Industrial Engineering Journal
ISSN: 0970-2555
Volume: 52, Issue 10, No. 1, October: 2023

AUTOMATED PROCESSING OF PROJECT PROPOSAL DOCUMENTS USING ARTIFICIAL INTELLIGENCE: SURVEY AND ITS RELEVANCE IN E-GOVERNANCE

Mrs. Manisha Tiwari, MPSTME, NMIMS University, Mumbai, India
Dr Padmanabha Aital, MPSTME, NMIMS University, Mumbai, India
Dr.Padmaja Joshi, C-DAC, Mumbai, India

Abstract:
Artificial intelligence (AI) has risen to prominence in recent years due to the increasing need for fully autonomous systems. The field of artificial intelligence is now thriving, with a profusion of new models arising in the R&D stages of several disciplines, including science, economics, and engineering. The rapid development in the artificial intelligence (AI) domain involving the application of natural language-based techniques, Machine learning, and deep learning has greatly enhanced document processing efficiency in various business domains. This study aims to survey the published work on AI-based techniques for the above application which includes the automated processing of documents submitted at the proposal stage of project development using natural language processing, machine learning, and deep learning. The study analyses such techniques in terms of relevance and efficiency including accuracy in relevant information extraction, classification, and question answering. The goal of the study of work is to assist the research community in enhancing current research initiatives and predicting future advances. Further, this work also explores the applicability of studied techniques in one of the e-governance systems as a case study.

Keywords: Requirement Engineering (RE), Artificial Intelligence(AI), Machine Learning(ML), Deep Learning(DL), Request for proposal (RFP), e-governance

Introduction

The adoption of new-age technologies including Artificial Intelligence in the domain of Project Management has played a vivid role in the cognitive simplification of complex tasks involving a diverse set of documents ranging from proposal documents to user reviews, stakeholders' opinions, etc. to attain the business objectives [1], [4]. In a typical scenario of a software project development through a tender process, the initial stages of the project involve sharing of project requirements either through formal documents like a Request for Proposal (RFP), Request for Quotation, or in the form of concept notes, the project requirements are aggregated through client interviews and subsequently condensed in the form of Functional Requirement Specification (FRS) and Software Requirement Specification (SRS) documents.

As per the terminology prevalent in the industry, the contract/proposal documents incorporate detailed information about the agreement between the owner and the contractor based on functionality, roles and responsibilities, delivery timelines, deliverables and pricing. This research work focuses on such contract documents for data extraction and understanding the proposal documents. The main focus of this paper is to simplify the life of a project manager or a bidder while bidding for the RFP [6]. The need to the identification of such requirements automatically in the life cycle of projects in various domains has been addressed in the literature [7], [59], [60], [65], [63]. The attempt in this paper is to study the variation in strategies deployed to study proposal documents used in projects belonging to unrelated domains. This included the calibration of various parameters to understand and extract granular project requirements. The objective of this study is to understand the current status of automated requirements extraction, classification, and further question-and-answer usage from proposal documents and see if it can be generic or domain-specific. The paper also attempts to explore the relevance of automation and its significance in increasing the efficiency of these evaluation techniques in terms of time, effort and cost. These proposal documents not only cover the essential requirements of the project but also include details of roles and responsibilities, service level agreements, timelines, formatting information, etc. Classifying such information using automated techniques such as machine learning and deep learning may help and simplify the overall project bidding process and the development process.

The goal of this paper is to give a comprehensive overview of AI approaches used for gathering requirements from proposal/contract documents and discuss the trade-offs between the existing approaches. The three components of this work's contribution are as follows:

1) In order to create a thorough presentation of existing machine learning/deep learning approaches utilized for information extraction, classification, and question answering from the proposal documents at the preliminary stages of project development across multiple domains, a review is undertaken.
2) In order to determine the primary contribution of the techniques, the identified literature is categorized according to particular criteria.
3) Research gaps are brought to attention to generate impulses for further research, especially in for sub-systems of the e-governance system as a case study.

The remainder of this paper is structured as follows. Section 2 discusses the usage of automation with AI in Project Management Section 3 introduces the concept of contracts or proposal documents especially the RFP or Request for proposal documents as it sets the background for understanding the need for research for the classification of requirements from them. Section 4 introduces the concept of information extraction, classification, and question answering in the context of AI techniques like machine learning and deep learning. Afterward, Section 5 summarizes the research literature and then divides it into other categories. Section 6 discusses findings and observations and their applicability in the e-governance system as a case study. With Section 7 we conclude the paper with a summary and future work.

The forthcoming sections are going to discuss the AI approaches and relevance in project management, therefore following questions are designed to streamline the research
discussion

Research Questions

This study's objective is to examine several AI methodologies used in project proposal documents. The following research questions (RQ) were established

RQ1: What is the significance of automation using Artificial Intelligence in project management?

RQ2: What current techniques/tools/frameworks using Artificial Intelligence are available in the context of proposal documents?

RQ3: Classify the current work in terms of Information Extraction, Text Classification, Question and Answering, and others.

RQ4: What are the significant benefits of incorporating such methodologies in the e-governance domain?

Significance of Artificial Intelligence in Project Management

In order to offer enterprises and everyone engaged with it anything of value, project management utilizes certain knowledge, abilities, tools, and procedures. Project management (PM) is unpredictable and extremely fluid, forcing numerous difficulties for organizations and professionals. It is observed that the combination of AI in the organization provides better results. Also, as per the reports, a steep increase in market grow question-answering technologies by 2024 is likely to be there [8]. AI as the name suggests is the technology used for building human intelligence in the system. Some applications where AI plays an important role are pattern recognition and anomaly detection, meaning extraction meaning from images, text, or speech, natural language processing and making predictions to provide suggestions, recommendations etc. Technologies used for AI include machine learning, deep learning etc. Term Robotic Process Automation, RPA stands for robotic process automation, which is when software automates business activities.AI, RPA and Soft computing together form the basis of Intelligent automation in businesses [9], [10]. Table I provides insight into state-of-the-art AI tools and techniques in Project Management.

Proposal/Contract/RFP Documents in Project Life Cycle

In general, any project when initiated must propose a set of objectives and requirements to be fulfilled before its development commences. Usually, these requirements are taken from various stakeholders and consolidated in one place in a form of a text document. Documentation is an indispensable asset across various business domains including construction, legal case, etc.

Many companies are awarded project contract based on the proposal received from them for the tender floated by the agency. The company bidding the tender needs to understand the requirements given in the tender including the prequalification criteria and then prepare the response in the form of proposal which includes technical solution, time duration for the completion of the project, costing, roles of various people etc. To prepare such a proposal, the team looks at various aspects such as the work done by the company in that domain, existing components,

The initial phase of project development includes sharing of the documents between the developer and the client and sometimes other interested parties or stakeholders. These documents define the basic objective and goals of the projects, including start and end dates, etc. We often call such documents proposal documents. To be more specific, a formal approach is used by organizations called a procurement process. In this process, there are 2 parties involved [24]

1. Company /organization that is interested in the development of the project
2. Vendor or supplier who can develop the project.

The company or organization that is interested in project development is one that requires projects, so it issues a proposal document known as an RFP. It evaluates the needs and develops a list of criteria, which is then turned into questions and sent to a number of vendors who offer the goods and services that the buyer is looking for.

Vendors /Suppliers/Bidders: The vendors study the proposal which comprises of project's technical requirements, roles and responsibilities, service level agreements, timelines format, and input required about the company and the proposed solution and verify if they have technical and financial expertise, and send their responses and return it to the buyer whereupon the vendors' responses are scored. As proposal documents are long verbose textual documents, project team effort is required to understand and categorize information present inside the documents. This is further used to create the responses. Now text classification and categorization can be performed at both the parties' ends, the company that creates the proposal and receives the responses and the Vendor who submits the responses. In either situation reading, analyzing and understanding, and most efficiently analyzing the data is essential. It requires a lot of human intervention and time and thereby affects the speed at which work is performed, affecting the overall project life cycle.

Table 1: Broad Areas of AI, their Description, and Applications in Project Management

Study Reference	Broad Area of AI and Description	Application in Project Management
[11]	Machine Learning, The paper focuses on effort and duration estimation models using intelligent data preprocessing. Three machine learning models including SVM, MLP, GLM,	Software project effort and duration estimation.

	and ensemble averaging, were used. Data Set Used- ISBSG dataset	
[12]	Machine Learning, The suggested method can help staff members with routine PM duties such as resource assignment, and work time estimation	Common PM tasks including resource allocation, task assignment, and task time prediction are planned, monitored, and carried out.
[13]	Theoretical AI approaches, This article explains how AI can still make a substantial improvement to PM. Significant areas of focus for AI application in PM are listed as follows: They are listed in the following order: business domain, application domain, predicted benefits, AI solution, use case, and AI system for PM.	Generic Use Cases for PM
[14]	Machine learning, Business Process Management, Optimization of process management	By using machine learning to identify trends as the process moves through the workflow, business process data improved decision-making.
[15]	Machine Learning, This investigation looks	Project Planning, Scheduling
[16]	Machine Learning, The goal of this study is to ascertain how AI can be used to manage the time and costs of a construction project in India.	Construction projects
[17]	Automated systems, Internet-of-Things and AI, Human-machine integration using AI approaches	Strategy, Relationship Marketing
[18]	Machine Learning, The goal of the study is to examine how business process management (BPM) will change as a result of digital transformation in relation to machine learning and artificial intelligence (AI).	Artificial intelligence and machine learning for business process management.
[19]	Machine Learning, Deep Learning, The topic of intelligent automation is covered, along with strategies for how top businesses can use it to increase productivity dramatically.	Credit card fraud management use case, Project Calling, Project Preparation, IA uses cases in the telecommunication industry, IA use case in public and government sectors
[20]	Machine Learning, A learning approach that can generate concurrent workflow	workflow management systems
[21]	Artificial Intelligence, Impact of AI on the economic growth of the business,	Robotics, artificial intelligence, business models, automation.
[22]	Deep Learning, Machine Learning, examined the available AI approaches and there.	Cases related to Tenders, engineering, and Design, Operations
[23]	Software Engineering, Artificial Intelligence, Agile teams can benefit from AI as a unique	Project backlog refinement, risk mitigation

	accelerator, which can assist raise project success rates.	

It has been observed that the selection of types of proposals also affects the overall project performance. [25] Irrespective of the domain, it is observed that tools and techniques related to artificial intelligence are used to extract useful information from the proposal documents. A need for an automated system thus has been addressed in the following work. In the next section, we will discuss the role of artificial intelligence specifically for proposal documents.

Techniques of Information Extraction, Classification, Question-Answering, Summarization for Proposal Documents

In this section, we will discuss the AI techniques related to Information Extraction, Classification, Question-Answering, and Summarization used in the literature for document processing in general, in the subsequent part of the paper we will discuss related techniques used for proposal documents(RFP) which is our primary focus.

A. Text Classification

The text classification job is utilized in NLP applications such as sentiment analysis, news categorization, question-and answer systems, etc. This technique assigns labels to the textual data and draws meaningful context between various sentences. As discussed it can be implemented either using rule based techniques or machine learning-based techniques [26]. Under circumstances where the dataset is extensive and, dependency is on designing features for the dataset, Neural approaches are also explored [27], [28], [30].

This technique is used for classifying the contents of a document, including its text portion, into different categories. The basic aim of text classification is to extract the meaningful data hidden in the unstructured text. Due to the complexity of the text information involving the variability in the context of the usage of a given word or a phrase, this classification exercise becomes a challenging one even for a human to extract and make sense of this unstructured data.

Text classification can be used at several levels, such as document, paragraph, sentence, and sub-sentence levels. [26], [29]

The need and importance of text classification can be understood from the point of view of an organization that is dealing with a huge amount of data. This data may contribute to several business proposals, business logic, and decisions. Due to its unstructured layout of data, organizing and analyzing becomes very hard and time-consuming.

1) Expert systems*:* Rule-based methods use predefined rules to classify text into different categories, as rules are required to be defined therefore this requires domain knowledge. Rule-based systems require both a set of facts or a data source and a set of rules for altering that data in order to operate. These rules are often known as “if statements” in some circumstances since they frequently function along the lines of “IF X happens THEN execute Y.”

An expert system is a type of computer program designed to handle complex issues and provide decision-making capabilities comparable to those of a human expert. By employing the reasoning and inference rules in accordance with the user’s questions, it accomplishes this by extracting knowledge from its knowledge base. How well an expert system performs is based on the knowledge that it has stored in its knowledge base. With more knowledge stored in the KB, the system performs better.

Several references in literature are available where expert systems are used for document processing [31], [32], [33]

Although the rule-based approach is simple, precise yet a lot of dependency on the heuristics which are created. In situations where the level of unknown and/or variance is low in the documents, an expert system based on rules is likely to yield good results. Rules-based methods, however, fall short when dealing with the diversity of document kinds or differences within document types that necessitate a significant amount of study and development. Machine learning for data extraction, on the other hand, is best suited, once more, to projects where the targeted documents have a high and/or unknown level of volatility and where a common approach is to cover a higher percentage of documents.

2) Machine learning–based approaches: In the field of artificial intelligence, Machine learning (ML) systems are those that are able to learn using algorithms and statistical models, this enables the system to identify and discover patterns in the data. this further enables the businesses to have a better insight into the data and subsequently better decision-making [34]. ML has different types of algorithms and therefore 3 different categories of algorithms are known based on the outcomes which include Supervised, Semi-Supervised, Unsupervised, and Reinforcement algorithms.

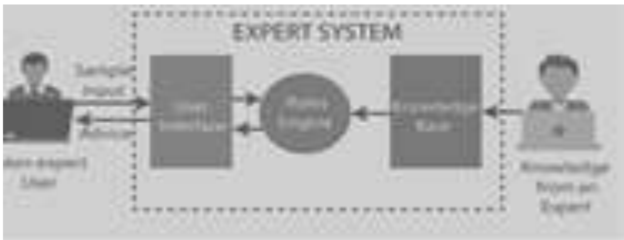


Fig. 1. Expert system

Supervised learning algorithms: Finding a mapping function to link the input variable (x) with the output variable is the goal of this kind of ML method (y). In this, machines estimate the outcome based on training data that has been carefully" tagged" for them. The labeled data refers to the input data.

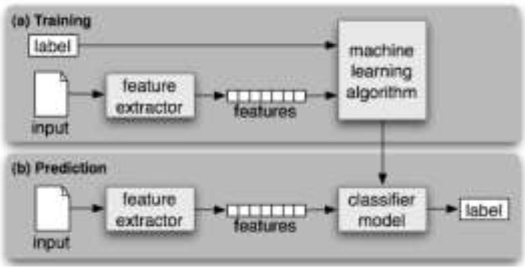


Fig. 2. Supervised learning

Unsupervised learning: In unsupervised learning, models are trained on sets of unlabeled data and then given free rein to use the data as they see fit.

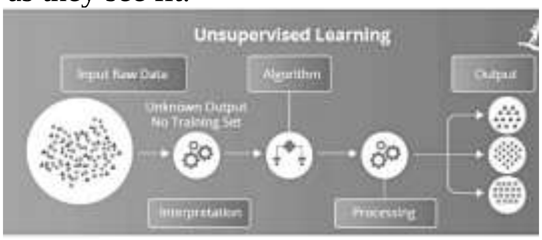


Fig. 3. Unsupervised learning

Using both labeled and unlabeled data, semi-supervised learning creates a function. Reinforcement Learning- It is a form of machine learning technique where a computer program or intelligent agent interacts with the environment and learns how to behave in it.

Machine learning–based approaches learn to classify the text based on observations of data. This is achieved by transforming the text contained in the documents from unstructured form to normalized data, appropriate for further analysis. These Machine Learning algorithms operate on a self-learning basis where they learn automatically from experience without any manual intervention and help efficiently solve complex problems.

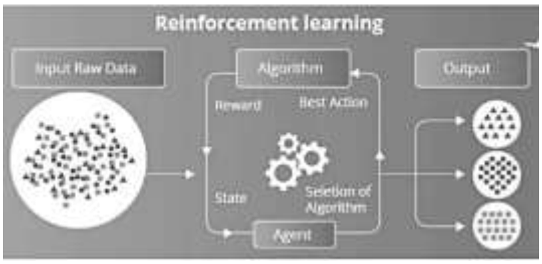


Fig. 4. Reinforcement Learning

Different machine learning algorithms are used for text classification, including supervised learning and unsupervised learning. Naive Bayes classifier, Logistic regression, K-Nearest Neighbor classifier, Decision Tree, and other algorithms are used in supervised learning, while K-means clustering and Polynomial regression are used in unsupervised learning [27].
Metric for text classification are listed below: when the models are created their performances are evaluated using different metrics, some of which are listed below. Detail explanation of the metrics is available in the given table. 1. Precision /Recall/F1 Score 2. Accuracy 3. Error Rate 4. Exact match 4. Mean Reciprocal Rank 5. Mean Average Precision 6. Area Under the curve

3) Classification of Functional and Nonfunctional Requirements: Analyzing requirements acts as a first step to a better understanding and overall success of any project. Considering software projects, requirements are further classified into functional and non-functional. It has been observed that not just functional requirements but non-functional requirements also play a very crucial role in the successful execution of projects [2], [3].

B. Deep learning based Text Classification

A branch of machine learning called deep learning mimics human understanding. The deep learning models have multiple layers of transformations; the lower layers may reflect minor features in the data, while the higher ones may represent more abstract ideas. Consequently, deep. Deep learning is the learning of multiple layers of transformations in order to learn multi-level representations.
Usage of deep learning spans various application areas in a different set of industries, which includes Self-driven cars, Healthcare, and Language translations. Our focus in this survey is to identify various deep learning techniques suitable for text classification and understand the variations amongst them. Also, we highlight a few of the major application areas where text classification has been used using deep learning. [27]

C. Question and Answering

A question-answering system generates a response to a question posed in natural language. Reading and comprehending a text by machine is a difficult task, hence it acts as a potential topic to explore in research.
To break down the Q/A process into simple steps, it is observed that we have a query processing module to analyze the question, classify the type of question to understand the context and formulate of question in a vector to be fed as an input into the next module which is document processing module which based on similarity and relevance filter out the most suitable documents which can give can give an appropriate response. This is further sent as input to the answer processing module which heuristically identifies the closest answers and then evaluates them using evaluation metrics like F1 scores, precision, etc.

Application of Artificial Intelligence for Project Proposal Documents

The tender/proposal contains a mix of various information that may include business, legal, managerial or technical information, and financial. The information inside the document is nested to give details of the technical requirements defining the scope of work along with contractual and legal information. The requirement specified in the documents is also interdependent on each other. Human expertise is required to understand the requirements and then assign them to their respective subsystem in a systematic way. It is observed [57] that precision in the extraction of requirements has a huge impact on firms' finance and reputation too.

Given the nature of tenders placed across various domains in the industry, it was observed that the role of the required documents such as the RFP is very critical in clearly specifying the project requirements. If we generalize the subject matter, essentially these are documentary information of projects in hand, sometimes instead of projects, we can term it as cases (e.g legal scenario). However, it is observed that despite of huge variation among these domains pain point is common. That is, extracting meaningful information from the documents. Using Artificial Intelligence techniques, it is observed that this task can be done in a much faster way. The many methods for extracting requirements from documents—including natural language processing, machine learning, and deep learning—are addressed in the section that follows.

Analyzing requirements acts as a first step to a better understanding and overall success of any project. Requirements for software projects are further divided into functional and non-functional categories. It has been observed that not just functional requirements but non-functional requirements also play a very crucial role in the successful execution of projects [2], [3]. The area of artificial intelligence's applications with respect to proposal document processing is explored. Project proposal contains a mix of various information that may include business, legal, managerial or technical information, and financial. The information inside the document is nested to give details of the technical requirements defining the scope of work along with contractual and legal information. The requirement specified in the documents is also interdependent on each other. Human expertise is required to understand the requirements and then assign them to their respective subsystem in a systematic way. It is observed [57] that precision in the extraction of requirements has a huge impact on firms' finance and reputation too.

Given the nature of proposal documents placed across various domains in the industry, it was observed that the role of the documents such as the RFP is very critical in clearly specifying the project requirements. If we generalize the subject matter, essentially these are documentary information of projects in hand, sometimes instead of projects, we can term it as cases (e.g legal scenario). However, it is observed that despite of huge variation among these domains pain point is common. That is, extracting meaningful information from the documents. Using Artificial Intelligence techniques, it is observed that this task can be done in a much faster way.

The study identified papers that used the application of Artificial Intelligence including NLP, ML/DL to obtain information from proposal documents. The study specifically focused on requests for proposal documents (RFP) also.

Table [2] summarizes the Application of AI techniques on proposal documents.

Various stakeholders including consultants, management, sales, module specialists, and software architects are involved in comprehending the RFPs, thus manually doing it poses many challenges like reading and understanding the large document, obliging the tight time frames also there is minimal or no interaction between customers and suppliers.[36] suggest an information model for segregating and depicting roles and responsibilities required in the RFP process from the suppliers' end including they also suggest a need for a knowledge base that acts as a repository of information about existing system features, cost incurred, and risk estimation. Although this work sets background work, it doesn't discuss the tools and techniques required to do it automatically. In the literature the work having reference to automation for proposal documents (particularly RFP), it was observed that work was distributed among 3 categories that are in effect interrelated to each other. This covers text classification, information extraction/retrieval, and question and answer. The following discussion is on the work in each of the categories and the same is summarized in Table 2

For the proposal documents, the literature suggests a variety of work with different intentions [37] exploring the basic task of extracting technical terms to identify emerging trends from the US govt. RFPs using machine learning and natural language processing. The Paper emphasizes the use of transfer learning to extract meaningful information from the proposal documents. Using the transfer learning approach existing models which are extensively pre-trained are reused. This acts as a starting point for solving the problem hence resulting in improved performance for the new task. Transfer learning is often known as fine-tuning, it is used in deep learning approaches resulting in learning time being faster as the model has already been trained with huge training data set, unlike the traditional machine learning model where performance relies on our training data. For this purpose, [37] using Hugging Face ‘transformers library and SpaCy to perform Name Entity Recognition [38], [39]. A neural network with an encoder and a decoder is referred to as a transformer model, it can learn the context and relationships in data and thus semantics of a long sentence in forward and reverse directions. This is also known as the attention mechanism. There are various transformer models like GPT and BERT whose references are also made in several parts of our survey. The Study presented in [40], [45] focused on the task of requirement classification from the proposal document of the construction domain. They highlighted that previous models and frameworks were domain-specific and hence the proposal for construction requires separate attention. The research [40] focuses on classifying text into categories of requirement and non-requirement. The authors demonstrated the classification using rule-based and machine-learning approaches. Rule-based classification techniques focus on the construction of If-then rules, in the professed work, requirement keywords were identified and statements in the text were assessed based on the rules. Although, rule-based approaches are simple to implement with smaller data it becomes very tedious to create rules of a large corpus. The Machine Learning Classification approach in [40] starts with preprocessing which included lowercasing, stop words removal, unique words removal, tokenization, lemmatization/stemming, and Parts of Speech (POS) tagging. Following this Sentence Embedding using Word2Vec and Bag-of-Words Model is used to enable the ML algorithm to accept words in a numerical format. This was further followed by 4 supervised ML algorithms including Naive Bayes, Support Vector Machine, Logistic Regression, and Feedforward Neural Network, SVM resulted in the best classification accuracy for their dataset. this work is further expanded in [45] and its implementation is shown in the context of the design-build project delivery method to draw more useful information from the proposal documents which includes extracting requirements and thus categorizing them into different categories like design, operation, and maintenance, etc. Before Model, Training [45] The vocabulary is condensed via feature selection utilizing chi-square, mutual information, and recursive feature removal. In order to increase the final model’s robustness and generalizability, the study also looked into four ensemble techniques: random forest (RF), bagging, boosting, and voting. A different take on the text classification for proposal documents is taken by [42] by extracting meaningful information about the projects from RFP response documents using AI-based techniques. For ML model training Maximum entropy classifier is used which is an exponential model. The MaxEnt is based on the Principle of Maximum Entropy, and it chooses the model with the highest entropy among those that fit our training data. A wide range of text classification issues, including language identification, subject classification, sentiment analysis, and others, can be resolved using the Max Entropy classifier. Other text classification work for proposal documents includes [46], [47], [48], [43] each work emphasized classifying requirements in the proposal documents. [46] explains how to extract and categorize requirements from software engineering contracts automatically using bidirectional encoder representations from transformers (BERT). Classified requirements as governance specific and architecture-specific.ML algorithms used for classification are – Naive Bayes, Random Forest, and Support Vector Machine (SVM). Imbalanced datasets were handled using hyper parameters like class weight etc. [10] Here, a unique Bi-LSTM-based regression model is proposed, and the deep learning technique greatly outperforms the Random Forest model. The author's take on this work is to provide a score on words and phrases identified from the RFP response document and classify them further into enablers and disablers which will help in creating better RFP responses. Apart from text classification on proposal documents, it was observed from the survey that work is taken a step ahead to generate answers for the questions extracted and interpreted from RFP documents. [49] proposes a meta-model to generate an RFP response draft. similar to this idea [50] SCM, soft cosine measure to identify the matching response to a question present in an RFP document. Thus semantically related questions and answers can be used to build an RFP response document. As in the context of the proposal document literature focuses on various NLP and ML techniques. Following is a brief discussion on other sentence embedding techniques and Machine Learning algorithm techniques present in literature. This discussion provides insights into various mentioned algorithms and techniques used by researchers to perform classification. Sentence Embedding/ Representation in NLP includes representing sentences in n-dimensional vector space. It evolved with various methods like Bag of Words (BoW) where sentences are marked as 0 or 1 based on whether they occur in a sentence or not however, the order of sentences and semantics were not considered. Term Frequency-Inverse Document Frequency (TF-IDF) emphasized how crucial a word is in a document Word2Vec/Doc2Vec, LSTM, Attention, and Transformer, BERT are neural embedding used to establish better semantic information between words of sentences. Word2Vec is a predictive model, words in documents are identified and plotted in a multidimensional vector space, and similar words tend to be closer to each other. It is further divided into two architectures: Continuous Bag of Words (CBoW), in which the model forecasts the current word from a window of nearby context words, and Continuous Skipgram, in which the model employs the current word to forecast the window of nearby context words. Doc2Vec computes a feature vector for each document in the corpus, as opposed to Word2Vec, which computes a feature vector for each word in the corpus.

Study Reference	Business Domain	Technology Used	Corpus	Intent	Results	Focused Task
[49]	Business Units of finance and accounting	NLP for Query Formulation, Information retrieval from Knowledge Base	15 RFP's from 6 Business Units	RFP Response Generation	76% Mean Query Precision86% Mean Query Recall	Question and Answering
[37]	USGovt RFP Technology Domain	Automated Emerging Trend Detection System(ETDS), In NLPword2vec algorithm (vectorization of words), Deep Learning Model, Uses method known as transfer learning or fine-tuning. (USe Pre-trained Models) Named Entity Recognition (NER) using transformer architecture, SPacy Models are used, precision was identified	Department of Defense (DoD) and Department of Homeland Security (DHS) corpus of RFP documents Identifying new technology trends requires the extraction of technological terms from RFP documents.		Information retrieval and classification	
[40]	Construction Domain	NLP and ML algorithms used to build a classification model. ML algorithms used: Naive Bayes, Support Vector Machine,	Labelled data set was prepared of 7 documents,1787 statements	Classification of text in requirements and non-requirements	In terms of Accuracy, Precision, Recall, and F1 score, SVM performs better than other models.Results from unigra	Text Classification

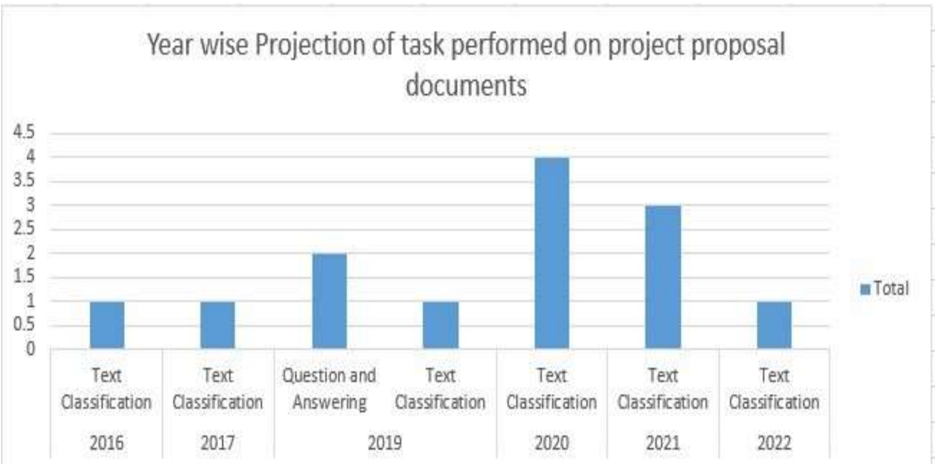
		Logistic Regression, Feedforward Neural Network.			ms are superior to those from higher n-gram features.	
[41]	Railways Contract Documents	T2K (Text-To-Information) is an expert system that uses an ontology learning system to automatically infer ontological knowledge from texts. The phrase ”extract module” was used.	Proposal/tender Documentwith 6000 requirements	Requirements identification and assignment to subsystem	Out of the 6000 Technical Requirements the program detected, 543 make reference to technical standards whose abbreviations are listed in the Knowledge Base.Precision 98%; Recall 98 %.”	Text classification
[42]	RFP of IT ser- vices domain	Model-training for ML It employs the maximum entropy (maxent) classifier. NLP Engine is developed using OpenNLP.	16 RFPs	Extract and categorize information pertinent to the project beginning from the RFP answer	Classification Accuracy=76	Text Classification
[50]	Independent dom ain	NLP Soft cosine Measure and Fast Text word embeddings	4600 Q&A	Identifying responses for questions in proposal documents	SCM performed better than WMD	Question and Answering

[45]	Construction Domain	Machine Learning algorithms- NB, SVM,LR,KNN, DT,CNN, RNN(LSTM), RNN(GRU)	3000 clauses in a proposal document	Extraction of specific scope for subcontractors and classification of design-build needs from the bid document into three specified categories: design, construction, operation, and maintenance	SVM and weighted majority voting outperformed other models	Text Classification
[46]	20 SE contracts from 9 application domains viz. healthcare, automotive, finance, banking, etc.	NLP and ML Techniques, traditional ML techniques, BiLSTM, BERT	20 SE contracts from 9 application domains,18,614 sentences extracted,20 to 250 words in each sentence.	Classified requirements as governance-specific and architecture-specific	39.73% are architectural requirements and 6027% are gover-nance specific	Text Classification
[47]	Insurance Indus-try	ABi-LSTM and a regression model based on transformers	1300 RFP responses	analyze and Score RFP responses	The random forest model is outperformed by the deep learning approach.	Text Classification
[48]	Railway Domain	Non-AI-based techniques- Patternbased approach (PABRE metamodel)	6 RFPs, 17,556 requirements	Classification based on real requirements or prose	Usefulnessof Model evaluated	Text classification
[43]	Construction Domain	Supervised machine learning approaches for classification model	2634 statements	Classify requirements and reduce reading time and improve quality	Logistic Regression gave the best accuracy at 94.12	Text classification

[44]	IT service provider company	Classification model using unigrAM FEATURES	300 Requirements, 30 + test data set	shows the effectiveness of the tool for intelligent requirement extraction and analysis	”RFPCog as an interactive and explorative tool for analysis, refinement, and browsing of	Text Extraction and Mapping

Table 2: Application of Artificial Intelligence for Project Proposal Documents

This is further divided into the Distributed Memory version of the paragraph vector (PV-DM) and the Distributed Bag of Words version of the paragraph vector (PV-DBOW). LSTM, or Long Short Term Memory Network, is a recurrent neural network that overcomes the drawback of RNN, or Recurrent Neural Network, by remembering the long-term dependencies of previous work. The cell state in LSTM holds the system’s state over time thus helping in embedding in a better way as better semantic relationships can be derived. To Derive better context and relationship in a sentence Attention and Transformers are readily used, BERT, Bidirectional Encoder Representations from Transformers, as the name suggests BERT accepts unlabeled text and starts encoding bidirectional representations to establish context on both the right and left context.
Figure 1 and 2 represents year Wise application of AI on Project Proposal documents, domain, and Year Wise application of AI on Project Proposal documents.



Industrial Engineering Journal
ISSN: 0970-2555
Volume: 52, Issue 10, No. 1, October: 2023

Fig. 5: Year-wise application of AI on Project Proposal Documents

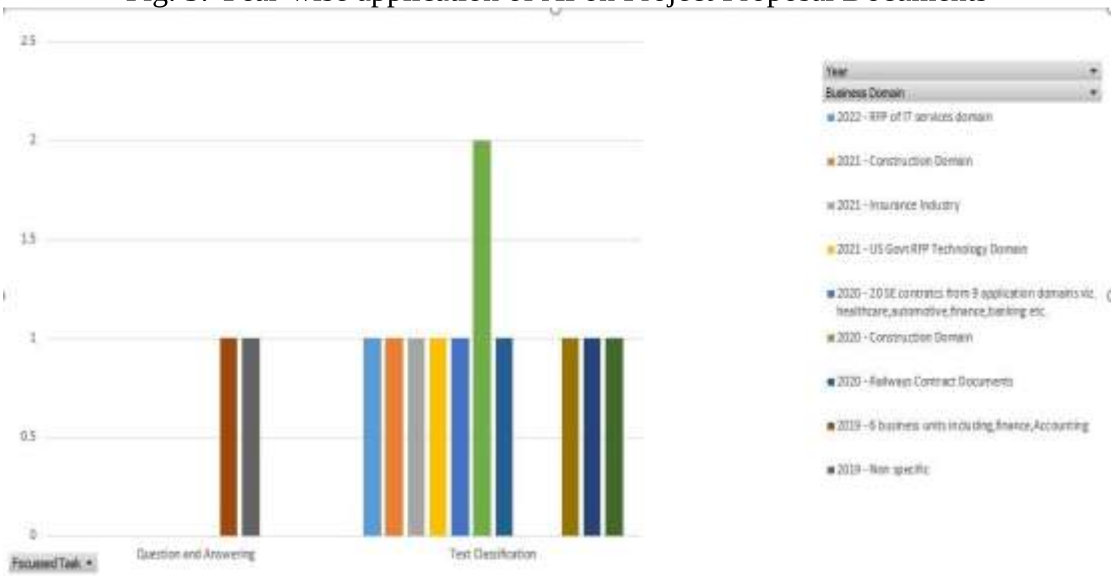


Fig. 6: Domain and Year-wise application of AI on Project Proposal Documents

A. Related Work

In the last few years, surveys related to requirements extraction from text documents and their classification have been conducted. This section provides an overview of related research and explains the key distinctions between this study and previous surveys.

[51] reported a survey of Natural language processing (NLP) techniques. To reach the objectives of requirement prioritizing and categorization, the author examined 27 research looking into the use of NLP in extracting software needs. The survey does not overlap with this work because it does not use any methods based on deep learning or machine learning.

[52] surveyed state-of-the-art of text classification. The study enumerated baseline elements of text classification and its associated techniques. The survey has qualitatively analyzed text classification common techniques and quantitatively observed common trends in text classification. As the discussion is generic therefore it does not overlap with our work, instead gives insight into evolving text classification trends.

[53] conducted a systematic review of 72 primary studies. The work was focused on text classification of clinical data.

[54] conducted a thorough analysis of machine learning for automated requirement classification. As they highlighted several machine learning techniques for text classification and reviewed publications based on the classification of functional and non-functional needs, this review work is helpful in our study. However, in contrast to our approach, which focuses specifically on proposal documents, they utilize the TERAPROMISE database.

[55] conducted a systematic literature review on using Machine Learning Algorithms for identifying the requirements. They focused their work on the stack overflow platform.

[56] conducted a systematic literature review to identify the classification of quality requirements only.

However, to the best of our knowledge exclusive focus of a survey on automatic text classification of proposal documents (RFP) is not visible in the literature. Therefore, this survey focuses on the area of automatic processing of proposal documents using Artificial Intelligence techniques.

Exploring Different Domains Using Automated Classification for their Requirement Documents

The possibility of classifying and comprehending the requirements for the tender documents or RFP has been explored for various domains. Given the unique nature of requirements in form of usage of text information in each domain, such as the legal or technology domain, even a sizeable amount of evidence on text classification present in the literature appears inadequate. This variance in the standards for sizing the requirements for formulation of a common scale has posed a long-standing challenge to the research community. In this section, domain-wise classification of work is presented with respect to proposal documents at the early project management stage.

A. Legal Domain

The legal domain has extensive usage of documents involving complex terminology and technical phrases. To extract meaningful information from these documents is time and effort intensive. Sequential classification is used to extract clauses from the legal contract by identifying paragraph boundaries in the raw text data in order to extract semantically useful information [59]. In any business domain, the text contained in the relevant legal documents involves crucial information related to various entities that may be a part of data privacy protection. Such entities are identified in legal documents using information Retrieval, Information Extraction, Natural Language Processing (NLP), and Ontology [60]. Further, to classify the text contained in the legal documents, text summarizing of these documents using context free grammar rules was used to extract information related to various clauses [61] [62] studies legal text to extract contract elements using machine learning techniques including logistic regression and support vector machine.

B. Railway Domain

During the tendering process when bid-related documents (RFP-Request for proposal are assessed, the process is time intensive and usually dependent on human expertise. [63], here author focuses on understanding the patterns in the requirements and also classifies the requirements into definitive requirements or prose. A similar idea is automatically implemented using an expert system [64] to automatically extract the requirements. The extracted requirements are further assigned to a specific enterprise system; a scoring system is also suggested to specify the significance of a requirement.

C. Construction Domain

One of the prominent domains where NLP and Machine learning are extensively used to understand the requirements in the document is a construction domain [65].

D. Technology Domain

Although technology is being used in every domain, but this separate subsection considers those proposal documents which especially are used for the development of IT projects. The authors [66], analyze RFPs of IT projects using NLP and machine learning and extract information on emerging technologies that are used for the development of software projects.

E. Healthcare Insurance

Healthcare proposal documents analysis are also identified in literature. [66] Proposes an automated framework to analyze RFP responses, a scoring methodology using machine learning Bi-LSTM based on the regression model. This model explores the responses of RFP documents that enable scores of the section which impacts positively and negatively on the proposal response.

Open Challenges in Automated Classification Research

Understanding data from unstructured documents consumes a lot of time for an organization. AI-based techniques can help in easing out this process. However, these techniques still possess certain challenges. In this paper, we are specifically looking at a type of unstructured documents used in project management at a very early stage of project development namely proposal/contract documents. The challenges of automated classification of contract documents are co-related with the challenges of classifying unstructured text data. Techniques for text classification have been explored, also its implementation in various fields, despite this fact the area of classification of requirements in the bidding/proposal stage faces certain challenges. Some challenges which need to be addressed so that many organizations can be benefited are:

Industrial Engineering Journal
ISSN: 0970-2555
Volume: 52, Issue 10, No. 1, October: 2023

(a) Data Set Requirements: In order to take better decisions for the growth of any organization, One of the important phases is to understand and accept the correct type of projects. To do so a huge volume of data past and present is required to take the train and test the system effectively then only we can step forward in the scoring of responses based on historical responses too.

(b) Domain and expert dependency - A huge dependency from domain-specific expert validation is to be combated, however for that a large knowledge base per domain is to be developed.

(c) Standard format across domains and organizations: There is no standard format for the proposal documents viz. RFP, RFQ, etc. Other challenges in the textual documents are existing taxonomies, keywords, and acronyms to ensure we work on clean data so as to achieve document classification and summarizing.

Proposed Framework

The Research questions addressed here helped us in understanding the concept, its relevance, and its usage in various domains. Based on our study we propose a framework for the software project proposal documents (RFP) of the e-governance domain.

A. e-governance*:* e-governance signifies the application of information and communication technology to provide various services offered by the government. In India initiative of e-governance began around the 1970s with the department of electronics and further with the National informatics center. In 2006, the National e-governance plan by the department of electronics and IT came up with an umbrella program to prepare India for a knowledge-based transformation with 21 mission mode projects including healthcare, education etc. One such mission mode project is e-procurement; whose objective is to provide a one-stop for all services related to government procurement. Now various components are there underneath this which includes e-purchasing, e-tendering, e-ordering, e-invoicing, etc. This lead to a simple, transparent system for all the stakeholders involved which included purchasing department, accounting department, end-user placing the orders, etc.

In the e-tendering system, a tender is used which can be described as an offer to perform some task at predetermined prices. These tenders are issues in the form of Request for proposal(RFP), Request for Quotation(RFQ), Request for Information(RFI) documents. Requests for proposals are basically an invitation for suppliers often through a bidding process to submit a proposal on a specific commodity or service. These e-governance-related RFP documents become the object of our interest as the input data sets. RFPs outline the requirements of purchasers, including their contractual and legal requirements, technical and functional requirements, and bid procedure and commercial specifications. Under the life cycle of the e-procurement process when RFP is published by a company that requires services, various vendor provides responses to stated requirements in the form of question and answers. Based on the responses received a selection of vendors takes place. Thus, it becomes an utmost significance for a vendor to create a response that caters to the technical, and nontechnical requirements, and also the suggested cost is low. Which may lead to favorable chances of winning the bid. Hence, we propose a framework for the e-governance domain. which can extract technical, economical and various other important information from these RFP documents and subsequently assess the managerial and economical benefit of the same w.r.t to the organization.

B. Steps of Proposed Framework for e-governance domain

We propose a 2-dimensional framework, Under the first dimension, an AI model is used to extract information that includes

1. Collection of data set- RFP documents of e-governance domain. As the e-governance domain in itself is huge, it comprises several subdomains. Selection of a particular sub-domain RFP.

2. Data Pre-processing and Validation

Industrial Engineering Journal
ISSN: 0970-2555
Volume: 52, Issue 10, No. 1, October: 2023

3. In order to find the algorithm that provides the most insight from the data, model building is frequently an iterative process that involves many rounds of testing. Deep Learning Model selection is required to retrieve quick and accurate information extraction from documents, it will be used to assess how accurately information is extracted from RFPs.

4. Further information filtering system will be built to give preferred responses to a set of questions appearing in the RFPs.

Under the second dimension of the proposed work, the predicted output will be studied to understand the following points:

1. How easily we can create a response to recurring or similar requirements?
2. How much cost and time are saved in this process which leads to quick response generation and how it might contribute to winning the bid?
3. How well a skillset can be shared to generate the solution response for an RFP.

Limitations of The Study

This paper presents a review of the application of Artificial Intelligence to fetch information from proposal documents. The authors have reviewed (Put number) papers that have used various NLP and ML/DL techniques and other models to extract the requirements out of proposal documents. An attempt is made to look at the literature critically with respect to sample size, accuracy value, and algorithms used in the study and the findings. The work studied for the survey is summarized in tables. The limitation of the work is that the author has tried to retrieve relevant data regarding the domain however, the uncertainty of not including all the possible related searches is a possibility. Also, the proposed framework for the e-governance domain is still in the planning phase, experimental setup and findings are future plans of the work.

Conclusion

Understanding and extracting requirements for a project is an integral task for any project management team. In this paper, we assessed the current state of the art related to the automatic processing of requirements from proposal documents for project development. We analyzed data found in the literature related to the automatic extraction of requirements from requirement documents at the proposal stage of the project management life cycle. We characterized the work done on different business domains and identified AI-based techniques used by each of them. As the required identification and understanding have an acute effect on project development, we closely reviewed the related engineering practices using natural language processing and machine learning. We observed that the automation task so far has been specific to the business domains, which gives us a clear indication that many other domains viz. Banking, government, etc can be explored in near future. Extraction of information for such documents will not only ease the process of making early decisions without human intervention but it will also help in making strategic decisions and analysis of the nature of projects and requirements coming over the year which results in deciding how we can distribute and utilize our resources too. Literature indicates that developments in requirements engineering pertaining to e-governance projects have not been explored. This survey, in our opinion, can act as the main investigator for understanding the need of characterizing requirements for e-governance projects and also understanding the effect on the project development life cycle.

As future work, we plan to implement the reviewed concepts and strategies in industrial settings for e-governance projects, and hence visualize the effect of automation and study the impact of the same on the overall project development cycle.

References

[1] Kanakaris, Nikos and Karacapilidis, Nikos and Lazanas, Alexis. (2019). On the Advancement of Project Management through a Flexible Integration of Machine Learning and Operations Research Tools. 362-369. 10.5220/0007387103620369.
[2] R. R. Maiti and F. J. Mitropoulos,"Capturing, eliciting, predicting and prioritizing (CEPP) non-functional requirements metadata during the early stages of agile software development," SoutheastCon 2015, 2015, pp. 18,doi:10.1109/SECON.2015.7133007.
[3] Maiti, Richard and Krasnov, Aleksandr and Wilborne, Deanna. (2018). Agile Software Engineering and The Future of Non-Functional Requirements.
[4] William Inmon, Anthony Nesavich, Tapping into unstructured data: integrating unstructured data and textual analytics into business intelligence,https://dl.acm.org/doi/abs/10.5555/1408345
[5] F. Hujainah, R. B. A. Bakar, M. A. Abdulgabber and K. Z. Zamli,"Software Requirements Prioritisation: A Systematic Literature Review on Significance, Stakeholders, Techniques and Challenges," in IEEE Access, vol. 6, pp. 71497-71523, 2018, doi: 10.1109/ACCESS.2018.2881755.
[6] Hendrik Meth, Manuel Brhel, Alexander Maedche, The state of the art in automated requirements elicitation, Information and Software Technology, Volume 55, Issue 10,2013, Pages 1695-1709, ISSN 0950-5849, [12] Kanakaris N., Karacapilidis N. and Lazanas A. (2019).
[7] G. Auth, J. Johnk and D. A. Wiecha, "A Conceptual Framework¨ for Applying Artificial Intelligence in Project Management," 2021 IEEE 23rd Conference on Business Informatics (CBI), 2021, pp. 161-170, doi: 10.1109/CBI52690.2021.00027.
[8] R. Ilieva and Y. Nikolov, "AI Integration in Business Processes Management," 2019 International Conference on Creative Business for Smart and Sustainable Growth (CREBUS), 2019, pp. 1-4, doi: 10.1109/CREBUS.2019.8840086.
[9] A. Alshaikhi and M. Khayyat, "An investigation into the Impact of Artificial Intelligence on the Future of Project Management," 2021 International Conference of Women in Data Science at Taif University (WiDSTaif), 2021, pp. 1-4, doi: 10.1109/WiDSTaif52235.2021.9430234.
[10] K. Rathod and A. Sonawane, "Application of Artificial Intelligence in Project Planning to Solve Late and Over-Budgeted Construction Projects," 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), 2022, pp. 424-431, doi: 10.1109/ICSCDS53736.2022.9761027.
[11] Sandra Maria Correia Loureiro, Joao Guerreiro, Iis Tussyadiah, Artificial intelligence in business: State of the art and future research agenda, Journal of Business Research, Volume 129, 2021, Pages 911-926, ISSN 0148-2963
[12] Joachim Herbst, A Machine Learning Approach to Workflow Management,ECML '00: Proceedings of the 11th European Conference on Machine LearningMay 2000 Pages 183–194
[13] https://vtindia.in/the-impact-of-artificial-intelligence-on-business/
[14] Mishra, S., Tripathi, A.R. AI business model: an integrative business approach. J Innov Entrep 10, 18 (2021). https://doi.org/10.1186/s13731021-00157-5
[15] H. K. Dam, T. Tran, J. Grundy, A. Ghose and Y. Kamei, "Towards Effective AI-Powered Agile Project Management," 2019 IEEE/ACM 41st International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER), 2019, pp. 41-44, doi: 10.1109/ICSENIER.2019.00019.
[16] Paech, B., Heinrich, R., Zorn-Pauli, G., Jung, A., Tadjiky, S. (2012). Answering a Request for Proposal – Challenges and Proposed Solutions. In: Regnell, B., Damian, D. (eds) Requirements Engineering: Foundation for Software Quality. REFSQ 2012. Lecture Notes in Computer Science, vol 7195. Springer, Berlin, Heidelberg.

[17] Mohammad Suprapto, Hans L.M. Bakker, Herman G. Mooi, Marcel J.C.M Hertogh, How do contract types and incentives matter to project performance? International Journal of Project Management

[18] Kam K.H. Ng, Chun-Hsien Chen, C.K.M. Lee, Jianxin (Roger) Jiao, Zhi-Xin Yang, A systematic literature review on intelligent automation: Aligning concepts from theory, practice, and future perspectives, Advanced Engineering Informatics, Volume 47, 2021, 101246, ISSN 1474-0346, https://doi.org/10.1016/j.aei.2021.101246. (https://www.sciencedirect.com/science/article/pii/S147403462100001X)

[19] Neha Soni, Enakshi Khular Sharma, Narotam Singh, Amita Kapoor, Artificial Intelligence in Business: From Research and Innovation to Market Deployment, Procedia Computer Science, Volume 167,2020, Pages 2200-2210, ISSN 1877-0509,https://doi.org/10.1016/j.procs.2020.03.272.

[20] Przemyslaw Pospieszny, Beata Czarnacka-Chrobot, Andrzej Kobylinski, An effective approach for software project effort and duration estimation with machine learning algorithms, Journal of Systems and Software, Volume 137, 2018, Pages 184-196, ISSN 0164-1212, https://doi.org/10.1016/j.jss.2017.11.066.

[21] Kolluri, Johnson and Razia, Shaik and Nayak, Soumya Ranjan, Text Classification Using Machine Learning and Deep Learning Models (June 4, 2020). International Conference on Artificial Intelligence in Manufacturing & Renewable Energy (ICAIMRE) 2019, Available at SSRN: https://ssrn.com/abstract=3618895 or http://dx.doi.org/10.2139/ssrn.3618895

[22] Shervin Minaee, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu, and Jianfeng Gao. 2021. Deep Learning–based Text Classification: A Comprehensive Review. ACM Comput. Surv. 54, 3, Article 62 (June 2021), 40 pages. DOI:https://doi.org/10.1145/3439726

[23] P. Semberecki and H. Maciejewski, ”Deep learning methods for subject text classification of articles,” 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), 2017, pp. 357-360, doi: 10.15439/2017F414.

[24] Jochen Hartmann, Juliana Huppertz, Christina Schamp, Mark Heitmann, Comparing automated text classification methods, International Journal of Research in Marketing, Volume 36, Issue 1, 2019, Pages 20-38, ISSN 0167-8116, https://doi.org/10.1016/j.ijresmar.2018.09.009.

[25] Kowsari, K.; Jafari Meimandi, K.; Heidarysafa, M.; Mendu, S.; Barnes, L.; Brown, D. Text Classification Algorithms: A Survey. Information 2019, 10, 150. https://doi.org/10.3390/info10040150

[26] Bogdan Walek, Vladimir Fojtik, A hybrid recommender system for recommending relevant movies using an expert system, Expert Systems with Applications, Volume 158, 2020, 113452, ISSN 0957-4174, (https://www.sciencedirect.com/science/article/pii/S0957417420302761)

[27] N. Nikzad–Khasmakhi, M.A. Balafar, M. Reza Feizi–Derakhshi, The state-of-the-art in expert recommendation systems, Engineering Applications of Artificial Intelligence, Volume 82, 2019, Pages 126147, ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2019.03.020. (https://www.sciencedirect.com/science/article/pii/S0952197619300703)

[28] M. J. Blosseville,G. Hebrail,M.´ G. Monteil,N. Penot,´ Automatic document classification: natural language processing, statistical analysis, and expert system techniques used together, SIGIR ’92: Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval June 1992 Pages 51–58https://doi.org/10.1145/133160.133175

[29] Goel, A., Bhujade, R.K. “A functional review, analysis and comparison of position permutation based image encryption techniques” (2020) International Journal of Emerging Technology and Advanced Engineering, 10 (7), pp. 97-99.

[30] Paech, B., Heinrich, R., Zorn-Pauli, G., Jung, A., Tadjiky, S. (2012). Answering a Request for Proposal – Challenges and Proposed Solutions. In: Regnell, B., Damian, D. (eds) Requirements Engineering: Foundation for Software Quality. REFSQ 2012. Lecture Notes in Computer Science, vol 7195. Springer, Berlin, Heidelberg.
[31] Beason, Sterling; Hinton, William; Salamah, Yousri A.; and Salsman, Jordan (2021) "Automated Analysis of RFPs using Natural Language Processing (NLP) for the Technology Domain," SMU Data Science Review:
Vol. 5: No. 1, Article 1.
[32] Dubey, G.P., Bhujade, R.K. "Improving the performance of intrusion detection system using machine learning based approaches ", International Journal of Emerging Trends in Engineering Research, 2020, 8(9), pp. 4947–4951
[33] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. arXiv. https://doi.org/10.48550/arXiv.1706.03762
[34] Fahad ul Hassan and Tuyen Le, Ph.D., A.M.ASCE, Automated Requirements Identification from Construction Contract Documents Using Natural Language Processing, Journal of Legal Affairs and Dispute
Resolution in Engineering and Construction Vol. 12, Issue 2 (May 2020),https://doi.org/10.1061/(ASCE)LA.1943-4170.0000379
[35] G. Fantoni, E. Coli, F. Chiarello, R. Apreda, F. Dell'Orletta, G. Pratelli, Text mining tool for translating terms of the contract into technical specifications: Development and application in the railway sector, Computers in Industry, Volume 124,2021,103357, ISSN 01663615,https://doi.org/10.1016/j.compind.2020.103357.
[36] Asha Rajbhoj, Padmalata Nistala, Vinay Kulkarni, Pulkit Batra, AIenabled Project Initiation: An approach based on RFP Response Document, ISEC 2022: 15th Innovations in Software Engineering ConferenceFebruary 2022 Article No.: 22Pages 1–5https://doi.org/10.1145/3511430.3511450
[37] Fahad ul Hassan; Tuyen Le; and Duc-Hoc Tran, Multi-Class Categorization of Design-Build Contract Requirements Using Text Mining and Natural Language Processing Techniques, Construction Research Congress 2020
[38] H. R. Motahari-Nezhad, J. M. Cappi, T. Nakamura, and M. Qiao, "RFPCog: Linguistic-Based Identification and Mapping of Service Requirements in Request for Proposals (RFPs) to IT Service Solutions," 2016 49th Hawaii International Conference on System Sciences (HICSS), 2016, pp. 1691-1700, DOI: 10.1109/HICSS.2016.213.
[39] Fahad ul Hassan, Tuyen Le, Computer-assisted separation of design-build contract requirements to support subcontract drafting, Automation in Construction, Volume 122,2021,103479, ISSN 0926-5805,https://doi.org/10.1016/j.autcon.2020.103479.
(https://www.sciencedirect.com/science/article/pii/S0926580520310591)
[40] A. Sainani, P. R. Anish, V. Joshi and S. Ghaisas, "Extracting and
Classifying Requirements from Software Engineering Contracts," 2020 IEEE 28th International Requirements Engineering Conference (RE), 2020, pp. 147-157, doi: 10.1109/RE48521.2020.00026.
[41] Subhadip Maji, Anudeep Srivatsav Appe, Raghav Bali, Veera Raghavendra Chikka, Arijit Ghosh Chowdhury, Vamsi M Bhandaru,An Interpretable Deep Learning System for Automatically Scoring Request for Proposals,2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)
[42] Costal, D., Franch, X., Lopez, L., Palomares, C., Quer, C. (2019). On the´ Use of Requirement Patterns to Analyse Request for Proposal Documents. In: Laender, A., Pernici, B., Lim, EP., de Oliveira, J. (eds) Conceptual Modeling. ER 2019. Lecture Notes in Computer Science (), vol 11788. Springer, Cham.

[43] Asha Rajbhoj, Padmalata Nistala, Vinay Kulkarni, Gomathi Ganesan, A RFP System for Generating Response to a Request for Proposal. ISEC'19: Proceedings of the 12th Innovations on Software Engineering Conference
(formerly known as India Software Engineering Conference) February 2019 Article No.: 9Pages 1–9https://doi.org/10.1145/3299771.3299779
[44] Ningwei Liu, Charles Feng, Shunning Wu, Aaron Chan, Joel Fulton, Automate RFP Response Generation Process Using FastText Word Embeddings and Soft Cosine Measure, AICS 2019: Proceedings of the 2019 International Conference on Artificial Intelligence and Computer ScienceJuly 2019 Pages 12–17https://doi.org/10.1145/3349341.3349362
[45] Nazir, F., Butt, W.H., Anwar, M.W., Khan Khattak, M.A.. The Applications of Natural Language Processing (NLP) for Software Requirement Engineering - A Systematic Literature Review. Kim, K., Joukov, N. (eds) Information Science and Applications. ICISA 2017. Lecture Notes in Electrical Engineering, vol 424. Springer, Singapore.
[46] Mohsin Irshad, Kai Petersen, Simon Poulding,A systematic literature review of software requirements reuse approaches, Information and Software Technology, Volume 93,2018, Pages 223-245,ISSN 0950-5849,https://doi.org/10.1016/j.infsof.2017.09.009.(https://www.sciencedirect.com/science/article/pii/S0950584916303615)
[47] Marcin Michał Mironczuk,´ Jarosław Protasiewicz, A recent overview of the state-of-the-art elements of text classification, Expert Systems with Applications, Volume 106, 2018, Pages 3654, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2018.03.058. https://www.sciencedirect.com/science/article/pii/S095741741830215X
[48] J. M. Perez-Verdejo, A. J. S´ anchez-Garc´ ´ıa and J. O. Ocharan-Hern´ andez,´ "A Systematic Literature Review on Machine Learning for Automated Requirements Classification," 2020 8th International Conference in Software Engineering Research and Innovation (CONISOFT), 2020, pp. 21-28, doi: 10.1109/CONISOFT50191.2020.00014.
[49] Goel, A., Bhujade, R., "A prototype model for image encryption using zigzig blocks in inter-pixel displacement of RGB value" International Journal of Advanced Trends in Computer Science and Engineering, 2020, 9(5), pp. 6905–6912
[50] Olsson, T., Sentilles, S. and Papatheocharous, E. A systematic literature review of empirical research on quality requirements. Requirements Eng 27, 249–271 (2022). https://doi.org/10.1007/s00766-022-00373-9
[51] Rajbabu K., Harshavardhan Srinivas, Sudha S., Industrial information extraction through multi-phase classification using ontology for unstructured documents, Computers in Industry, Volume 100, 2018, Pages 137147, ISSN 0166-3615, https://doi.org/10.1016/j.compind.2018.04.007.
[52] Y. Saito, A. Monden, and K. Matsumoto, "Evaluation of Non-Functional Requirements in a Request for Proposal (RFP)," 2012 Joint Conference of the 22nd International Workshop on Software Measurement and the 2012 Seventh International Conference on Software Process and Product Measurement, 2012, pp. 106-111, DOI: 10.1109/IWSM-MENSURA.2012.23.
[53] P. Shah, S. Joshi, and A. K. Pandey, "Legal Clause Extraction From Contract Using Machine Learning with Heuristics Improvement," 2018 4th International Conference on Computing Communication and Automation (ICCCA), 2018, pp. 1-3, DOI: 10.1109/CCAA.2018.8777602.
[54] S. P. Nayak and S. Pasumarthi, "Automatic Detection and Analysis of DPP Entities in Legal Contract Documents," 2019 First International Conference on Digital Data Processing (DDP), 2019, pp. 70-75, DOI: 10.1109/DDP.2019.00023.
[55] Katiyar, N., Bhujade, R.," Authentication & spam analysis architecture over C-hlms " International Journal of Scientific and Technology Research, 2019, 8(9), pp. 612–615

[56] Chalkida, I., Androutsopoulos, I., and Michos, A. (2017). Extracting contract elements. Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law.
[57] Costal D., Franch X., Lopez L., Palomares C., Quer C. (2019) On the´ Use of Requirement Patterns to Analyse Request for Proposal Documents. In: Laender A., Pernici B., Lim EP., de Oliveira J. (eds) Conceptual Modeling. ER 2019. Lecture Notes in Computer Science, vol 11788. Springer, Cham.
[58] G. Fantoni, E. Coli, F. Chiarello, R. Apreda, F. Dell'Orletta, G. Pratelli, Text mining tool for translating terms of the contract into technical specifications: Development and application in the railway sector, Computers in Industry, Volume 124, 2021, 103357, ISSN 0166-3615, https://doi.org/10.1016/j.compind.2020.103357.
[59] Y. Kim, J. Lee, E. -B. Lee and J. -H. Lee, "Application of Natural Language Processing (NLP) and Text-Mining of Big-Data to Engineering-Procurement-Construction (EPC) Bid and Contract Documents," 2020 6th Conference on Data Science and Machine Learning Applications (CDMA), 2020, pp. 123-128, DOI: 10.1109/CDMA47397.2020.00027.
[60] Y. Saito, A. Monden, and K. Matsumoto, "Evaluation of Non-Functional Requirements in a Request for Proposal (RFP)," 2012 Joint Conference of the 22nd International Workshop on Software Measurement and the 2012 Seventh International Conference on Software Process and Product Measurement, 2012, pp. 106-111, DOI: 10.1109/IWSM-MENSURA.2012.23
[61] Zhongxuan Wang, Letian Wu, and Xiangcheng Ji. 2021. An Interpretable Deep Learning System for Automatic Intracranial Hemorrhage Diagnosis with CT Image. Proceedings of the 2021 International Conference on Bioinformatics and Intelligent Computing. Association for Computing Machinery, New York, NY, USA, 338–357.
DOI:https://doi.org/10.1145/3448748.3448803
[62] P. Semberecki and H. Maciejewski, "Deep learning methods for subject text classification of articles," 2017 Federated Conference on Computer Science and Information Systems (FedCSIS), 2017, pp. 357-360, DOI: 10.15439/2017F414.
[63] https://huggingface.co/docs/transformers/main/en/index
[64] A Systematic Literature Review on Using Machine Learning Algorithms for Software Requirements Identification on Stack Overflow, Arshad Ahmad,1,2Chong Feng,1Muzammil Khan,3Asif Khan,1Ayaz Ullah,2Shah Nazir,2and Adnan Tahir, , https://www.hindawi.com/journals/scn/2020/8830683/
[65] Mehta, R., Bhujade, R.K.," Advance cbir by using cts features with relevance feedback"International Journal of Scientific and Technology Research, 2019, 8(9), pp. 1057–1060
[66] Kubeka, Sibusiso and Ade-ibijola, Abejide, "Automatic Comprehension and Summarisation of Legal Contracts," Advances in Science, Technology and Engineering Systems Journal