iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VII **Month of publication:** July 2025

DOI: https://doi.org/10.22214/ijraset.2025.73399

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VII July 2025- Available at www.ijraset.com

Indian Legal Text Summarization Using InCaseLaw BERT

Aswin Asokan[1], Bineesha K P[2], Shyam Prakash M[3], Dr. Jisha P Abraham[4], Prof. Richu Shibu[5]
Department of Computer Science and Engineering, Mar Athanasius College of Engineering (Autonomous), Kothamangalam

Abstract: ***Summarizing Indian legal documents such as court judgments and orders have significant difficulties due to its complexity and length. Recent advances in Natural Language Processing (NLP) pave the way to overcome this challenge. This paper presents the design, implementation, and performance evaluation of the summarization of Indian legal texts using domain-specific transformer models. We begin the work with an introduction to domain-specific transformer models for the summarization of legal texts. Through this work, we have used pre-trained transformer models fine-tuned on various In- dian court judgments to generate concise summaries categorized into facts, arguments, judgments, analysis, and statutes, ensuring read- ability. Key components of the work include fine-tuned transformer models for sentence selection, categorization, and paraphrasing, as well as Google's Gemini model for assisting users with their inquiries. The work aims to assist users in accessing and reduce the time taken for research of these complex texts.***
Keywords: India Legal Text Summarization, Natural Language Processing, Transformers

I. INTRODUCTION

The Indian legal documents are often complex in nature, and extracting useful information is a challenging task for users. This causes difficulties for indi- viduals researching such documents, especially common people. Traditional legal documents are time-consuming and need specialist expertise, inviting ex- tra costs. While general text summarization techniques exist they are bound to make mistakes and may omit important parts from the original text. These challenges point out the need for efficient tools that can enhance readability and reduce time spent on each document. This work takes these challenges into account and tries to make the task easier by implementing an Indian legal text summarizer and support system using domain specific transformer models. The work uses text extraction capa- bilities of InCaseLaw BERT transformer model to detect important sentences from the original text, categorize them into various sections. These sections are paraphrased to make it more readable and shown as a concise summary. To improve the usability, the system integrates Google Gemini, a large language model (LLM) based conversational model set to answer queries regarding Indian legal content to help users. The performance of the system was evaluated using F1 and ROUGE scores, which confirmed its ability to efficiently provide users with proper summaries while preserving key sentences from the original text.

II. OVERVIEW

Artificial intelligence (AI) and Natural Language Processing (NLP) methods are widely used in the legal domain nowadays. But most of the current legal text summarization systems generates a single paragraph by including some of the important sentences from the original text. The aim of the work is to create section wise concise summary of a document given to the system. This section covers some of the technologies we have referred for building the system.

A. Domain Specific BERT Models

Bidirectional Encoder Representations from Transformers (BERT) is a power- ful framework built by Google for natural language processing tasks. The main ability of the model is that it can understand context of a sentence by consid- ering both the preceding and following words. The bi-directional approach of BERT models makes it widely used in NLP tasks.

There are existing pre-trained domain specific BERT models such as In- CaseLaw BERT and InLegal BERT. These are trained on a corpus of Indian legal text. This enables the models to detect key terms and sentences from a given Indian legal text. As shown in [1] pre-trained BERT models such as Legal BERT and CaseLaw BERT can be used for Sentence Boundary Detection. The work mentioned in [3] describes how they have developed the InCaseLaw and InLegal BERT models using a corpus of Indian legal text. We chose the above models as they are trained on Indian corpus and can be used efficiently for sentence extraction tasks. Also [2] mentions different methodologies that can be used for legal text summarization.

B. Paraphrasing for Readability

When the model extracts text from original document the sentences may not have any grammatical consistency. This is a common issue with extractive sum- mary. Paraphrasing is used to improve readability of sentences as mentioned in Comparative analysis of paraphrasing performance of ChatGPT, GPT-3, and T5 language models using a new ChatGPT generated dataset: ParaGPT [5]. And there are models by Vorobev and Kuznetsov, who trained a paraphrasing model using ChatGPT-generated text pairs, showing that LLM-assisted para- phrasing can significantly improve readability while preserving meaning [6].

C. Usage of LLM for legal assistance

The journal by Kevin, Craig & Johnson, Niyi [4] mentions the usage of Large Language Models (LLM) chatbots for free legal assistance. The work describes the benefits, challenges and existing implementations of chatbots in the legal domain. In our work we've used Gemini, an LLM chatbot developed by Google and set its parameters to answer within its knowledge on Indian legal domain.

D. Conclusion

Existing literature suggests the potential to enhance the efficiency in generating concise legal summaries and access to legal assistance. However, challenges related to the models used in the work need to be addressed to ensure reliable summaries. The challenges we have encountered will be discussed in the results section.

III. OBJECTIVES

The primary objective of this work is to develop an AI-assisted system that simplifies the interpretation of Indian legal documents through automated summarization and interactive support. The system is designed to generate efficient, structured and concise summaries of uploaded documents making research on large and complex legal documents easier. The key objectives of the work are outlined below:

A. Improving Accessibility to Legal Information

Legal documents are often complex and time-consuming to read. The system aims to close the gap by generating structured summaries from large legal documents, making it easier for users to understand judgments, case laws, and statutes. The tool supports broader access to justice by making legal information more digestible for a wider audience.

B. Enabling Interactive Legal Understanding

Beyond static summaries users may need further assistance for clear understanding of the uploaded document. The system integrates an interactive layer where users can ask questions regarding specific sections of the document. This is particularly beneficial for users without legal training.

C. Delivering a User-Centric Interface for Legal Navigation

To ensure ease of usage the system features a simple user interface that supports uploading documents, presenting structured summaries, and conversational interactions. The user interface is designed with accessibility in mind, enabling users from diverse backgrounds including legal practitioners, and the general public to interact with the system effortlessly.

D. Supporting Legal Education and Research Efficiency

During research purposes, legal scholars, students, and professionals may face the challenge of reviewing large documents for relevant information. This system supports them by extracting key information and presenting it as a structured summary from large documents. This way they can reduce the intellectual load when reviewing multiple documents.

E. Conclusion

By addressing these objectives, the system seeks to enhance the accessibility, clarity, and usability of Indian legal content through a structured and interactive AI-based system. The system not only helps legal practitioners in need of reduced workload in research but also individuals in need of legal assistance.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VII July 2025- Available at www.ijraset.com

IV. METHODOLOGY

This section explains the design, data preparation, development, and implementation of Indian legal text summarization system.

A. System Architecture

The system architecture consists of 5 major components:

1) User Interface (UI): A web based interface where users can upload legal documents and receive summaries. It also serves as the chatbot interface
2) PDF Parsing: Uses PyPDF2 library to extract text from Indian Legal documents in PDF format
3) InCaseLaw BERT model: Generates section-wise textual data by extracting key sentences from the document and categorizing them to each section, omitting unwanted sentences
4) ChatGPT paraphraser on T5 base: A paraphraser model to improve the readability of the extracted text
5) Google Gemini: Google's large language model (LLM) set to answer queries based on Indian Legal System

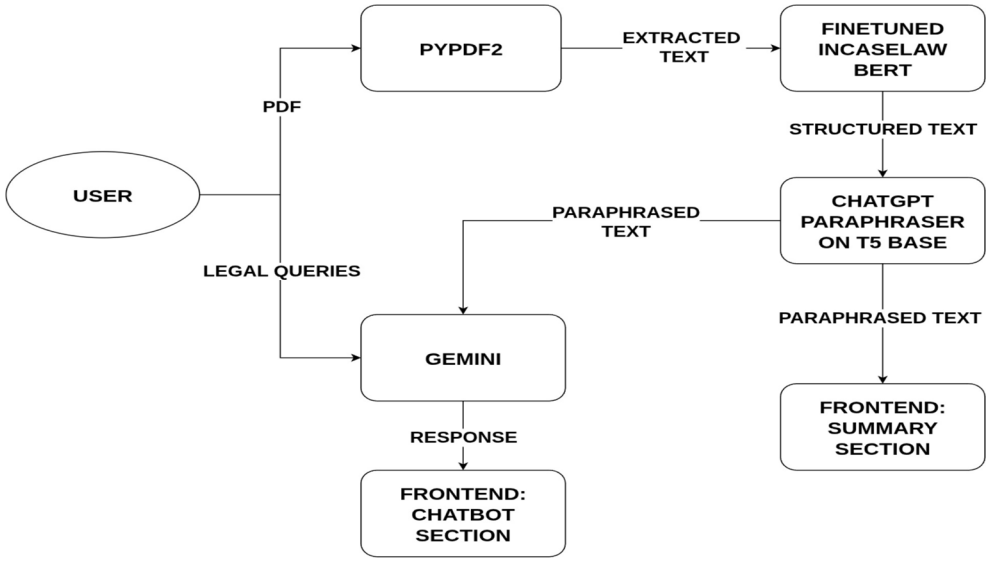


Figure 1: Overview of Indian Legal Text Summarization System

B. Data Preparation

To generate accurate summaries the model had to be trained using a proper dataset. For this, the dataset prepared by Abhay Shukla, Paheli Bhattacharya, Soham Poddar, Rajdeep Mukherjee, Kripabandhu Ghosh, Pawan Goyal, & Saptarshi Ghosh [7] was used. The dataset contains Indian Supreme Court case documents & their 'extractive' summaries, written by two law experts. The dataset was converted from TXT files to CSV and sentences were given section-wise labels. This was used for the training and performance evaluation of the model.

C. Summarization Pipeline

In this work we have adopted a hybrid summarization strategy. First, the fine-tuned InCaseLaw BERT model will extract important sentences from uploaded documents and assign them labels based on the sections they belong to. The model was trained to predict the sections in each key sentence. Then this text is passed to the ChatGPT paraphraser on the T5 base. This model paraphrases the given sentences to improve readability and reduce redundancy. By implementing a summarization model in this way instead of a single paragraph we were able to generate structured summaries with section-wise information (facts, arguments, judgments, analysis, and statutes) extracted from the original text.

D. Interactive Support Layer

For enhancing user experience, we have integrated Google Gemini's large language model (LLM) to provide dynamic query-answering capabilities. The Gemini model is set to answer questions based on its knowledge of the Indian Legislature. With this users can ask questions based on the summary generated by the fine-tuned model or questions regarding the Indian legal system.

E. Implementation

The system is implemented using the following technologies:

1) Backend: The backend is developed using Python utilizing Hugging Face transformer models. It also uses PyPDF2 library for PDF parsing.

2) Frontend: React based web interface

3) APIs: Google Gemini API to handle query answering. RESTful API endpoints are defined using FastAPI to connect frontend with backend models.

F. Conclusion

The methodology detailed above provides a structured approach to implementing an Indian legal text summarization system using NLP methods. By integrating a fine-tuned InCaseLaw BERT model and Gemini for query answering the system ensures accurate summaries and assistance to users.

V. RESULTS AND DISCUSSION

This section discuss the evaluation of the system focusing on its performance, accuracy and usability. The system's sentence extraction and classification ability was evaluated using F1 score and the summary was evaluated using ROUGE score.

A. Evaluating using F1 score

The F1 score is used to define a model's ability to predict and classify given data. In our case, the sentences had to be classified into facts, arguments, judgments, analysis, and statutes. The F1 score combines the precision and recall scores of a model, while the accuracy metric computes how many times a model makes a correct prediction across the entire dataset. Through fine-tuning the InCaseLaw BERT and InLegal BERT transformer models we were able to generate the given results.

Table 1: Results of fine-tuned InCaseLaw BERT

Loss	Accuracy	Precision	Recall	F1	Epoch
0.5052	0.8206	0.8197	0.8206	0.8122	1.0
0.5862	0.8661	0.8669	0.8661	0.8620	2.0
0.6192	0.8929	0.8911	0.8929	0.8912	3.0
0.7843	0.8741	0.8752	0.8741	0.8697	4.0
0.8596	0.8755	0.8773	0.8755	0.8709	5.0

Table 2: Results of fine-tuned InLegalBERT

Loss	Accuracy	Precision	Recall	F1	Epoch
0.5132	0.8086	0.8075	0.8086	0.7978	1.0
0.5326	0.8661	0.8658	0.8661	0.8615	2.0
0.6809	0.8755	0.8789	0.8755	0.8705	3.0
0.7457	0.8849	0.8871	0.8849	0.8813	4.0
0.8572	0.8782	0.8834	0.8782	0.8729	5.0

Table 3: Training summary of the models used

Metric	InCaseLaw BERT	InLegalBERT
Total Training Time	4590.15 seconds	4623.0386 seconds
Samples/Second	7.32	7.266
Steps/Second	0.915	0.908
Final Training Loss	0.2268	0.2541
Total Epochs	5	5

As the InCaseLaw BERT model showcased slightly better results than the InLegal BERT model during the fine-tuning process, we utilized that model for the implementation of this work.

B. Evaluating using ROUGE score

The ROUGE score is a set of metrics used to evaluate machine-generated text. The work uses extractive capabilities of the InCaseLaw BERT model to select important sentences and the ChatGPT paraphraser on T5 base is used to improve its readability by paraphrasing the sentences. The summary generated was evaluated using ROUGE metrics.

- ROUGE-1: Unigram (word-level) overlap.
- ROUGE-2: Bigram overlap.
- ROUGE-L: Longest common subsequence (structural overlap).
- ROUGE-LSum: Longest common subsequence (document level).

Table 4: ROUGE Score Evaluation

ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-L
Sum 0.7700	0.6606	0.6136	0.7305

C. Challenges and Limitations

Despite the results, the system still has some limitations:

- Sentence Boundary Detection (SBD) Errors: BERT models often get confused with sentence boundaries. For example, when there are short forms such as "Rs.", "Mr.", "Dr.", and legal citations, the model may detect it as the end of a sentence. This leads to incorrect sentence segmentation, which in turn affects the accuracy of summarization.
- Computational Efficiency: Processing large legal documents with pages exceeding 30 may take some time to get processed.
- Lack of Section-Wise Annotated Datasets: There is a lack of publicly available Indian legal datasets with summaries categorized by sections. Due to this the model may skip key sentences in a section or may even skip an entire section.

D. Future Enhancements

To improve the system some enhancements are planned:

- Automated document classification: Implementing automated classification of legal documents into categories like civil, criminal, corporate, and constitutional law for quick reference allow users to filter judgments based on case types, jurisdictions, and relevant legal statutes.
- Multilingual support: A multi-lingual model can be implemented to support court judgments in different languages throughout India.
- Similarity checking: Ability to upload multiple PDF files to find similarities between cases, to analyze how the judgments differ in each case.
- Community-driven review system: Enabling a community-driven review system where legal experts can verify and refine summaries can increase the credibility of the summaries generated by the system.
- Mobile app development: Developing and publishing a mobile application helps the system to reach more users.

E. Conclusion

In this paper, we have presented how a transformer based model can be used to generate structured and concise summaries from large Indian legal documents. We evaluated the system with both F1 score and ROUGE scores to ensure its ability to provide accurate summaries. While the system displays promising results, further refinements are required to improve its efficiency and effectiveness. Future works will focus on improving the model through expanding datasets, multilingual and multi-document support, as well as mobile application development to improve accessibility.

REFERENCES

[1] Reshma Sheik, Sneha Rao Ganta, S. Jaya Nirmala "Legal sentence boundary detection using hybrid deep learning and statistical models", National Institute of Technology, Trichy, Tiruchirappalli, Tamil Nadu, India, 14 March 2024

[2] Reshma Sheik, S. Jaya Nirmala "Deep Learning Techniques for Legal Text Summarization", 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) | 978-1-6654-0962-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/UPCON52273.2021.9667640

[3] Paul, Shounak & Mandal, Arpan & Goyal, Pawan & Ghosh, Saptarshi. (2022). Pre-training Transformers on Indian Legal Text. 10.48550/arXiv.2209.06049.

[4] Kevin, Craig & Johnson, Niyi. (2025). The Use of Chatbots in Providing Free Legal Guidance: Benefits and Limitations.

[5] Kurt Pehlivanoğlu, M., Gobosho, R. T., Syakura, M. A., Shanmuganathan, V., & de-la-Fuente-Valentín, L. (2024). Comparative analysis of paraphrasing performance of ChatGPT, GPT-3, and T5 language models using a new ChatGPT generated dataset: ParaGPT. Expert Systems, 41(11), e13699. https://doi.org/10.1111/exsy.13699

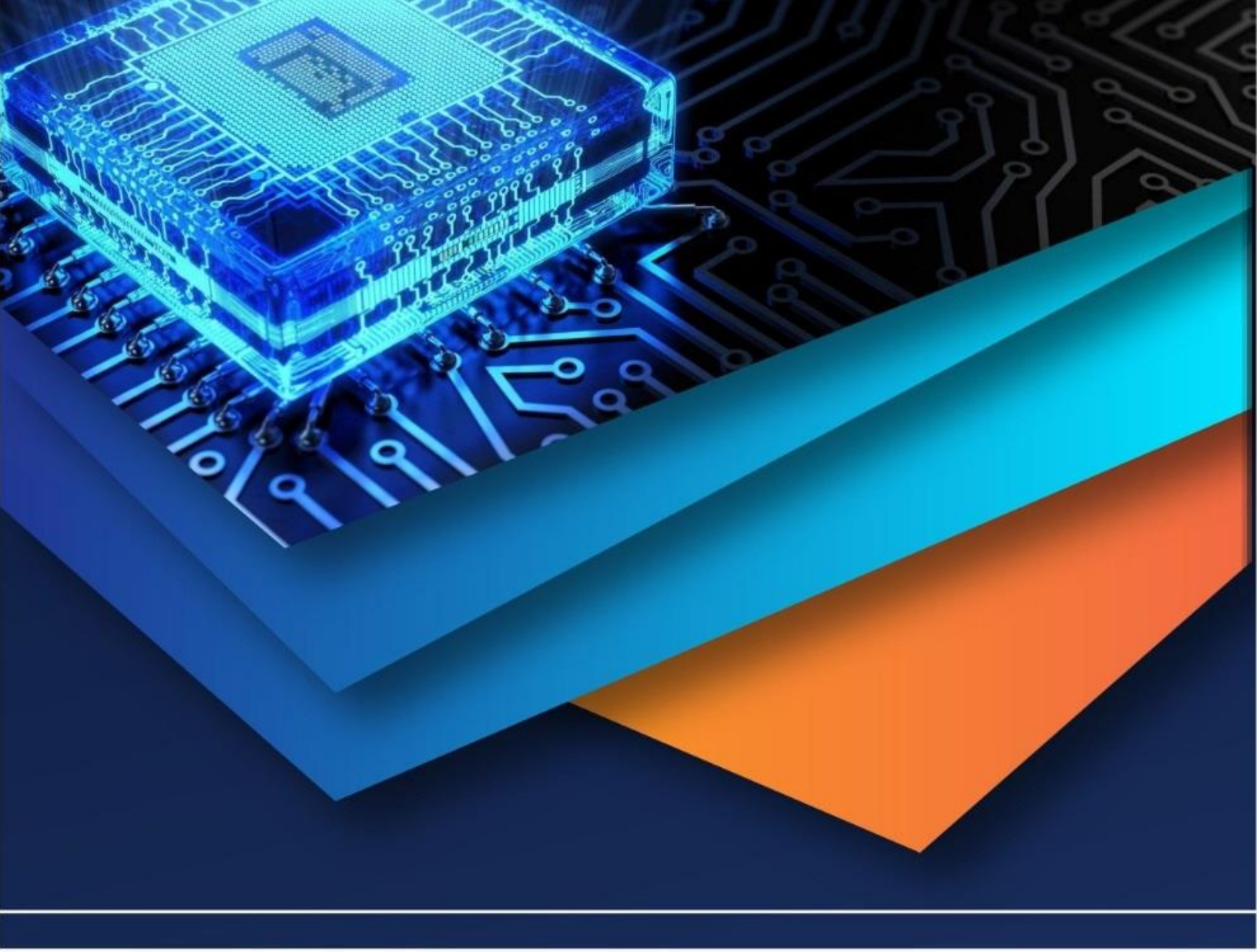
[6] Vorobev, Vladimir, and Kuznetsov, Maxim. A paraphrasing model based on ChatGPT paraphrases. Proceedings of the [conference name not specified], 2023.

[7] Abhay Shukla, Paheli Bhattacharya, Soham Poddar, Rajdeep Mukherjee, Kripabandhu Ghosh, Pawan Goyal, & Saptarshi Ghosh. (2022). Legal Case Document Summarization: Extractive and Abstractive Methods and their Evaluation [Data set].The 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (AACL-IJCNLP).

[8] Alami Merrouni, Zakariae & Frikh, Bouchra & Ouhbi, Brahim. (2023). EXABSUM: a new text summarization approach for generating extractive and abstractive summaries. Journal of Big Data. 10. 10.1186/s40537-023-00836-y.

[9] Kusabi, V., Jain, P., Palesha, R., Khandelwal, K., & Potdar, G. P. (April 2024). Advancements in Legal Document Processing and Summarization. International Journal of Novel Trends and Innovation (IJNTI), 2(4), a158–a166. ISSN 2984-908X.

[10] Liu Y. Fine-tune bert for extractive summarization. arXiv preprint arXiv:1903.10318 2019.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY