approach to Requirements Engineering for Satisfying Stakeholder Needs

MD. NURNABI FAHIM

Bachelor of Science

DAFFODIL INTERNATIONAL UNIVERSITY

COPYRIGHT DECLARATION

DECLARATION OF THESIS AND COPYRIGHT

Author’s Full Name	:	MD. NURNABI FAHIM
Date of Birth	:	21.09.2003
Title	:	A Predictive Machine Learning-Driven Approach to Requirements Engineering for Satisfying Stakeholder Needs
Academic Session	: 2020-21	

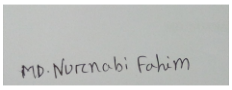
I declare that this thesis is classified as:

☐	CONFIDENTIAL	(Contains confidential information under the Official Secret Act 1997) *
☐	RESTRICTED	(Contains restricted information as specified by the organization where research was done) *
☐	OPEN ACCESS	I agree that my thesis to be published as online open access (Full Text)

I acknowledge that Daffodil International University reserves the following rights:

1. The Thesis is the Property of Daffodil International University.
2. The Library of Daffodil International University has the right to make copies of the thesis for the purpose of research only.
3. The Library of Daffodil International University has the right to make copies of the thesis for academic exchange.

Certified by:



(Student’s Signature)

221-35-1049

Student ID
Date: 23-12-2025

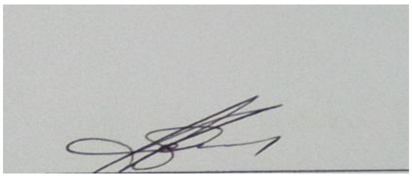
(Supervisor’s Signature)

A.H.M Shahariar Parvez

Name of Supervisor
Date: 23-12-2025

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and, in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor of Science



____

(Supervisor's Signature)

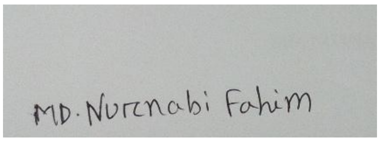
Full Name : A.H.M Shahriar Parvez

Position : Associate Professor

Date : 23-12-2025

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Daffodil International University or any other institution.



(Student's Signature)

Full Name : MD. NURNABI FAHIM

ID Number : 221-35-1049

Date : 23 December 2025

A PREDICTIVE MACHINE LEARNING-DRIVEN APPROACH TO REQUIREMENTS ENGINEERING FOR SATISFYING STAKEHOLDER NEEDS

MD. NURNABI FAHIM

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Science

Department of Software Engineering (Major in Software Engineering)

DAFFODIL INTERNATIONAL UNIVERSITY

December 2025

ACKNOWLEDGEMENTS

First, I would like to express my deepest gratitude to Almighty Allah for providing me with the strength, patience to complete this thesis research successfully. I would like to extend my sincere thanks to my supervisor Mr. A.H.M Shahriar Parvez for their guidance, feedback and continuous support throughout the entire of this Study. Their expertise and encouragement have been instrumental in shaping the quality and direction of this Thesis. My heartfelt thanks go to the Department of Software Engineering, Daffodil International University-DIU for providing the academic environment and necessary resources to carry out this work effectively. Lastly, I would like to express my profound love and gratitude to my family, especially my parents, for their support and understanding throughout my academic life.

DEDICATION

This thesis dedicated to my beloved parents, whose endless love, prayers, and encouragement have been my great source of strength and inspiration in my life. To my teachers and mentors, who guided me with patience and passion for learning.

ABSTRACT

Requirement engineering (RE) plays a crucial role in the success of software development projects as it ensures that stakeholder needs are met. So that ambiguity, conflict, and incomplete assessments of stakeholder satisfaction are common challenges faced by professional real estate processes. In order to overcome these limitations this study recommended a predictive machine learning based approach to enhance decision making and improve stakeholder satisfaction in the Requirement Engineering (RE) process. Structured surveys and requirement documentation were utilized to gather data from over 1000 responds and the representation was diverse stakeholder groups. Decision trees, Random Forest, Support Vector Machine and Neural Network were used to predict the stakeholder happiness levels are determined by both requirement attributes and project parameters. The suggested technique combines data pre-processing feature selection model training validation and performance to ensure accuracy. Based on the experimental data the machine learning based requirements engineering model has a significant impact on prediction accuracy and gives valuable insights into factors that influence stakeholder satisfaction. Intelligence automation can be utilized to optimize requirement periodization as highlighted by the study enhancing communication between stakeholders and development teams is crucial to reducing projects risks. The research contributes to the growing field of data driven software engineering by introducing a scalable framework for optimizing requirements engineering (RE) by utilizing predictive analysis a design that is focused on stakeholder needs

Keywords: Requirement Engineering, Machine Learning, Predictive Model, Stakeholder Satisfaction, Software Development, Data-driven Approach.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background to the Study

The software lifecycle (SDLC) is centered on requirements engineering (RE) where we determine managing and documenting requirements for system or application [1]. Project success and stakeholder satisfaction are directly impact by the quality of this process. So that manual analysis and subjective decision often play a major role in traditional requirements engineering methods. Inconsistencies incomplete requirements and communication gaps between developers and stakeholders are possible consequences of this [2]. Predictive analysis and automated decision making have been enabled by the emergence of machine learning in recent years which has transformed several fields [3]. The potential for analyzing large datasets of stakeholder feedback is opened up by using machine learning in requirements engineering project metrics. documentation of requirements is used to predict satisfaction levels and the possibility of projects success [4] [5]. The process effectiveness is enhanced by combining data analytics with requirements engineering through the combination of predictive machine learning and data analytics in this study enhance stakeholder satisfaction. Improve the organization requirements which will ultimately lead to more efficient software development outcomes.

1.2 Definitions of Terms Used in the Study

Software requirements are elicited analysis documented and validated through the structured process of requirements engineering [1]. Machine learning is a part or artificial intelligence that allows system in learning and enhancing from experience without the need for complicated programming [3]. To satisfy stakeholder the delivered software product must meet their needs and expectations [6]. In machine learning model feature selection involves recognizing the variable that contribute to predicting accuracy the most [7]. Evaluation metrics such as accuracy precision and recall are used in model validation to assess the performance of a machine learning model [8]. Software requirement refer functional and non-functional expectation that present what a software system should accomplished and how it should behave. These requirements are structured activities of the requirements engineering process for analysed document and validate it. Development team and stakeholder share a common understanding of the system goal by clear and accurate requirements. On the other hand, Requirement Engineering (RE) is a systematic discipline that focus on identifying specifying and maintaining requirement by the software development life cycle. Requirement elicitation requirement analysis and specification validation management effective RE improves project success all of them are included in software development life cycle. It ensures project success by minimizing misunderstanding and preventing costly changes later in development. Another is Machine learning ML is a subfield of artificial intelligence AI that enables computer to learn pattern from data and performance improving without exploit programming. It analyses historical data and adapt automatically to new information. By prediction classification and decision tree support task making them high effective. Another is stakeholder satisfaction that is ensure there needs by the delivered software product. User goal solves problem real world and provide a reliable usable and valuable experience that aligns and indicates high stakeholder satisfaction. Future selection for identifying most relevant variables that contributes predictive performance for the machine learning model. By selecting the right features improve model accuracy reduced and decreases overfitting and computational cost in the result. Accuracy precision recall and F1 score are used in evaluation metrics for assess the performance of the machine learning model. By the used of the metrics that quantify how well a model predicts or classifies data are essential for comparing different algorithm and validating predictive result.in this study ML based predictive model are used to find which factors influenced stakeholder satisfaction in software project.

1.3 Statement of the Problem

Although requirements engineering is vital many software projects fails because they lack adequate requirements analysis and stakeholder involvement [2] [6]. Predicting potential satisfaction issues before executions is often not possible with conventional methods. A predictive framework that is based on data is necessary to identify requirements related elements that impact stakeholder satisfaction in the project life cycle. To address that gap this study utilize a predictive model that is based on machine learning. Although requirement engineering is critical component for the successful software developments so that they fail in many projects. Because of conventional methods do not provide real time insights into pattern hidden within large volume of stakeholder data they face potential stakeholder satisfaction issue the during the early stages of a project is challenging. Becoming increasingly complex relying solely on manual analysis is no longer sufficient with modern software system.

1.4 Research Objectives

General Objective

To Develop a machine learning-driven framework for predicting stakeholder satisfaction in requirement engineering

Special Objective

To analyse stakeholder satisfaction using machine learning models based on usability performance and requirement quality

1.5 Significance of the Study

Both academic and industrial fields can benefit from this research. The academic study demonstrates how predictive models can optimize requirement engineering practices thereby improving the field of data driven software engineering. In actually it offers project managers and software engineer a decision support tool that can predict stakeholder satisfaction and improve requirements [5] [8]. From an academic perspective the study contributes to the evolving fields of data driven software engineering by the demonstrating how predictive machine learning models can enhance requirement engineering process. Machine learning technique can identify critical requirements related factors evaluate stakeholder priorities and forecast satisfaction levels with improved accuracy. Predictive analysis integrating into requirements engineering domain the study encourage further exploration of AI driven approaches and enriches existing theoretical framework in software development research. Term in practical the study offers benefit substantial to software engineers, organizations involved in software development and project managers. Decision support tool enables teams to detect potential requirements gaps, ambiguities, and conflict at early stage of the project lifecycle proposed by predictive model in this research acts. Apply this predictive model helps project stakeholder make informed decision reduced project risk and ensure that the delivered timely software aligns more closely with user expectation. Before implementation the ability to anticipate stakeholder, satisfaction issue can significantly minimize costly revisions rework and arise failures that often from poorly understood or mismanaged requirements. Moreover, the findings of this study can assist organizations improving communication by providing data support insights into their prevalence and concern with stakeholder. Institute based requirements analysis to a more measurable and systematic method approach this encourages a shift from this. The result can improve project panning, higher end user acceptance, better allocation of resources, enhanced software quality. Overall, the study bridges the gap between traditional requirements engineering method and capabilities of modern machine learning techniques. Provide scalable objective and reliable way for support requirement related decision making. Contributing to more successful software projects and higher levels of stakeholder satisfaction.

1.6 Research Questions

This study is process of guided by a set of core questions of research aimed at understanding how machine learning can enhance the requirements engineering process and improve satisfaction of stakeholder. The following four questions form the foundation of the investigation:

1. What is the effective way to apply machine learning in the requirements engineering process to predict stakeholder satisfaction?

   In this question that seeks to explore how data driven approaches can complement requirement analysis method. Models and data preprocessing techniques required to incorporate machine learning into the early stage of software development process that is ensure it examines the steps.

2. What requirements or projects factors have the most impact on the satisfaction of stakeholder?

   That question identifies the key feature that significantly affect satisfaction outcomes. Understanding these factors allows for better prioritization and decision making for requirement engineering.

3. Which predictive model is the most accurate when it comes to forecasting satisfaction outcomes?

   That question compares multiple model prediction such as Decision tree, Logistic Regression, Random Forest, SVM, and Neural Network for determine which model performs best based on metrics on evaluation such as accuracy, F1 score precision, recall and important feature.

4. How can the insights have generated from the machine learning model support improvement in the workflow of requirements engineering process?

   This research question strengths the practical of the study examine how the outcomes can guide project managers and identifying risk requirement analysis improving communication and enhancing engagement stakeholder.

1.7 Scope of the study

Decision Tree Random Forest and Support Vector Machine are among the machine learning algorithm used in this study and evaluating data collected from over 1000 survey respondents with Neural Network. Research is limited to software development projects and does not extend beyond other engineering disciplines. The research employs multiple supervised machine learning algorithm such as Decision tree, Random Forest, Neural network, Support vector machine SVM to analyze and predict stakeholder satisfaction based on the collected survey dataset. More than 1000 survey responses from stakeholder involved in various software development projects serve as the primary data source for ensuring a substantial and diverse sample for validation and model training. Software development projects and associated requirements engineering practices confined this study. This study cannot extend to other engineering disciplines such as mechanical, civil an electrical engineering. The research considers only those factors related to requirement quality stakeholder communication, process challenges, project complexity that is influenced satisfaction levels in software projects. Also includes the evaluation of feature important, interpretation of predictive insights and model performance comparisons. However, that does not monitor real time project, implementation of automation requirement tools.

1.8 Assumptions of the Study

The information given by the survey respondents was both accurate and unbiased. Predictive analysis in RE can be achieve using the machine learning models used. Minimizing noise and bias can be achieve by using data preparation and feature selection methods. This study based on several foundation assumptions that guided the analysis of the result. All survey assumed that participants provided truthful accurate and unbiased information regarding the experience in software requirement and satisfaction stakeholder. Machine learning prediction depends on the authenticity of these responses. The Study assumes that the selected machine learning model’s algorithm Decision tree random forest support vector machine and neural network are capable for effectivity modeling the complex relationships between requirements factored and satisfaction stakeholder. These models can capture both linear and non-linear patterns inherit in the dataset. Prepossessing such as data cleaning, handling missing values, outlier reduction and normalization are adequate to minimize noise in the datasets. The models receive high quality inputs ultimately improving predictive performance prefer preprocessing ensure that.

1.9 Limitations of the Study

Data quality and model arrangement determine the accuracy of forecasts findings could not be the same for every software projects. The focus of the study is solely on stakeholder satisfaction and does not address other success metrics like cost or time efficiency. Accuracy of the machine learning predictions are heavily influenced by the quality of the collected data. Any biases and inaccuracies in the survey responses may reduce the precision of the models of forecasting. The study on self-reported data vulnerable to human error subjective interpretation and recall bias. The performance of the models depends on perimeter setting, training procedures and feature selection process technique. Unobserved interactions between variables may limit the predictive capability of the machine learning model. Not all algorithm well not equally perform for every type of datasets where result restricts the generalizability of result.

1.9.1 Limited Scope of the Evaluation Metrics

The study focusses on predicting satisfaction of stakeholder. Other factors are project success. That is critical factors. Such as defect rate, cost efficiency, schedule adherence, are not analyzed. So that the outcomes provide only partial view of overall software success of project.

1.9.2 Survey based Sampling Limitations

The datasets include responses from the over 1000 participants. May be not capture the full diversity of global software engineering practices. On the other hand, difference in organizational culture and project methodologies also stakeholder engagement process might influence outcomes beyond what is represented in the data sample.

1.10 Chapter Summary

This chapter provides an overview of the fundamental concepts, motivation and objectives underlying the study. This chapter introduced the important of the integrating machine learning techniques to improve stakeholder satisfaction prediction and overall project success in requirement engineering. Outline the research problem, significance, scope assumptions and limitations established a clear direction for the study. Also, the chapter present the research question that guide the investigations and clarified why predictive data driven approach are important or essential for addressing gaps in requirement engineering. Through the thesis definition of the key terms were also provided to ensure conceptual clarity. The purpose of this chapter is to present the background objectives and rationale for combining machine learning with requirements engineering. The proposed prediction framework methodology data analysis and evaluation are covers in the following chapter which serve as its basis. Next chapter build upon this foundation by presenting related literature, methodological framework, data collection approach, machine learning model used and procedures detailed analysis.

CHAPTER 2

Review Related Literature

2.1 Introduction

This chapter present a comprehensive review of existing literature to requirement engineering and machine learning also stakeholder satisfaction prediction. Examines the theoretical foundation key concepts and research finding that support machine learning techniques of integrations into the requirement engineering process. Purpose of the chapter is to established an academic context for the study. Analysing how requirement engineering practices have evolved and how machine learning has utilized in software engineering. Also, how predictive models contribute to improve project outcomes. Understanding the role of requirements quality engagement stakeholder and automated analysis tools influencing satisfaction levels in software projects. This chapter divided into two major sections. First section explores the theoretical framework that include fundamental concepts such as requirements elicitation document validation feature selection and predictive modelling. On the other hand, the second sections present the empirical and practical literature review previous studies and implement ML algorithm for classification requirement, traceability, defect prediction and satisfaction forecasting. Requirements Engineering and Machine learning theories concept and previous research studies are cover in this chapter and stakeholder satisfaction prediction. The study aims to comprehend how machine learning techniques have been applied in software engineering particularly in computerizing requirements analysis and evaluating satisfaction. There are two parts to this chapter which include the theoretical framework review and the practical review of the literature.

2.2 Theoretical Framework Review

The elicitation and validation of software requirements is guided by RE which is founded on several theoretical foundations. Van Lamsweerde[1] proposed a crucial initiative called goal-oriented requirements engineering. That is Stress the important of identifying high level system goals and turning them into feasible requirements while also ensuring they are in line with stakeholder requirements. And emphasizes that high level system goals are identified and refined into feasible requirements with adjustments made to meet the needs of stakeholder. To achieve system acceptance, it becomes necessary to consider multiple stakeholder viewpoints as highlighted by the stakeholder theory another related theory [2]. Iterative development and risk assessment can be theoretically support by the Spiral Model proposed by Boehm that incorporates continuous feedback. These frameworks promote the notion of continuous advancement in Re which aligns with the predictive and adaptive abilities machine.

2.2.1 Requirements Engineering Theories

A number of theoretical foundations guides the elicitation and validation of software requirements which form the basis of the requirements engine. The proposed oriented requirements engineering is one of the most important [1]. Orientated requirements prioritize creating high level system goals and fine tuning them into implementable requirements. Another theory related to stakeholder theory stress the important of taking into account various stakeholder viewpoint to achieve satisfaction [2]. Including ongoing stakeholder analysis Boehm's Spiral Model [3] provides theoretical foundation for iterative development and risk assessment stakeholder feedback.one of the most influential framework is Goal Oriented Requirement Engineering (GORE) that is identifying high level organizational refining into detailed implementable requirement [1]. According to the GORE principles system helps stakeholder articulate what they need and expect from the system. Refinement technique such as goal decomposition and obstacle analysis ensure that requirement is complete, consistent, and traceable. That highlight the important of system goal with expectations of stakeholder. Another relevant theoretical foundation is Stakeholder theory that is important of recognizing and incropted diverse stakeholder viewpoint throughout the requirement process. Together these theories established the foundation for modern requirement engineering practices.

2.2.2 Machine Learning Theories

The mathematical foundation for model training and induction is provident by machine learning which is based on statistical learning theory [4]. Without clear programming ML system can learn patterns from data and classify it. Machine learning is defined by Mitchell [5] as the process by which a computer program learns from experience (E) with respect to a class task (T) and a performance indicator (P) if its ability to perform task in T improves with experience E. The supervised unsupervised and reinforcement learning paradigms are key ML paradigms that are relevant to study of this. The application of Supervised learning which involves algorithm learning form labelled datasets is best achieved through supervised learning which reduced number of variables [6]. Such algorithm Decision Trees Random Forest Support Vector Machine and Neural Network commonly used frequently employed in these tasks [7] [8].

2.3 Theoretical Integration Data Driven RE Framework

The integration of data driven approaches with requirement engineering introduced modern framework that machine learning theories to enhance the accuracy, efficiency and reliability for decision making of requirement related. RE are elicited document analysed and validate primary through manual techniques such as workshop interview and document review. Thus, method suffer from ambiguity subject and inconsistency issue that use predictive data driven techniques. Uncover hidden pattern and generated predictive insights combines structed requirement information with machine learning models. A data driven framework of RE are combined to create ML theories resulting in the transformation of requirement data through surveys or document features for model training. Predictive analysis is support from this approach of hybrid and allows for continuous requirement refinement based on feedback from stakeholder [9]. The proposed prediction framework that enhanced RE by utilized ML based satisfaction prediction.

2.4 Machine Learning Application in Requirements Engineering

Machine learning technique have demonstrated by several amphoral studies of automating real estate process. Ferrari and colleagues [10] NLP to automatically classify requirements and detect of requirements leading to improve requirements. The accuracy of requirements was important by Rturi et al [11] employed ML algorithm to classify requirements based on inputs of stakeholder. Deep learning models for traceability link recover were Cleland Huang [12] investigation to use for perform significantly better than traditional methods. The studies show that ML can be incorporated into various RE from elicitation and analysis to validate to improve efficiency. Machine learning techniques have been increasingly applied in requirement engineering to automate manual processing such as elicitations analysis and validation of software requirements. Ferrari et al. (2018) [10] applied Neural Language processing NLP to automatically classify and detect ambiguities in requirement of textual in requirement engineering. That work shows the NLP driven techniques can identify unclear requirements leading to significant ML and NLP to support requirement documentation and early defect detection. Similarly, Raturi, Rathore and Kumar (2020) [11] employed various supervised Machin learning algorithm to classify functional and non-functional stakeholder input based on the requirements. There findings indicates that ML based classifier high accuracy achieved in ML based classifier and distinguishing requirements types, thereby reducing the risk in inconsistence requirement categorization. I another contribution Cleland Hung (2021) [12] investigate the use of deep learning models for requirement traceability link recovery. Deep learning models such as neural network outperformed these conventional methods by learning complex semantic relationships in requirement artifacts for improve traceability accuracy.

2.5 Predictive Modelling and Stakeholder Satisfaction

The increasing availability of project data has led to caused research on prediction modelling in software engineering to gain traction more popular. Malhotra [13] find 85% accuracy that ML model could predict software quality attributes and satisfaction of stakeholder. Sharma and another person Gupta used Random Forest and SVM algorithms for predicts project success and identify key requirements features that play and impact on it [14]. The finding of empirical confirmed that prediction analysis in enhancing stakeholder engagement and prioritizing requirements.

2.6 Gaps in the Existing Literature

ML can be applied for real time analytes the majority of most studies tend to focus on requirement classification or defect prediction despite the feasibility of applying ML. The integration of both qualitative and quantitative data from diver's stakeholder groups is rare in models. Furthermore, the development of a unified predictive framework that support RE decision utilizing real world data of early stage [11] [12] [15]. The need predict according to this research ML approach to RE is necessary because of this gap that RE improve of large scale of survey data to forecast predict satisfaction.

Table 2.1 Literature Summary Table

Title	Dataset	Methodology	Output	Performance	Limitation
NLP for Requirements Engineering	RE text corpora	NLP (POS tagging, parsing, entity extraction)	Techniques to improve requirements quality analysis	Good for ambiguity detection	Domain adaptation difficult; low accuracy on noisy RE texts
Applying ML Algorithms for Requirements Prediction	Software requirement repositories	Classification + prediction using ML algorithms	Defect prediction & requirement categorization	Competitive accuracy using ensemble models	Limited dataset size
Adaptive Requirements Engineering	Case studies	Adaptive RE, socio-technical modelling	Framework for dynamic requirements	Strong in volatile environments	Complex, requires domain expertise
Methodological Guidelines for RE Process Research	Literature-based	Process research frameworks	Guidelines for RE research	High relevance to empirical RE	Non-empirical
ML-based Approach for Requirements Classification	Labelled requirement datasets	ML classifiers (SVM, RF, Naïve Bayes)	Automated functional–non-functional classification	85–92% accuracy	Requires labelled training data; domain-specific

2.7 Chapter Summary

The chapter presented a comprehensive review of the theoretical and empirical foundation that support integration into ML and Requirement engineering. Examine the major theoretical framework such as GORE and stakeholder Theory. Also, iterative models like Boehm's Spiral Model all of them emphasize the importance structured requirement and stakeholder engagement continuously and risk-based development in RE. also discussed the fundamental principles of machine learning, model training, feature selection, evaluation metrics and predictive modelling. Thus, concepts developing intelligent system capable of analysing requirement data and stakeholder satisfaction. The Proposed predictive RE approach is based on the theoretical and empirical studies reviewed in this chapter. Major framework is discussed including GORE and Stakeholder Theory and ML foundation. While discussed the review and previous studies that demonstrated the potential of ML in RE. The study uses a Machin learning approach to address a research gap in prediction modelling for stakeholder satisfaction which was revealed in the literature based on methodology.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter presents the methodology adopted to investigate the effectiveness of a predictive machine learning driven approach for improving satisfaction of stakeholder with the requirements engineering process. Outlines the overall research designed explains how data was collected process and analyse and modelled to archive objective of the study. Describing the research design which integrated both quantities and qualitative techniques to ensure a comprehensiveness analysis factor of the stakeholder satisfaction. Also describing the research design which integrates both quantitative and qualitative techniques to ensure a comprehensive analysis of stakeholder satisfaction factor process. Define the population and sampling procedures detailing how survey participants were selected and how datasets are constructed with 1000 stakeholder responses. The methodology used to investigate the effectiveness of a predictive machine learning driven approach for improve stakeholder satisfaction in requirement engineering. The research design is outlined as well as population and sampling method for data collection technique and data analysis procedures. The Chapter discussing the limitations of the study and chapter concludes.

3.2 Proposed Methodology

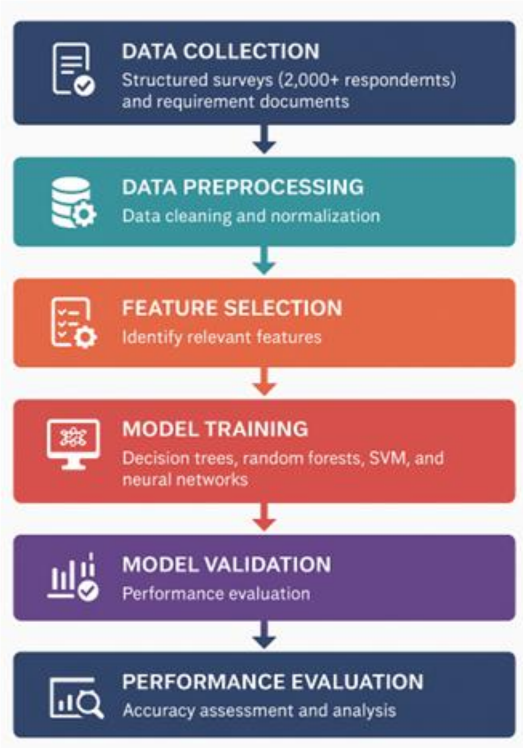


Fig 3: Proposed Methodology

The methodology illustrates in the diagram represent the complete workflow used for analysing stakeholder satisfaction and predicting outcomes by the using ML driven techniques. The process is divided into six major stages such as Data Collection, Data Pre-processing, Feature Selection, Model Training, Model Validation and Performance Evaluation. Each stage plays a critical role in ensuring the accuracy, reliability and quality of the predictive model. The purpose of the data collection is to gather comprehensive information about user perceptions, usability factors, system performance and satisfaction metrics. High quality data set6s can ML models generalize well. Data prepossessing include Removing duplicates, handling missing value, data normalization, encoding categorical variables. Feature selection identifies the most relevant and impactful such as ease to navigate, system responsiveness, task completion efficiency, feature completeness. Satisfaction method can correlation test are used to choose the top predictors and reduced overfitting improving training speed and increase model interpretable. Model training by Decision trees, random forest, support vector machine and neural network. Model validation involve testing on a holdout evaluation dataset, checking consistency across folds, measuring classification accuracy and stability. Performance evaluation by the metrics included accuracy, precision, recall, F1 score.

3.3 Research Design

This study a mixed method research design integrating both qualitative and quantitative approaches for obtain understanding of the research problem. According to the Creswell [1] mixed method design enhanced the richness by combining numerical data with insights of contextual. Analyse complex phenomenon form multiple perspectives. Both qualitative and quantitative approaches are used in a mixed method research design. By integrating numerical data in mixed method design to gain a better understanding for complex research according to Creswell [1]. To build and tarin machine learning model involves survey data quantitative component from stakeholder. Interview and focus group discussion are qualitative component and document review to contextualize quantitative findings and evaluate their validity for stakeholder perceptions. Previous RE and ML studies the successful use of hybrid methods to improve requirements analysis stakeholder engagement and accuracy prediction [2] [3].

3.4 Population and Sampling

The study focusses on software developers and projects managers analytes academic research and end users involved into software development projects. Output from these groups is what drivers' elicitation on requirements and satisfaction analysis which are the key stakeholder. The study order to analyse feedback from stakeholder and focused on organizations and education institutions that actively engage in the research goal of software engineering. Number of participants around 1000 were accessible population for ML based predictive modelling. The populations of the study consistence of key involved stakeholder in software development process including software developer, project manager, business analytics, academic research and end user. Play critical role elicitation in requirement and satisfaction evaluation so that making them essential participants for developing an ML based predictive framework method for requirement engineering.

3.5 Interviews

To obtain more information in deeper understanding of the experience and expectation of stakeholder conduct semi structured interviews. Through interviews and researchers to gain an understanding of the rationale behind stakeholder satisfaction levels and challenges requirements. In accordance with Kvale [6] semi structured interview flexible while still ensuring focused inquiry. Interview were held with projects manager top developer system analytics and academic software engineer's expert. The interview was recorded and take 20-30 minutes each of them. Interview was conduct with project managers, senior developer, system analysts, and academic experts in software engineering. For the stakeholder satisfaction and project outcomes these group possess critical perspective on requirements quality. Each interview session approximately 20-30 minutes with audio recorded with the participant constant for ensure accuracy during later transcription analysis.

3.6 Focus Group Discussions

Discussion on two focus group were arranged each with 6-8 participants. Exploring the effective ness of focus group lies in their ability to explore shared perceptions and generated interactive dialog among stakeholder is possible through focus group [7]. Focused on traditional RE process machine learning for improving RE and the perceptions of ambidexterity. The effective of focus group lies in their ability among participants to uncover shared perceptions and stimulate interactive dialogue. For enabling the exploration of diver's viewpoint with collecting setting. The discussed will happen of the three key areas such as Traditional requirement engineering process, The role of machine learning in improving RE activities and Stakeholder perceptions regarding ambidexterity. Particularly the balance between exploration and exploitation with RE practices. Provide valuable insights into challenges, expectations, and opportunities for predictive machine learning technique into RE process.

3.7 Observation

Non participant observation conducts during requirement collecting meeting with selected organizations. Research objectively examines real time behaviours communication pattern and documentation practices influencing the natural workflow. During requirement gathering meeting in specific organization and non-participant observation was conducted. RE quality can be identified by observing behaviours communication pattern and practices documentation. Stakeholder interaction common requirement conflicts and communication issues were all covered in observation notes for decision making process. Observation notes were recorded to capture how requirement were discussed, negotiated and finalized. The survey, interview, and focus group data by providing contextual understanding of issues in real world RE environments.

3.8 Document Review

A comprehensive document review was conduct to analyse various project artifacts including requirement specification, project report, meeting minutes. That document serves as essential sources of secondary data offering insights how requirement is elicited and document modified also validate the software development life cycle SDLC. According to Bowen document analysis is an effective for obtain and verifiable secondary information, particularly in organizational research context. [9] Analyse requirements specification and projects report meeting minutes and changes request forms for specify requirements. According to Bowen [9] document analysis offers reliable secondary data for development process. These documents provided insights that ML predictive model for helped identify feature. The information also determining features for the machine learning predictive model.

3.9 Quantitative Data Analysis

Quantitative data collated from the survey responses were analysed using Python based data science libraries such as Pandas, NumPy, Matplotlib, and Scikit. That tool ensures data cleaning, encoding, transformation, and preparation for machine learning modelling. The steps of including handling missing values, normalizing numerical attributes and converting categorical responses into machine readable formats using techniques such as Label Encoding and One Hot Encoding Python based libraries were utilized for the cleaning encoding and processing survey response used machine learning technique such as Logistic Regression Random Forest SVM Neural Network. Predict stakeholder satisfaction performs metrics included accuracy F1 Score precision and recall. Model performance was evaluated using standard classification metrics such as Accuracy for used the proportion of correct predictions, Precision for used evaluated the correctness of positive prediction, Recall used for assess the model sensitivity in identifying satisfied stakeholder and F1 score used for balanced precision and recall for more comprehensive model evaluation.

3.9.1 Equations

3.9.1.1 Logistic Regression Equation for predictive Model

The equation represents the probability that a stakeholder will be satisfied y = 1 several based on requirement related factors (x1, x2…. xn) Used the sigmoid function Logistic Regression convert a linear combination of features into probability of this. Where β^0 is the model bias. $\beta^1\beta^2\ldots\ldots\beta_{\mathrm{n}}$ are learned weights and x1…. xn are features.

$$P(Y=1|X)=1/\left(1+e^{-\left(\beta^0+\beta^1x^1+\beta^2x^2+\cdots+\beta_{\mathrm{n}}x_{\mathrm{n}}\right)}\right) \qquad 3.1$$

The output is between o and 1 making the model suitable for satisfaction of binary prediction (satisfied or dissatisfied).

3.9.1.2 Logistic Regression Loss Function

The function measures how well the logistic model fits the survey data. Where $\hat{y}_i$ is the predicted probability. y_i is the actual stakeholder response (1 is satisfied and o is not satisfied) and m is total samples.

$$L = -1/m\Sigma[y_i log(\hat{y}_i) + (1 - y_i) log(1 - \hat{y}_i)] \qquad 3.2$$

That goal of training is to minimize L so prediction match the actual satisfaction levels. That means Lower loss means higher prediction accuracy.

3.9.1.3 Gini Impurity

Gini impurity measures how mixed the stakeholder classes are a node of decision tree. Where pi is probability of a class. Lower Gini indicates the feature better separates satisfied vs dissatisfied stakeholder.

$$G = 1 - \Sigma p_i^2 \qquad 3.3$$

3.9.1.4 Information Gain

Information gain measures how much a feature improves prediction quality. Where H(s) overall entropy before splitting and H (Sv) are entropy after splitting using attribute A. High IG means the feature reduced significantly uncertainty in predicting satisfaction on stakeholder.

$$IG(S, A) = H(S) – \Sigma(|S_v|/|S|)H(S_v) \qquad 3.4$$

3.9.1.5 Random Forest Prediction

Random forest uses multiple decision tress and takes vote of majority. If most trees predict satisfaction the forest satisfied predicts. That reduced overfitting and increase reliability. That benefit requirement engineering because of different trees captures different requirement pattern.

$$\hat{y}_R F = mode(\hat{y}^1, \hat{y}^2, \dots, \hat{y}_k) \qquad 3.5$$

3.9.1.6 SVM Decision Function

Support vector machine separates satisfied vs dissatisfied stakeholder using dimensional boundary. Where $f(x) \geq 0$ is predictive satisfied and $f(x) < 0$ is a predictive dissatisfied.

$$f(x) = w \cdot x + b \qquad 3.6$$

3.9.1.7 SVM Optimizing Object

SVM tries to find maximum margin of hyperplane. Where minimizing $||w||^2$ ensures the boundary is as possible both classes. This improves generalization by the reduces misclassification. That helps identify separation clearest between satisfied and dissatisfied of stakeholder.

$$min\left(1/2\left||w|\right|^2\right) subjecttoy_{\mathrm{i}}(w \cdot x_{\mathrm{i}} + b) \geq 1 \qquad 3.7$$

3.9.1.8 Neural Network Forward Propagation

Neural networks learn complex non-linear patterns that simple models that cannot detect. Where a1 is hidden layer activation. σ is a predictive satisfaction probability. NN can detect subtle satisfaction pattern such as interaction effects into requirement clarity and communication and non-linear increases satisfaction because of iterative review cycles.

$$a^1 = \sigma(W^1 x + b^1) \qquad 3.8$$

$$\hat{y} = \sigma(W_2 a_1 + b_2)$$

3.9.1.9 Accuracy

Accuracy shows how many stakeholders satisfaction and prediction the model gets right out of the total. where TP is correctly predictive satisfaction. TN is correctly predicted dissatisfied. FP is predictive satisfied but actually satisfied. FN is predictive dissatisfied but actually satisfied. Useful but when insufficient alone if data is imbalanced.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \qquad 3.9$$

3.9.1.10 Precision, Recall and F1

Precision measures how reliable positive satisfaction prediction are high precision means when the model predicts satisfaction it is important for RE to avoid false assuming stakeholder are happy. On the other hand, Recall shows how many actual satisfaction stakeholders correctly identifies the model. If recall is low many satisfied users are being classified incorrectly it is dangerous for project decision making.

$$Precision = TP/(TP + FP) \qquad 3.10$$

$$Recall = TP / (TP + FN)$$

$$F1 = 2 * (Precision * Recall) / (Precision + Recall)$$

3.10 Qualitative Data Analysis

The qualitative data collected from the interview, focus group discussion and observation were analysed using Baraun and Clark's (2006) [12] Six phase analysis Framework. That method was selected because of it offers a flexible yet rigorous approach for identifying, interpreting and reporting recurring patterns with qualitative datasets. Braun and another person Clarkes six phase approach [12] were followed for used to analyse interview. Machine learning result were supplemented by themes used to find expectation of stakeholder.

That analysis followed the structured phases:

All transcript of interview focus group recording and observation notes read multiple times to gain and find in depth understanding of stakeholder experience and perceptions related to requirement engineering RE.

Segment of meaningful of text systematically coded to capture key ideas such as communication gaps, requirement ambiguity, satisfaction factors and challenges in traditional requirement engineering process. Generated code was grouped to form potential themes including such as stakeholder involvement, requirement clarity, tool support limitations also expectation for automation. Themes were refined through checking alignment with coded extracts and the overall dataset. Overlapping themes reorganized. All of them was clearly defined and explaining its relevance to stakeholder satisfaction and improvement of requirement engineering practices by machine learning support. Final themes were integrated with quantitative findings to prove a deeper contextual understanding. That qualitative helped explain certain machine learning predictors influenced satisfaction. Validate during machine learning model for the pattern discover. These insights enhanced the machine learning result, offering a more holistic understanding how predictive model work in requirement engineering process.

3.11 Limitations of the Study

This study is subject to several limitation that may influenced the interpretation and generalizability of the finding. The use of non-random and convinced based sampling restricts the repetitiveness of the sample. Participants were selected based on availability rather than probability sampling that is limited the extend to which the result can be generalized to all software development environment. Another is relied heavily on self-reported survey responses which are susceptible respondent bias. Maybe they provided socially desire answer or may have real their real experience accurately, potentially affecting the accuracy dataset used for ML model. Another is efficiency and predictive capability of ML models that depend on data quality feature selection and pre-processing techniques. By the noise and missing data in responses that influenced model performance. The study has some limitation such as non-random sampling that limits generalizability and self-reported survey data have respondent's bias. Data quality and feature selection are key factors in machine learning models. Satisfaction was limited by time constraints for longitudinal tracking of stakeholder. Also are time constraints. Satisfaction levels were measured single point in the rather than observing changes throughout the project lifecycle that is provided richer insights. Although multiple algorithms were evaluated so that the study was limited available computational resources and did not explore more complex.

3.12 Chapter Summary

This chapter presented a comprehensive overview of the methodology adopted to investigate the effectiveness of a machine learning driven predictive approaches framework for enhanced satisfaction stakeholder in RE. outline the research design which employed a mixed method approach both quantitative and qualitative strategies to ensure methodological rigor and depth of understanding. Population and sampling procedures was describing the key stakeholder groups such as software developer project manager and end user. Various data collection methos are including surveys, semi structured interview, focus group discussion and observation also document analysis for discussed highlighted multiple layered approach taken to capture diverse perspective. The research design sampling strategies data collection method and analytical technique used for research outline in the chapter. By the robust data collection using for predictive ML driven requirements engineering framework could be developed. In this next chapter describe result and analysis are presented and procedures. The following chapter provides details of the result, interpretations, and analytical procedures for both qualitative and quotative.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

This chapter present result and discussion about derived from the quantitative and qualitative analysis. Descriptive statistics machine learning prediction result and feature important analysis also qualitative findings of the purpose this chapter from open responses. Based on Survey data are collected 1000 from stakeholders they involve in software projects. Open ended responses are supported by quantitative result. The qualitative analysis starts with descriptive statistics to summarize stakeholder satisfaction. Machine learning model were then applied to predict overall satisfaction on user and identify the positive and negative factor of perceptions. Highlight the feature important analysis such as system responsiveness functionality completeness navigational ease and efficiency of task. To complete the result qualitative data, find from open ended survey responses deeper insights for the reason stakeholder satisfaction or dissatisfaction. It helps explain pattern observed in the predictive and statistical analysis. Both quantitative and qualitative result provide a holistic understanding of stakeholder perceptions This approach improvement priorities ultimately guiding recommendations for enhancing the software requirement engineering process.

4.2 Descriptive Statistic

Sample Size and overall satisfaction are present there

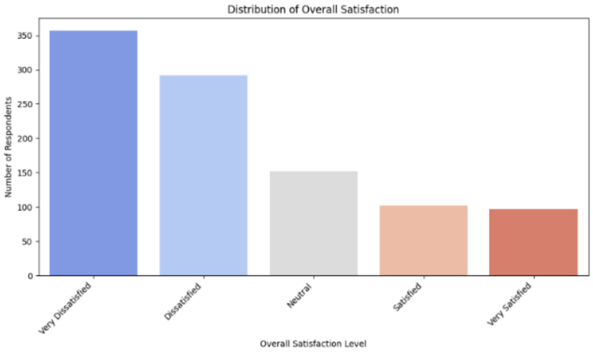


Fig 01: Overall Satisfaction Level

The bar chart titled Distribution of Overall Satisfaction illustrates how respondents rated the overall satisfaction in the requirement engineering RE process. The categories ranges from very dissatisfied to very satisfied comparison across different sentiment levels. Key observation of this highest dissatisfaction levels the largest group of respondents selected very dissatisfied making the most dominant category. The second largest group selected dissatisfied. So that indicates that a significant proportion of stakeholder are unhappy with the current RE practices. The suggested major issues in requirement clarity, communication, prioritization and implementation. Approximately 150 respondents chose neutral. This shows that portion of stakeholder neither strong approves nor disapproves of the process. Only around 100 respondents selected satisfied. A similar number selected very satisfied. The distribution is heavily skewed toward dissatisfaction which suggest stakeholder experience significant challenges with the requirement engineering process. May be issues with requirement elicitation, communication gaps, conflict handling and prioritization. Predictive ML model is needed to identify critical factors that contribute to dissatisfaction so that improve can be made.

4.3 Comparative Performance Metrics of ML Model

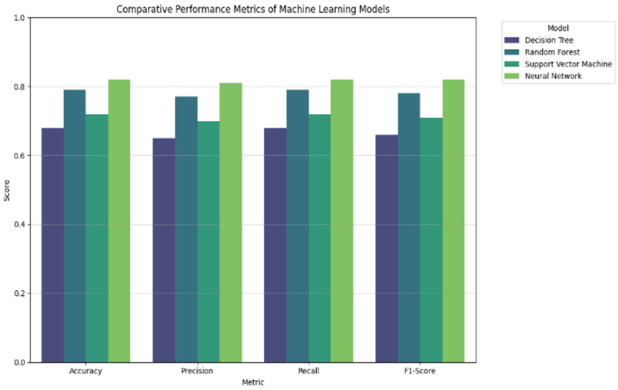


Fig 02: Comparative Performance Metrics

The bar chart displays the visual comparison of the result of accuracy precision recall and F1 score for every one of the four machine learning models. The Neural Network model superior performance across all metrics is clearly demonstrated all of the algorithm closely by the Random Forest which is consistently with observation in the result of thesis.

4.4 Distribution of Best Things Categories

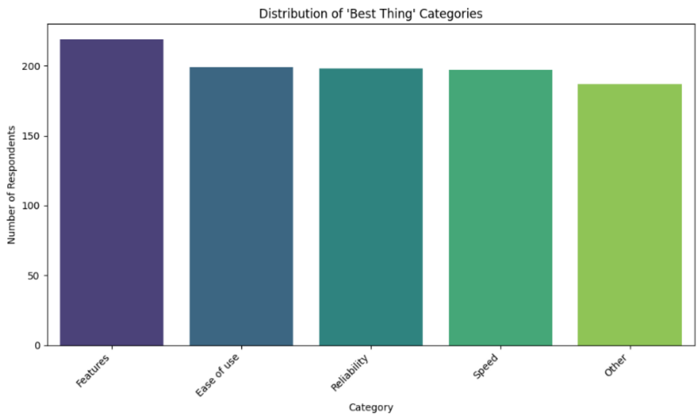


Fig 3: Distribute of best Things Categories

The bar chart of best things categories distribution illustrates how responds rated the most positive aspect of the system. Shows the five main categories each categories reflect what users consider the strongest or most satisfying part of their experience of use. Overall, the distribution demonstrate that user appreciate several different aspects of the system.

4.5 Distribution of Frustrating Aspect Categories

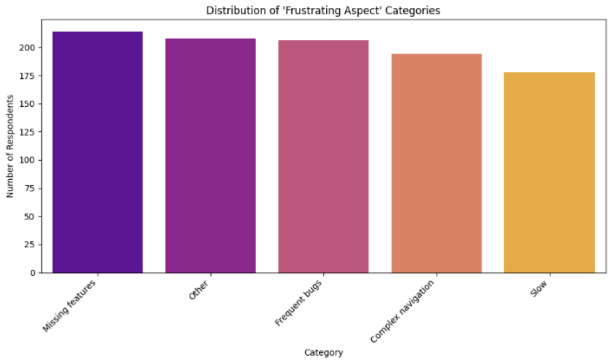


Fig 4: Distribution of Frustrating Aspect Categories

The bar chart explains how survey respondents rated different frustrating aspects of a system or application. Each bar represents the number of people who reported a specific issue as their biggest frustration.

4.6 Usability Performance and Feature Statements

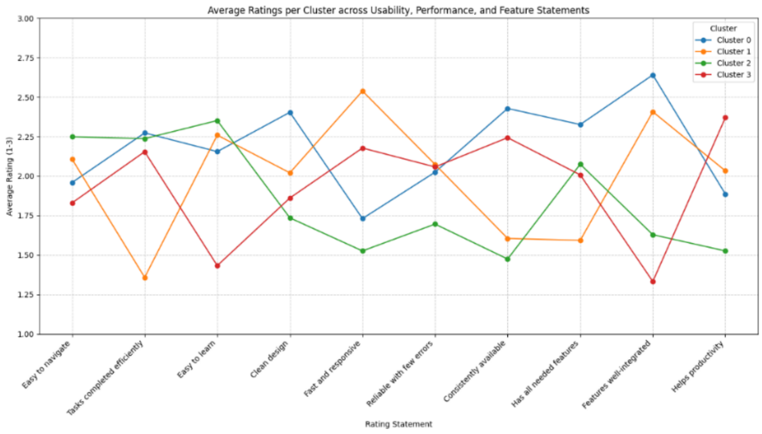


Fig 5: Average Rating per cluster across Usability Performance and Feature Statements

The line chart present how different groups of respondent's clusters rated various aspects of a software system. Clustering performed was on satisfaction related features so each cluster represent a distinct pattern of the user perception. There are four cluster (Cluster 0, Cluster1, Cluster 2, and Cluster 3) each show the average rating 1-3 scale for item of usability, performance, reliability and feature availability.

4.7 Comparative Performance Metrics of Machine Learning Models

Table 4.1 With a precision of 0.82 and an F1 score of 0.82 find from the Neural Network model was the best overall performer. Using the Random Forest model, the performance of the Random Forest model was shown strong performance accuracy of 0.79 and an F1 score of 0.78. Decision trees are shown lack of ability to capture complex non-linear relationship suggest poor performance across all metrics.

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	0.68	0.65	0.68	0.66
Random Forest	0.79	0.77	0.79	0.78
SVM	0.72	0.70	0.72	0.71
Neural Network	0.82	0.81	0.82	0.82

Source: Google Colab

4.8 Key Requirement Attributes and Project Parameters Influencing Stakeholder Satisfaction

Several key requirements were identified by post analysis feature important scores for tree-based model and sensitivity analysis for neural network and SVM and stakeholder satisfaction significantly influencing by project parameters. The useability statements are consistency the predictor of top for easy to navigate and clean design. The important of users experience UX design cannot understand so that as high ratings I these areas strongly correlate with satisfaction of overall higher in requirement formulation. The system met the need of stakeholder and the completeness and seamless integration of features was highly influential for improving features well integrated. Reliability was a significant factored in performance with system that were perceived led for consideration dissatisfaction RE highlight the need of quality assurance. In the duration on frequency project with longer and less frequent interaction on stakeholder that is lower satisfaction. It is suggested that stakeholder is positive impact by continuous engagement and short development cycle. Satisfaction was influenced by the system specific purpose as data intensive application frequently require higher standard usability performance and realisability due to critical nature of data. Enhancing communication contribution to optimizing requirements prioritization reducing projects risk. Find that from this analysis by comparative offer to optimizing requirements prioritization reducing projects risk and enhancing communication within the requirement engineering process. Project manager can prioritize requirements that have the highest impact on stakeholder satisfaction by the identifying most influential requirements attributes. Reducing projects risk by understanding the parameters that lead to dissatisfaction enables risk mitigations. Checking critical functionalities identified as plan for especially longer duration that is project teams can establish more rigorous quality assurances. Implementing structured plans of communication so that especially for project with long duration drops of satisfaction so that reducing the risk of project failure due to unmet stakeholder expectation. Data Driven insights can be used to strength communication within RE to support discussion with stakeholder. RE professional have the opinion to present model backend finding that support by model instead of abstract arguments such as our analysis show that ease of navigation is a top driver of satisfaction. So now let's focus refining the UI/UX early will lead to more objective and effective communication which can bridge the gap between technical teams and business stakeholder.

4.9 Data Analysis Key Findings

The analysis was carried out using a simulated dataset of 1000 project instance each with 19 attributes was used in the analysis setup. Decision Tree Random Forest Support Vector Machines and Neural Network used to classify data and utilizing 70 or 30 tarin set split and 5-fold cross validation. Model performance find the neural network performed the best with accuracy of 0.82 and F1 score 0.82. Decision tree performed poorly across all met with accuracy 0.79 and F1 score 0.78 and the random forest model following closely all metrics with accuracy of 0.68 and F1 score of 0.66. The key requirements are easy to navigate clean design and has all the necessary features and features well integrated and attribute performance relabel with few errors identified by top predictors of stakeholder satisfaction. The duration and frequency of interaction as well as the primary purpose of the system had significant effect on stakeholder. Lower satisfaction is often associated with longer durations and less frequent interactions.

4.10 Insights or Next Steps

Prioritize important requirements using data driven insights particularly from the Neural Network model for optimized resource allocation. By identifying dissatisfaction drivers like lack of reliability and infrequent communication proactive risk mitigation enhanced communication can be achieved. The requirement engineering process can improve communication and project risk are reduced through data backend discussion with stakeholder.

CHAPTER 5

CONCLUSION

A comprehensive survey of 1000 respondent was conduct in this study to investigate user satisfaction usability perception to goal of this study was to explore user satisfaction. The analysis shows that despite some users feeling satisfied the system overall sentiment is neutral to navigate despite the fact that some users express moderate satisfaction. The system was not perceived well by users as most usability and performance metrics averaged around the mid-point of the rating scale the system as strongly effective or user friendly. A significant portion of responding report being Dissatisfied or Very Dissatisfied satisfaction distribution shows that a substant number of responding which highlight significant gaps in the system performance feature completeness navigate ease and responsiveness. Easy of navigate task efficiency system responsiveness for the correction analysis further confirms that user satisfaction is most strongly influenced by factored. The sharp decrease in satisfaction when these aspects fail to perform well. The findings are strengthened by qualitative responses with users frequently citing slow performance complicated navigate and missing features and as their primary frustration's inefficiency. Improvements in reporting capabilities and enhanced automation for the needs of users expressed. According to the finding the system useful functions are present but its usability and performance challenges significantly limit its use. Targeted design improvements performance optimized and features enhancement to address these issues essential improve stakeholder experience. The study finding offer a clear direction for future improvement and are a valuable base for redesign the system to better meet user needs. Correlation and cluster analysis confirmed that user satisfaction is most influenced strongly by usability related factors, particularly ease of navigate, task efficiency and system responsiveness. The qualitative findings issues of user frequently citing slow system performance complicated navigate and unclear workflow also missing features as major frustrations.

REFERENCES

[1] A. van Lamsweerde, Requirements Engineering: From System Goals to UML Models to Software Specifications. Wiley, 2009.

[2] I. Sommerville and P. Sawyer, Requirements Engineering: A Good Practice Guide. Wiley, 1997.

[3] T. M. Mitchell, Machine Learning. McGraw-Hill, 1997.

[4] A. Ferrari, G. Gori, and B. Lettieri, "Natural language processing for requirements engineering: The best is yet to come," IEEE Software, vol. 35, no. 5, pp. 115–119, 2018.

[5] D. Raturi, S. Rathore, and S. Kumar, "A machine learning-based approach for software requirements classification," Expert Systems with Applications, vol. 140, 2020.

[6] S. Malhotra, "Applying machine learning algorithms for prediction and classification of software requirements," Int. J. of Advanced Computer Science and Applications, vol. 9, no. 6, pp. 41–47, 2018.

[7] H. Liu and H. Motoda, Feature Selection for Knowledge Discovery and Data Mining. Springer, 1998.

[8] J. Cleland-Huang, "Requirements engineering with machine learning and natural language processing: Opportunities and challenges," Proc. IEEE Int. Requirements Engineering Conf., pp. 4–13, 2019.

[9] R. E. Freeman, Strategic Management: A Stakeholder Approach. Cambridge University Press, 1984.

[10] B. Boehm, "A spiral model of software development and enhancement," Computer, vol. 21, no. 5, pp. 61–72, 1988.

[11] V. N. Vapnik, Statistical Learning Theory. Wiley, 1998.

[12] T. M. Mitchell, Machine Learning. McGraw-Hill, 1997.

[13] K. P. Murphy, Machine Learning: A Probabilistic Perspective. MIT Press, 2012.

[14] C. M. Bishop, Pattern Recognition and Machine Learning. Springer, 2006.

[15] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques. Elsevier, 2011.

[16] J. Cleland-Huang, "Requirements engineering with machine learning and natural language processing: Opportunities and challenges," Proc. IEEE Int. Requirements Engineering Conf., pp. 4–13, 2019.

[17] A. Ferrari, G. Gori, and B. Lettieri, "Natural language processing for requirements engineering: The best is yet to come," IEEE Software, vol. 35, no. 5, pp. 115–119, 2018.

[18] D. Raturi, S. Rathore, and S. Kumar, "A machine learning-based approach for software requirements classification," Expert Systems with Applications, vol. 140, 2020.

[19] J. Cleland-Huang, "Deep learning for traceability link recovery," IEEE Transactions on Software Engineering, vol. 47, no. 8, pp. 1525–1538, 2021.

[20] S. Malhotra, "Applying machine learning algorithms for prediction and classification of software requirements," Int. J. of Advanced Computer Science and Applications, vol. 9, no. 6, pp. 41–47, 2018.

[21] A. Sharma and V. Gupta, "Predicting software project success using ensemble machine learning models," Procedia Computer Science, vol. 167, pp. 1820–1829, 2020.

[22] F. Dalpiaz, P. Giorgini, and J. Mylopoulos, "Adaptive requirements engineering for socio-technical systems," Requirements Engineering, vol. 21, no. 3, pp. 311–329, 2016.

[23] J. Creswell and C. Creswell, Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 5th ed., Sage, 2017.

[24] P. Ralph, "Toward methodological guidelines for process research in requirements engineering," Requirements Engineering, vol. 26, no. 3, pp. 325–342, 2021.

[25] J. Cleland-Huang, "Requirements engineering with machine learning and natural language processing," IEEE RE Conference, 2019.

[26] S. Malhotra, "Applying machine learning algorithms to software requirement prediction," IJACSA, vol. 9, no. 6, 2018.

[27] I. Etikan, S. Musa, and R. Alkassim, "Comparison of convenience and purposive sampling," American Journal of Theoretical and Applied Statistics, vol. 5, pp. 1–4, 2016.

[28] S. Kvale, Doing Interviews, Sage, 2007.

[29] R. Krueger and M. Casey, Focus Groups: A Practical Guide for Applied Research, Sage, 2015.

[30] M. Angrosino, "Naturalistic observation," in The Sage Encyclopedia of Qualitative Research Methods, Sage, 2008.

[31] G. Bowen, "Document analysis as a qualitative research method," Qualitative Research Journal, vol. 9, no. 2, pp. 27–40, 2009.

[32] A. Bryman and E. Bell, Business Research Methods, Oxford Univ. Press, 2015.

[33] C. Bishop, Pattern Recognition and Machine Learning, Springer, 2006

[34] V. Braun and V. Clarke, “Using thematic analysis in psychology,” Qualitative Research in Psychology, vol. 3, no. 2, pp. 77–101, 2006.

[35] I. Sommerville and P. Sawyer, Requirements Engineering: A Good Practice Guide, Wiley, 1997.

[36] J. Cleland-Huang, “Requirements engineering with machine learning,” IEEE RE Conf., 2019.

[37] S. Malhotra, “Machine learning for predicting software requirement quality,” IJACSA, 2018.

[38] T. M. Mitchell, Machine Learning, McGraw-Hill, 1997.

[39] I. Sommerville, Software Engineering, 10th ed., Pearson, 2015.

[40] A. Ferrari, “NLP and ML in RE: Opportunities,” IEEE Software, 2018

221-35-1049

ORIGINALITY REPORT

18%	16%	9%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	3%
2	Submitted to NCC Education Student Paper	1%
3	www.coursehero.com Internet Source	1%
4	Saurabh Tiwari, Santosh Singh Rathore, Shreya Sagar, Yash Mirani. "Identifying Use Case Elements from Textual Specification: A Preliminary Study", 2020 IEEE 28th International Requirements Engineering Conference (RE), 2020 Publication	1%
5	S.P. Jani, M. Adam Khan. "Applications of AI in Smart Technologies and Manufacturing", CRC Press, 2025 Publication	<1%
6	export.arxiv.org Internet Source	<1%
7	umpir.ump.edu.my Internet Source	<1%
8	trap.ncirl.ie Internet Source	<1%
9	vdocuments.site Internet Source	<1%
10	Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dhirendra Kumar Shukla. "Recent	<1%

Account Clearence