



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 15, Issue 3, March 2026

AutoPRD: Automated Product Requirement Document Generation Using NLP and Machine Learning

Mrs.B.Uma Rani¹, T Mohan Krishna², T Sai Tejasvi³, C Surendhar Reddy⁴

Assistant Professor, Department of CSE, ACE Engineering College, Hyderabad, Telangana, India¹

IV B.Tech Students, Department of CSE, ACE Engineering College, Hyderabad, Telangana, India^{2,3,4}

ABSTRACT: In contemporary software development, the preparation of Product Requirement Documents (PRDs) is a critical activity that assists teams in understanding product features, system behavior, and development requirements. Nevertheless, manual preparation of PRDs from unstructured discussions like team conversations, meeting discussions, and requirement documents is a time-consuming and error-prone process. To overcome this problem, this paper presents the AutoPRD system, an automated tool that automatically generates structured Product Requirement Documents from unstructured text inputs through Machine Learning and Natural Language Processing (NLP) techniques. The system obtains inputs from different sources like text input, documents, and communication tools, analyzes the input data using NLP techniques, and extracts critical requirement details. A trained transformer model is employed to extract significant product features, functional requirements, and system information. Using this data, the system automatically generates structured PRD templates with feature descriptions, user stories, and technical details. The proposed method minimizes the effort required for manual documentation, maximizes productivity, and maintains consistency in product requirement documentation. The system is also capable of handling multiple document templates, version management, and export options, which makes it an ideal tool for product managers, developers, and software teams.

KEYWORDS: Product Requirement Document, Natural Language Processing, Machine Learning, DistilBERT, Requirement Extraction, Document Automation.

I. INTRODUCTION

Product Requirement Documents (PRDs) are critical in today's software development processes as they outline product features, development objectives, and system requirements. They also ensure that product managers, designers, and developers remain on the same page by outlining product functionality and expected results. Nevertheless, the process of manually creating PRDs is time-consuming and laborious, especially when requirements are sourced from different meetings, chat sessions, and technical discussions.

As the use of digital collaboration tools continues to rise, there is an abundance of requirement-related information in unstructured formats. The process of extracting meaningful information from such conversations and structuring it into PRDs is difficult, resulting in delays, inaccuracies, and lost requirements.

To overcome these challenges, this paper presents AutoPRD, an automated system that automatically generates Product Requirement Documents from unstructured inputs using Machine Learning and Natural Language Processing (NLP) techniques.

II. LITERATURE SURVEY

Recent studies have investigated the application of Artificial Intelligence, Natural Language Processing (NLP), and Large Language Models (LLMs) in Requirement Engineering to automate tasks such as requirement extraction, analysis, and documentation.

The research paper “LLMs for Requirements Engineering: A Systematic Literature Review” examines the application of LLMs in requirement analysis, documentation, and validation. However, it mostly reviews existing research and lacks implementation.[1]

The research paper “From Requirements to Test Cases with NLP” applies NLP techniques such as Named Entity Recognition (NER) and rule-based extraction to automatically create test cases from requirement descriptions. Although successful, it mostly applies to requirement descriptions in specific formats.[2]

The ReXCL tool is specifically designed to apply heuristic techniques and machine learning algorithms to extract and classify information from requirement documents. However, it still lacks general validation.[3]

In summary, existing research mostly targets requirement extraction and task-specific automation. There is limited research on the automatic generation of structured Product Requirement Documents (PRDs) from unstructured conversations, which is addressed by the proposed AutoPRD system utilizing NLP techniques.

III. METHODOLOGY

The proposed AutoPRD system’s methodology is centered on the automatic creation of Product Requirement Documents from unstructured text inputs through the application of Natural Language Processing and Machine Learning algorithms.

The proposed system starts with the acquisition of user inputs, which could be in the form of text descriptions, meeting transcripts, communication messages, or uploaded documents. The inputs are initially preprocessed using text extraction algorithms to extract the raw text. The text is then subjected to further preprocessing steps like tokenization, stop-word removal, and text normalization to clean the text.

After the data is preprocessed, Natural Language Processing algorithms are applied to analyze the text and extract key product requirements, feature descriptions, and functional specifications. A transformer model is applied to classify and extract relevant information about product features and development tasks. The extracted information is then structured according to predefined PRD templates.

The system automatically creates different parts of the Product Requirement Document, including product overview, functional requirements, user stories, and technical specifications. Finally, the automatically created PRD is displayed to the user through a web interface, where it can be viewed, modified, and saved in different formats.

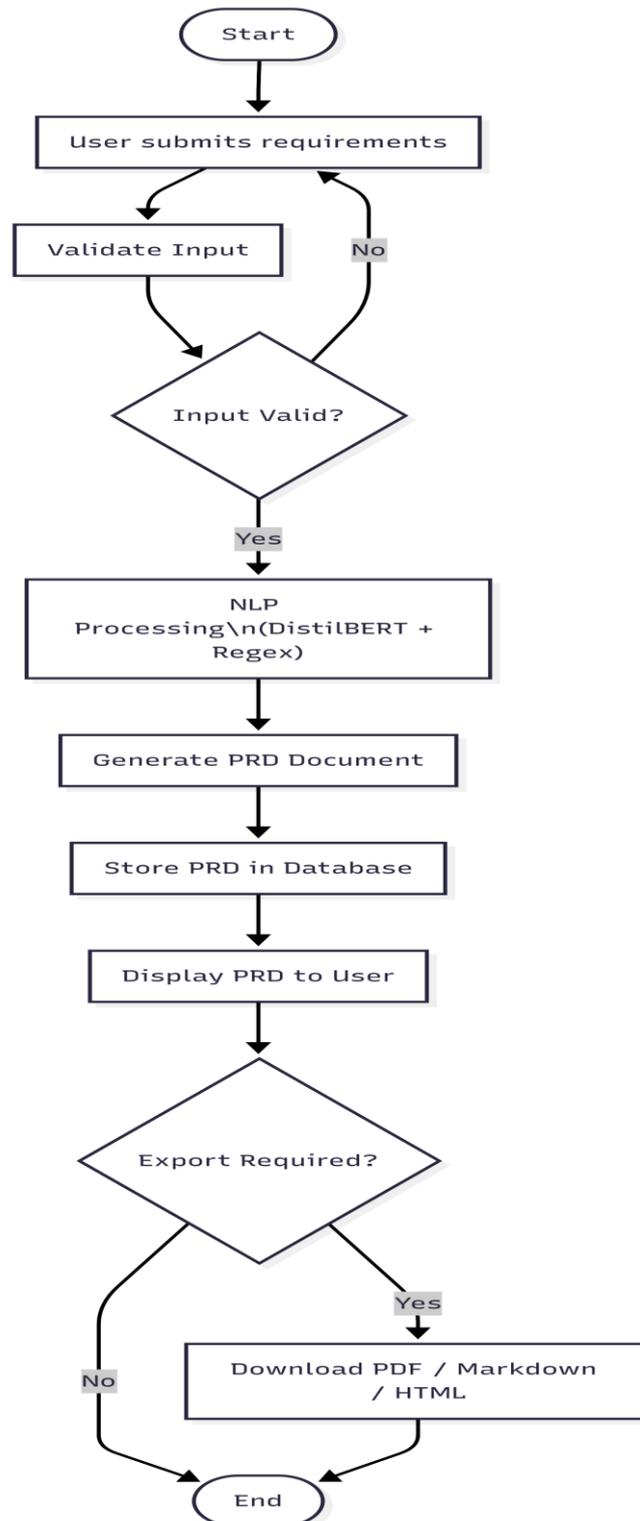


Fig 3.1 Automated Product Requirement Document Generation Methodology

IV. SYSTEM ARCHITECTURE

The architecture of the AutoPRD system is designed as a modular pipeline that integrates the data collection, NLP processing, machine learning analysis, and document generation functions.

The AutoPRD system architecture comprises the frontend interface, backend module, NLP model pipeline, and database storage module. The frontend module provides the user with the opportunity to input the requirements, view the generated PRD documents, and manage documents. The backend module processes the user input, performs the NLP model analysis, and stores the documents using the database module.

The NLP model pipeline uses the transformer model to analyze the user input data and obtain requirement information. The requirement information is then passed to the document generation module, where the information is formatted into the PRD document template.

The database module stores the user input, generated documents, and document versions to support user collaboration and document tracking. The generated PRD document can be exported into various formats such as PDF, Markdown, and HTML.

The AutoPRD system architecture efficiently processes user input data and provides document generation capabilities for software development teams.

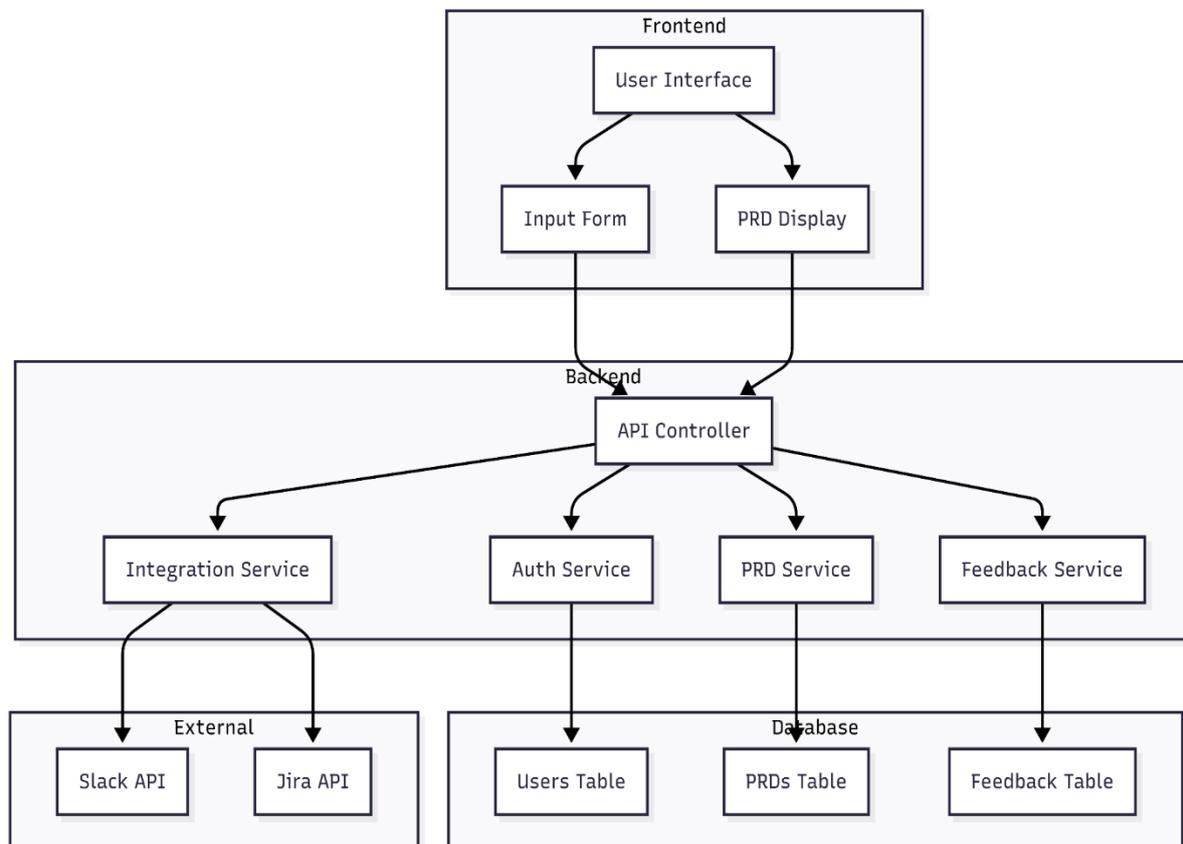


Fig 4.1 System Architecture

V. OUTPUT SCREENS

The AutoPRD system has an interface that is user-friendly, enabling users to generate product requirement documents based on the information given. The main interface enables users to submit their product requirements using text input fields or document upload.

Once the user submits the information, the system uses NLP technology to process the text, enabling it to identify the significant information related to the product requirements. The information is then used to generate PRD document sections such as product overview, features, and user stories.

The generated document is then made available to the user through the application interface, enabling the user to edit it if need be. The user is able to export the generated PRD document in various formats, such as PDF, HTML, and Markdown.

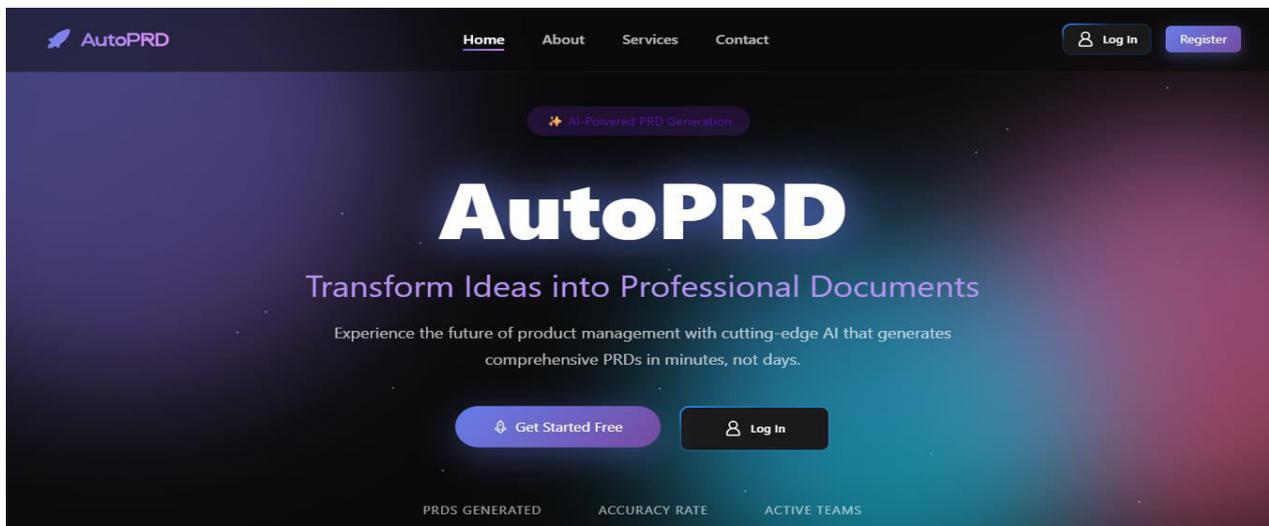


Fig 5.1 AutoPRD Home Interface

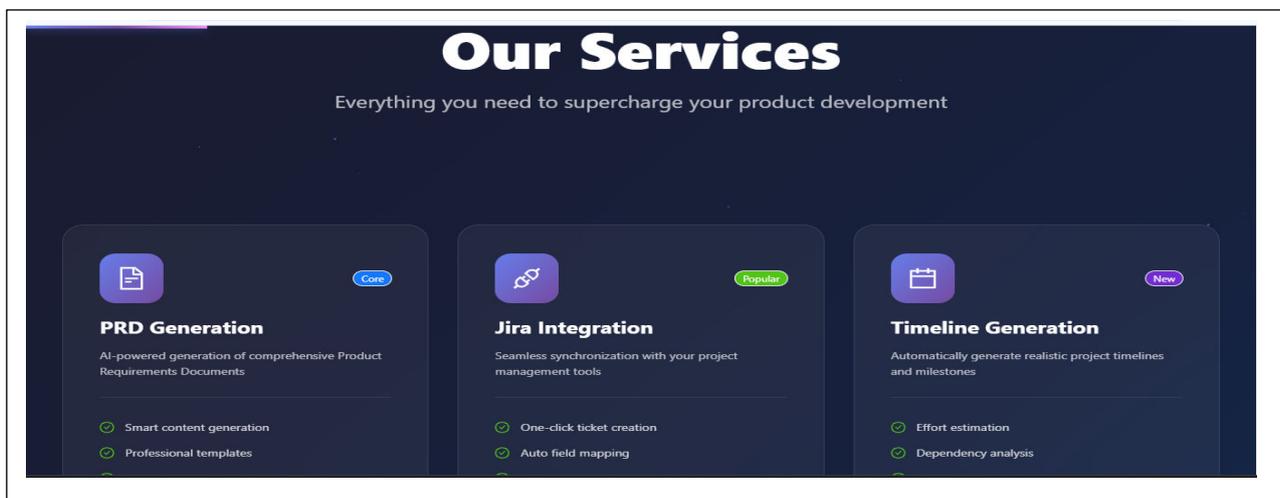


Fig 5.2 Services Page

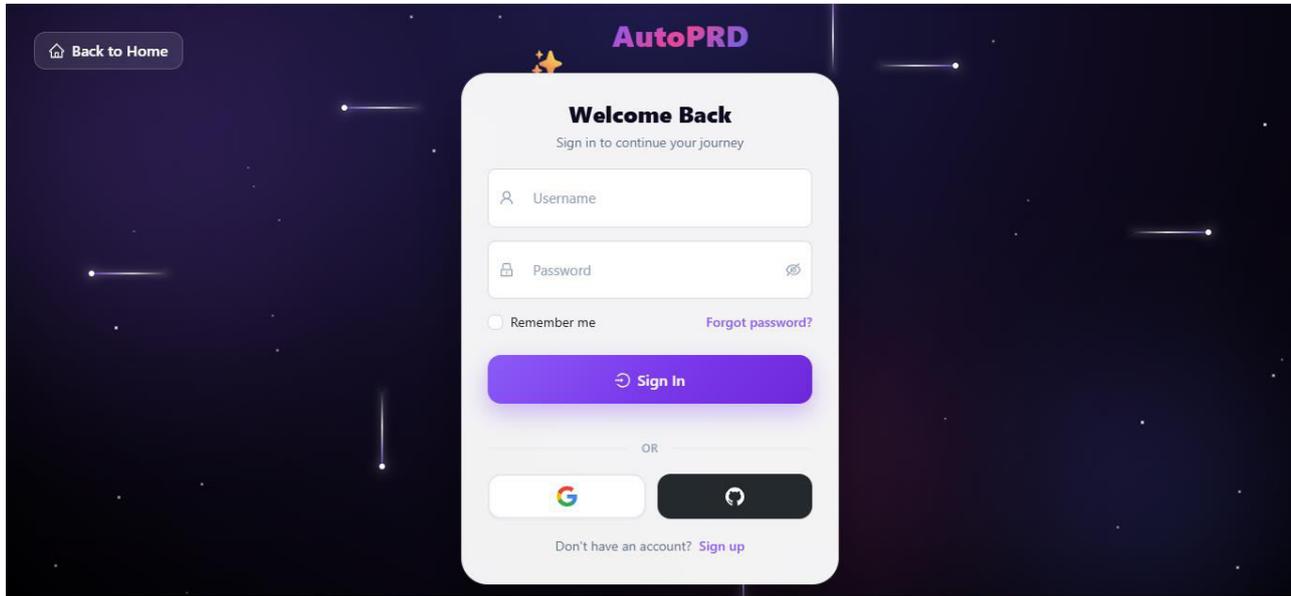


Fig 5.3 Login Interface

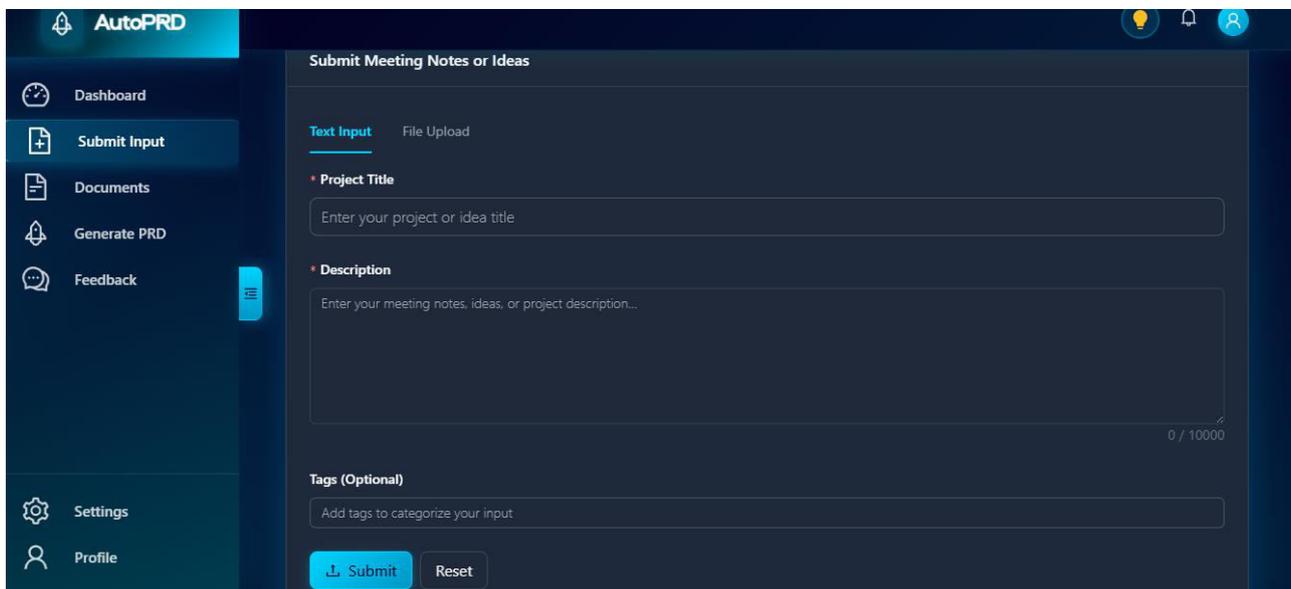


Fig 5.4 PRD Generation Dashboard

VI. CONCLUSION

In this project, an automated system named AutoPRD was created to produce Product Requirement Documents using Machine Learning and Natural Language Processing approaches. The process of manually writing PRDs is a time-consuming process that involves collecting data from different sources and presenting it in a structured form. The proposed system makes this process easier by automatically analyzing unstructured text inputs and extracting relevant information about key requirements. By leveraging NLP approaches and transformer models, the system is capable of extracting product features, functional requirements, and development tasks. The PRDs produced by the system enable development teams to easily comprehend product requirements and facilitate collaboration between different

stakeholders. In general, the system illustrates how artificial intelligence can be applied to automate documentation work and boost productivity in software development environments.

VII. FUTURE SCOPE

Although the AutoPRD system has successfully automated the generation of PRD, some changes can be made to improve the system in the future. One possible change that can be made to the system is the incorporation of sophisticated language models to enhance the efficiency of the system. Another possible change that can be made to the system is the extension of the system to support multiple languages, enabling different teams around the world to make efficient use of the system. A possible change that can be made to the system is the incorporation of the system with other project management systems, like Jira or Slack, to enable the generation of requirements based on conversations. Another possible change that can be made to the system is the incorporation of intelligent recommendation systems that can recommend requirements that are not specified or can point out possible conflicts that might exist in the specifications of the product.

REFERENCES

- [1] Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv.
Link: <https://arxiv.org/abs/1811.00196>
- [2] Sanh, V., Debut, L., Dernoncourt, F., and Wolf, T. (2020). DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. arXiv.
Link: <https://arxiv.org/abs/1910.01108>
- [3] Hinton, G., Vanhoucke, V., and Dean, J. (2015). Distilling the knowledge in a neural network. arXiv.
Link: <https://arxiv.org/abs/1503.02531>
- [4] Kurtanović, Z., and Maalej, W. (2017). Automatically Classifying Functional and Non-Functional Requirements Using Supervised Machine Learning. IEEE RE Conference.
Link: <https://ieeexplore.ieee.org/document/8049171>
- [5] Zhao, L., Alhoshan, W., Ferrari, A., Letsholo, K., Ajagbe, M., Chioasca, E., and Batista-Navarro, R. (2020). Natural Language Processing for Requirements Engineering: A Systematic Mapping Study. arXiv.
Link: <https://arxiv.org/abs/2004.01099>
- [6] Slankas, J., and Williams, L. (2013). Automated Extraction of Non-Functional Requirements from Documentation. NaturaLiSE Workshop.
Link: <https://ieeexplore.ieee.org/document/6611715>
- [7] Raffel, C., Shazeer, N., Roberts, A., et al. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer (T5). JMLR.
Link: <https://arxiv.org/abs/1910.10683>
- [8] Brown, T., Mann, B., Ryder, N., et al. (2020). Language Models are Few-Shot Learners. NeurIPS.
Link: <https://arxiv.org/abs/2005.14165>
- [9] Ferrari, A., Spagnolo, G., and Gnesi, S. (2019). PURE: A Dataset of Public Requirements Documents. arXiv.
Link: <https://arxiv.org/abs/1904.12724>
- [10] Abualhaija, S., Arora, C., and Ernst, M. D. (2019). Automated Extraction of Functional Requirements from User Stories. IEEE RE Conference.
Link: <https://ieeexplore.ieee.org/document/8920498>



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details