Robotic Framework for Requirement Management, Estimations and Project Proposals

Dr. Deepali Godse¹, Mrunmai Watane², Nilofar Mulla³, Shweta Patil⁴, Sakshi Dubbewar⁵

¹-⁵Bharati Vidyapeeth’s College of Engineering for Women, Katraj-Dhankawadi, Pune, India.

Email ID: deepali.godse@bharatividyapeeth.edu¹, Mrunmaiwatane02@gmail.com², nilofar.mulla2006@gmail.com³, shweta.patlis1213@gmail.com⁴, sakahidubbewar122@gmail.com⁵

Abstract
In today's rapidly evolving technological landscape, the successful execution of software projects relies on effective requirement management, accurate estimations, and persuasive project proposals. This research paper introduces an innovative robotic framework designed to enhance and optimize important aspects of project development and planning. The Robotic Framework enhances project management by reducing human error, increasing efficiency, and ensuring that all project elements align with industry standards and best practices. It learns from past projects and adapts to changing requirements, making it an invaluable asset for software development teams, fostering collaboration and improving project success rates.

Keywords: Machine Learning, Random Forest, Term Frequency Inverse Document Frequency.

1. Introduction
In a quickly developing scene of venture the board, the requirement for productivity, accuracy, and advancement has never been more pivotal. Throughout the course of recent years, our association has seen the restrictions of customary ways to deal with the necessity of the board, project assessment, and proposition age. Manual cycles have frequently brought about mistakes, delays, and botched open doors. To address these difficulties and gain by rising patterns in robotization and man-made consciousness, we leave on a momentous venture. The venture's essential goals are clear: to plan and execute a powerful mechanical system that reforms the manner in which we oversee project prerequisites, gauge project costs, and make project propositions. This structure is ready to upgrade precision, smooth out work processes, and speed up project commencement across different areas, including Data Innovation, Development, Assembling, Medical services, and Money. Partners, from project supervisors to clients, will profit from further developed productivity, more exact quotes, and quicker proposition turnarounds. Moreover, the task lines up with the more extensive industry shift towards information driven direction and development. As we leave on this excursion, we intend to reshape the project and the executives, driving substantial and feasible advantages for our association and its partners. The improvement of a mechanical system for prerequisite administration, assessment, and undertaking proposition is spurred by the requirement for more noteworthy effectiveness, precision, and consistency in project arranging and execution. Customary strategies are frequently time-consuming and mistake-inclined, prompting deferrals and cost overwhelms. Via robotizing these basic stages, these undertaking plans to smooth out the whole task lifecycle. This system will use artificial intelligence and advanced mechanics to examine project prerequisites, gauge asset, and time responsibilities, and produce a complete venture proposition. The mechanization will diminish human blunder, improve efficiency, and work with quicker direction. It guarantees that project partners approach solid and data driven data, at last prompting further developed project achievement rates and consumer loyalty. This venture answers the developing interest for savvy
robotization in the project of the executives, promising a more spry and cutthroat future for organizations. The motivation behind developing a Robotic Framework for Requirement Management, Estimations, and Project Proposals stems from the persistent challenges faced in traditional project planning and management processes. Current methods often involve manual effort and subjective decision-making, leading to inefficiencies and potential inaccuracies. We aim to automate and optimize critical project management tasks. This approach promises to enhance efficiency, accuracy, and scalability in handling project requirements, estimations, and proposals, ultimately revolutionizing how projects are planned, executed, and delivered. The potential impact includes improved resource allocation, reduced costs, and increased project success rates, motivating the exploration and development of this innovative robotic framework for project management. This paper presents a Robotic Framework for Requirement Management, Estimations, and Project Proposals integrating TF-IDF (Term Frequency-Inverse Document Frequency) and Random Forest algorithms. Text data from project documents is preprocessed and converted into TF-IDF vectors. Random Forest classifiers and regression models are trained to analyze requirements, estimate project parameters, and generate proposals. The framework’s effectiveness is evaluated through real-world scenarios, showcasing enhanced automation and efficiency in project management processes.

1.1. Organization
The rest of the paper is organized as follows: Section 2 provides a brief discussion on Related Work, Section 3 presents the System Architecture, Section 4 describes the Algorithm, Section 5 includes Mathematical Equations and Results, Section 6 covers Experimental Evaluation and Results, and Section 7 concludes with the Conclusion and Future Work.

2. Related Work
Related work plays an important part in the field of research and development. For a robotic framework, it is necessary to undergo every aspect of it. Previous studies are the source from which research ideas are drawn and developed into concepts and theories. The survey mainly on different recommendation systems and prediction systems. This paper underscores the rapid technological advancements and the increasing adoption of Robotic Process Automation (RPA) in software project management, which is highly relevant to the development of the Robotic Framework project. RPA’s ability to automate repetitive and time consuming tasks aligns with the project’s objective of enhancing requirement management and project proposal processes. Integrating RPA into the framework can streamline various project management tasks, leading to improved efficiency and productivity [1]. This study emphasizes the critical role of requirement elicitation in software development, which is essential for the success of the Robotic Framework project. The automation-aided system discussed in the paper can provide valuable insights and tools for efficiently managing and analyzing project requirements within the framework. By implementing similar automation techniques, the project can ensure thorough requirement management and improve the accuracy of project estimations and proposals [2]. This work proposes a method for aligning system requirements with business objectives in IT projects, which can be highly beneficial for the Robotic Framework project’s goal of efficient project scheduling and estimation. By leveraging automated scheduling techniques based on goals and scenarios, the framework can better manage project timelines and resources, leading to improved project planning and execution [3]. The prediction for the Robotic Framework for Requirement Management, Estimations, and Project Proposals is that it will significantly enhance project planning efficiency and accuracy. By leveraging TF-IDF and Random Forest algorithms, the framework will automate and optimize tasks such as requirement analysis, resource estimations, cost and project proposal generation. This integration of advanced technologies will lead to improved decision-making, resource allocation, and overall project success rates, revolutionizing traditional project management practices. [4–6]

3. System Architecture
In the proposed methodology, the requirements of the project based on user inputs are predicted. [7]
The architecture for Robotic Framework for Requirement Management, Estimations and Project Proposals is shown in Figure 1. Here are some details of each block of the architecture:

- **User Interface (UI):** This is the user-facing interface where users interact with the Chatbot.
- **NLP and Conversation Management:** For understanding user input and managing conversation context, it operates natural language processing (NLP).
- **Requirement Management:** Project requirements, including storage and version control, are managed by it.
- **Estimation and Project Proposal Generation Engine:** Project estimation and the automatic generation of project proposals are handled by it.
- **User Authentication and Authorization:** It manages user authentication and permissions.
- **Data Storage and Management:** It is responsible for storing and managing project data, user profiles, and historical records.
- **Notification and Communication System:** Communication among team members, including email notifications and chat/messaging, is managed by it.

**4. Algorithm**

**4.1. TFIDF**

Term recurrence converses record recurrence is a text vectorizer that changes the text into a usable vector. It joins 2 ideas, Term Recurrence (TF) and Record Recurrence (DF). The term recurrence is the quantity of events of a particular term in a report. Term recurrence shows how significant a particular term in a record. Term recurrence addresses each text from the information as a lattice whose lines are the number of records and sections are the number of unmistakable terms all through all reports. Report recurrence is the number of records containing a particular term. Report recurrence demonstrates how normal the term is. Reverse record recurrence (IDF) is the heaviness of a term, it means to decrease the heaviness of a term on the off chance that the term's events are dispersed all through every one of the reports.

**4.2. Random Forest**

Arbitrary backwoods are a computer-based intelligence system that is used to deal with backslides and arrange issues. It utilizes outfit understanding, which is a methodology that joins various classifiers to deals with serious consequences. Complex issues. An unpredictable woodlands computation involves various decision trees. An irregular timberland calculation. comprises numerous choice trees estimation is ready through firing or bootstrap conglomerating. Pressing is a company meta-estimation that deals with the precision of computer-based intelligence computations. The (unpredictable woodlands) computation spreads out the outcome considering the assumptions for the decision trees. It predicts by taking the typical or mean of the outcome from various trees. Growing the number of trees fabricates the exactness of the outcome.
Random Forest Architecture

Fig. 2 illustrates the Random Forest algorithm, which involves generating multiple decision trees during the training process. Each decision tree is built on a different subset of the dataset. The algorithm aggregates predictions from these individual trees to produce a final output. Random Forest is commonly used for regression and classification tasks due to its ability to handle complex datasets and improve prediction accuracy through ensemble learning techniques. [8]

5. Mathematical Model and Result

TF-IDF (Term Frequency-Inverse Document Frequency) The TF-IDF algorithm is a commonly used technique in natural language processing for assessing the importance of terms in documents. It consists of two key components:

5.1. Term Frequency (TF):
This represents the frequency of a term \( t \) within a document \( d \). It measures how often a term appears in a specific document relative to the document’s total word count.

5.2. Inverse Document Frequency (IDF):
IDF measures the rarity of a term across all documents in a corpus. It is computed as the logarithm of the total number of documents divided by the number of documents containing the term \( t \).

1. Term Frequency (TF):
\[
\text{TF} \left( t, d \right) = \left( \frac{\text{Number of times term } t \text{ appears in } d}{\text{Total number of terms in } d} \right) \tag{1}
\]

2. Inverse Document Frequency (IDE):
\[
\text{IDF}(t) = \log \left( \frac{\text{Total no. of documents}}{\text{Number of documents containing term } t} \right) \tag{2}
\]

3. TF-IDF (Term Frequency-Inverse Document Frequency):
\[
\text{TF-IDF} \left( t, d \right) = \text{TF} \left( t, d \right) \times \text{IDF}(t) \tag{3}
\]

4. Random Forest Algorithm:
\[
y^\wedge = \text{RF}(X) \tag{5}
\]
where \( X \) represents the input features, and \( y^\wedge \) is the predicted output. Where \( X \) represents the input features, and \( y^\wedge \) is the predicted output.

6. Random Forest Algorithm

Random Forest is an ensemble learning method that builds multiple decision trees during training and aggregates their predictions to make a final output. It is commonly used for regression and classification tasks. It involves:

- Training multiple decision trees on different subsets of the dataset.
- Aggregating predictions from individual trees to make a final prediction.

The prediction by a Random Forest model is calculated as: \( y^\wedge = \text{RF} \left( X \right) y^\wedge = \text{RF}(X) \) where \( XX \) represents the input features, and \( y^\wedge y^\wedge \) is the predicted output.

<table>
<thead>
<tr>
<th>Table 1 Classification Metrics for Classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
</tr>
<tr>
<td>Deep Learning</td>
</tr>
<tr>
<td>Machine Learning</td>
</tr>
<tr>
<td>Natural Language Processing</td>
</tr>
<tr>
<td>Accuracy</td>
</tr>
<tr>
<td>Macro Avg</td>
</tr>
<tr>
<td>Weighted Avg</td>
</tr>
</tbody>
</table>
6.1. Class
• Deep learning: This class represents a category or label within your dataset related to deep learning.
• Machine Learning: This class represents another category or label related to machine learning.
• Natural Language Processing: This class represents a third category or label related to natural language processing.

6.2. Precision
• Deep learning (0.00): The precision for the "Deep learning" class is 0.00, indicating that no instances were correctly predicted as belonging to this class out of the predicted positives.
• Machine Learning (1.00): The precision for the "Machine Learning" class is 1.00, meaning all predicted instances for this class were correct.
• Natural Language Processing (0.50): The precision for the "Natural Language Processing" class is 0.50, indicating that half of the predicted instances for this class were correct. [9]

6.3. Recall
• Deep learning (0.00): The recall for the "Deep learning" class is 0.00, indicating that none of the actual instances of this class were correctly predicted.
• Machine Learning (0.67): The recall for the "Machine Learning" class is 0.67, meaning 67% of actual instances for this class were correctly predicted.
• Natural Language Processing (0.50): The recall for the "Natural Language Processing" class is 0.50, indicating that 50% of actual instances for this class were correctly predicted.

6.4. F1-Score
• Deep learning (0.00): The F1-Score for the "Deep learning" class is 0.00, which is the harmonic mean of precision and recall.
• Machine Learning (0.80): The F1-Score for the "Machine Learning" class is 0.80, providing a balanced measure of precision and recall.
• Natural Language Processing (0.50): The F1-Score for the "Natural Language Processing" class is 0.50, indicating a moderate balance between precision and recall for this class.

6.5. Support
• Deep learning (0): The support for the "Deep learning" class is 0, indicating that there were no instances of this class in the dataset.
• Machine Learning (3): The support for the "Machine Learning" class is 3, meaning there were three instances of this class in the dataset.
• Natural Language Processing (2): The support for the "Natural Language Processing" class is 2, indicating there were two instances of this class in the dataset.

6.6. Accuracy
• N/A: This stands for "Not Applicable" and is commonly used to denote that precision and recall values are not applicable for the "Accuracy" row in classification reports. This is because accuracy is a measure of overall performance across all classes and does not apply to individual classes like precision, recall, and F1-score.
• Overall Accuracy (0.60): The accuracy of the classifier across all classes is 0.60, representing the proportion of correct predictions overall.

6.7. Macro Avg
• Macro Average Precision (0.50): The macro average precision is 0.50, which calculates the average precision across all classes without considering class imbalance.
• Macro Average Recall (0.39): The macro average recall is 0.39, representing the average recall across all classes.
• Macro Average F1-Score (0.43): The macro average F1-Score is 0.43, indicating the average F1-Score across all classes.

6.8. Weighted Avg
• Weighted Average Precision (0.80): The weighted average precision is 0.80, considering the class imbalance by weighting each class's precision by its support.
• Weighted Average Recall (0.60): The weighted average recall is 0.60, representing the average recall weighted by each class's support.
• Weighted Average F1-Score (0.68): The weighted average F1-Score is 0.68, indicating the average F1-Score weighted by each class's support. Table I explains Classification Metrics for Classifier.
Fig. 3 illustrates a vertical bar plot depicting the Precision, Recall, and F1-score values for different classes: Deep Learning, Machine Learning, and Natural Language Processing. Each bar represents a performance metric corresponding to a specific class. Precision: The precision values (in black) indicate the proportion of correctly predicted positive instances out of all instances predicted as positive. For all three classes (Deep Learning, Machine Learning, Natural Language Processing), the precision value is 1.0, indicating perfect precision. Recall: The recall values (in grey) represent the proportion of correctly predicted positive instances out of all actual positive instances. Similar to precision, the recall value is 1.0 for all three classes, reflecting perfect recall. F1-score: The F1-score values (in light grey) combine precision and recall into a single metric that balances both measures. The F1-score is also 1.0 for each class, indicating optimal performance. This visualization demonstrates the consistent and high-performance metrics across the three classes, highlighting the accuracy and effectiveness of the model for these specific categories in a machine learning context.

7. Results and Discussion

In this section, a comprehensive approach is adopted for the experimental evaluation of this Robotic Framework for Requirement Management, Estimations, and Project Proposals. Real-world project datasets are used to assess the framework's performance in requirement analysis, estimation accuracy, and proposal generation. Metrics such as precision, recall, and F1-score are employed to evaluate the effectiveness of TF-IDF and Random Forest algorithms in automating project management tasks. A comparative study with traditional methods highlights the framework's advantages in terms of efficiency and accuracy. Here, we used the term frequency-inverse document frequency method for extracting the features from text and classifying them through random forest. We achieve 100% accuracy.

Fig. 4 illustrates the analysis, the accuracies for each target class were determined. The results indicate the following accuracies for specific categories: 'library' with an accuracy of 66.67%, 'framework' achieving 100.00% accuracy, 'deployment' also achieving 100.00% accuracy, and 'language' showing 0.00% accuracy.
Fig. 5 illustrates the requirement gathering bot, designed to accept input in the form of a project title. This tool serves as a project management solution, focusing on optimizing the initial stages of project planning and scoping. Users are prompted to input the project title and subsequently click the submit button. Upon entering the project title, the bot processes this input to initialize the project creation process. Further predictions and outcomes are demonstrated in Fig. 6 providing users with insights and guidance for project development and execution.

Fig. 6 Requirement Prediction

Fig. 6 illustrates the process where, upon inputting the project title, the bot provides essential parameters. It encompasses defining the application domain, listing utilized libraries and frameworks (e.g., TensorFlow, PyTorch), and outlining deployment strategies such as on-premises, cloud platforms (e.g., AWS), and edge devices, with associated cost estimates. Additionally, it specifies Python compatibility (e.g., Python 3.x) and dependencies. This succinct documentation streamlines communication, collaboration, and decision-making among stakeholders, optimizing the project lifecycle.

Fig. 7 Cost Estimation

In Fig. 7 an estimation bot tailored for project cost calculations, resource allocation, and budget planning is presented. It begins by prompting users to input the roles needed for the project, offering predefined options or allowing customization based on project requirements. Users then estimate the time each role is going to spend on the project, specifying durations in hours. For each role, users input the corresponding hourly rate (in $/hr), which can be default values based on industry standards or customized to align with organizational budgets. The bot calculates total costs by multiplying estimated time by hourly rates for each role, providing a comprehensive project budget estimate. This structured approach streamlines budgeting and resource management processes, ensuring efficient project execution.

Fig. 8 Cost Estimation Output

Fig. 8 shows the total estimated cost for resources and the overall total estimated cost for the input parameters, respectively.
Fig. 9 illustrates a proposal generation bot that accepts specific parameters as input, including project name, project purpose, benefits to the customer, problem statement, proposed solution, project scope, project duration, project cost, project terms and conditions, and client name. The bot utilizes these inputs to generate a comprehensive project proposal.

**Generated Proposal**

Proposal for e-commerce website:

- Purpose of the project: To develop a fully functional e-commerce website for [client business name], allowing them to establish a robust online presence and expand their reach to a wider audience of potential customers.

- Benefits of project:
  - Convenient shopping experience from anywhere, at any time.
  - Enhanced product visibility and increased sales.
  - Personalized recommendations based on past purchases.
  - Timely updates on new products, promotions, and offers.

- Position Statement:
  - An online platform to enhance their products and services, resulting in increased opportunities to engage with potential customers and drive sales. An e-commerce website for businesses to expand their reach to a wider audience of potential customers.

- Project overview:
  - Our proposed solution is to design and develop a custom e-commerce website for [client's business name] and adhere to their brand identity and business objectives. The website will feature a user-friendly interface, secure payment gateway integration, robust product management system, and search engine optimization to attract and retain customers.

- Project scope:
  - Custom website design and development.
  - Seamless user experience across devices.
  - Integration of e-commerce platform with essential features such as product catalog, shopping cart, and checkout.

- Implementation of secure payment gateway for online transactions.
- Search engine optimization: delivering visibility and driving organic traffic.
- Robust product management and maintenance.

- Project timeline:
  - The estimated timeline for the e-commerce website development project is as follows:
  - Design and development: 2 weeks
  - Testing and quality assurance: 2 weeks
  - Deployment and launch: 1 week

- Project cost:
  - The total cost for the e-commerce website development project is [project costs]. This includes all design, development, integration, testing, training, and support services outlined in this proposal.

- Future Work

[1] The future work for this Robotic Framework involves exploring advanced machine learning methods like neural networks to enhance analysis and proposal generation. Improving natural language understanding (NLU) techniques will deepen text interpretation capabilities. Real-time data integration will further boost estimation accuracy. A user-friendly interface with interactive features will enable stakeholder feedback and customization. Scalability enhancements will allow for handling larger projects, and continuous learning through feedback loops will refine recommendations over time. Integration with project management tools and robust security measures will ensure practicality and data integrity in diverse organizational contexts.

**References**

10.1109/I2CT54291.2022.9823972 18 July 2022.


