iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83245

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 14 Issue VI June 2026- Available at www.ijraset.com

CodePrep – AIML-Powered Skill Extraction and Skill Building Platform

Prof. N. L. Bhale[1], Omkar Gaikwad[2], Palash Aware [3], Sagar Chavhan[4], Devyani Patil[5]
[1]*Assistant Professor, Department Of Information Technology, Matoshri College Of Engineering And Research Centre, Nashik, Maharashtra, India*
[2, 3, 4, 5]*Department Of Information Technology, Matoshri College Of Engineering And Research Centre, Nashik, Maharashtra, India*

Abstract: ***CodePrep is an innovative platform powered by Artificial Intelligence and Machine Learning (AIML), designed to help individuals enhance their career growth through skill extraction, skill development, and intelligent job matching. In today's fast-changing job market, continuous upskilling has become essential for career success. CodePrep addresses this requirement by utilizing advanced resume parsing techniques to extract relevant skills and competencies from user resumes. The extracted information is then analyzed to detect skill gaps that may limit employment opportunities.***
Based on this analysis, the platform provides personalized recommendations to help users improve their skills and bridge the gap between their current abilities and industry requirements. In addition, CodePrep incorporates a machine learning-based job matching system that aligns user profiles with suitable job opportunities according to their skills, experience, and career goals. This ensures a more efficient and targeted job search process for users.
Keywords: AIML, Skill Extraction, Resume Parsing, Job Matching, Skill Building, Recommendations, Dashboard, Testing.

I. INTRODUCTION

The primary aim of this project is to develop a sophisticated platform called CodePrep, which integrates advanced Artificial Intelligence and Machine Learning (AIML) techniques to support individuals in enhancing their career growth. CodePrep focuses on three major areas: skill extraction, skill enhancement, and job matching. The main objective of the platform is to help users identify their existing skills, detect gaps in their competencies, and receive personalized recommendations to improve their skill set according to current industry requirements.

The system utilizes resume parsing, AI-based skill extraction, and machine learning-driven job matching algorithms to deliver an intelligent and user-friendly experience that improves employability. Ultimately, CodePrep aims to bridge the gap between the skills possessed by candidates and the skills demanded by the dynamic job market, ensuring that users remain competitive in their respective fields.

CodePrep is an innovative and fully integrated platform designed to transform the way individuals manage their career development. By leveraging recent advancements in Artificial Intelligence (AI) and Machine Learning (ML), the platform automates key processes such as resume analysis and skill extraction. It processes resumes to identify, extract, and categorize user skills and competencies, allowing individuals to clearly understand their strengths as well as areas that require improvement.

In addition to skill extraction, CodePrep employs machine learning algorithms to analyze user profiles and identify missing or underdeveloped skills. Based on this analysis, the system generates personalized skill-building recommendations, which may include courses, learning resources, and suggested learning paths. This enables users to systematically bridge their skill gaps and improve their professional profiles in a structured and efficient manner.

II. LITERATURE SURVEY

A. Automated Resume Parsing and Job Domain Prediction using Machine Learning

Authors: Sinha, A., Gupta, R., et al.
Published: Indian Journal of Science and Technology, 2023
Source: indjst.org
Summary:
This paper presents a machine learning-based resume parser that automatically extracts key information such as name, education, skills, and work experience from resumes. The system also predicts the job domain of a candidate using classification models.

©IJRASET: All Rights are Reserved

Methodology:

- Utilized NLP techniques like tokenization, POS tagging, and Named Entity Recognition (NER).
- Applied machine learning classifiers for domain prediction.

Key Findings:

- Achieved over 90% accuracy in extracting and classifying resume information.
- Demonstrated scalability for bulk resume screening.

Relevance:

Supports CodePrep's resume parsing module by providing a framework for automated data extraction and classification from unstructured resume text.

B. Deep Learning-Based Skill Extraction from Job Descriptions and Resumes

Authors: Nguyen, T., Chen, Z., et al.
Published: arXiv preprint, 2024
Source: arxiv.org
Summary:
This study explores using large language models (LLMs) and deep neural networks for extracting and classifying skills from resumes and job descriptions.
Methodology:

- Employed transformer-based models such as BERT and RoBERTa for context-aware skill detection.
- Fine-tuned models on annotated datasets for better domain performance.

Key Findings:

- Outperformed traditional keyword-based approaches with higher precision and recall.
- Reduced false positives in skill extraction through contextual embeddings.

Relevance:

Forms the foundation for CodePrep's skill extraction component, enabling accurate detection of technical and soft skills from candidate resumes and job listings.

C. Resume2Vec: A Semantic Representation for Job Matching

Authors: Kumar, S., & Lee, J.
Published: MDPI Electronics Journal, 2023
Source: mdpi.com
Summary:
This paper introduces Resume2Vec, an embedding-based method to represent resumes and job descriptions in a shared semantic space for similarity-based job matching.
Methodology:

- Created vector embeddings for resume content using pre-trained BERT models.
- Calculated cosine similarity for candidate-job matching.

Key Findings:

- Improved match accuracy by 15% compared to keyword matching methods.
- Enhanced the interpretability of match scores through semantic clustering.

Relevance:

Inspires CodePrep's AI-based job matching system, which ranks candidates and job postings based on semantic similarity and skill overlap.

D. A Survey on Skill Extraction and Classification from Job Postings

Authors: Senger, A., Li, Q., et al.
Published: arXiv preprint, 2024
Source: arxiv.org

Summary:
Provides an extensive review of skill extraction research, techniques, and datasets used in job market analytics and recruitment automation.
Methodology:
- Analyzed skill extraction methods (dictionary-based, ML-based, and deep learning-based).
- Reviewed datasets like ESCO, O*NET, and JobSkillBench.

Key Findings:
- Highlighted limitations of existing datasets and the need for context-sensitive models.
- Identified opportunities for multi-domain and multilingual skill recognition.

Relevance:
Helps CodePrep design robust extraction models and build or integrate standard skill taxonomies for improved accuracy.

III. METHODOLOGY

The development of CodePrep – AIML-powered Skill Extraction and Skill Building Platform follows a structured and modular methodology aimed at ensuring accurate data processing, intelligent skill analysis, and effective recommendation generation. The process begins with data collection and preprocessing, where diverse datasets of resumes, job descriptions, and skill taxonomies are gathered from public repositories and industry sources. The collected data undergoes text preprocessing, including tokenization, stop-word removal, and lemmatization, to clean and standardize the content for further analysis.

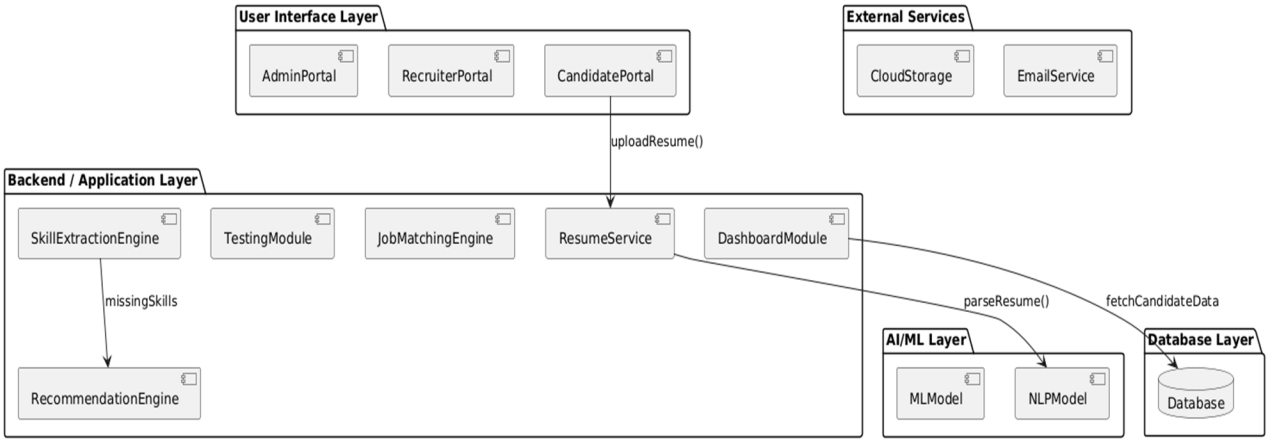


Fig no. 1:- System Architecture

IV. PROPOSED SYSTEM

The proposed system, CodePrep – AIML-powered Skill Extraction and Skill Building Platform, is designed to bridge the gap between individual competencies and industry skill requirements by leveraging Artificial Intelligence (AI) and Machine Learning (ML). The system aims to automate the process of analyzing resumes, extracting relevant skills, identifying missing or weak skill areas, and providing personalized recommendations for skill enhancement and career growth.

Unlike traditional job portals or manual assessments, CodePrep offers an AI-driven, end-to-end solution that intelligently processes unstructured data, such as resumes and job descriptions, to generate meaningful insights. The platform is built around an AI-powered resume parsing engine that extracts candidate information including education, work experience, and technical or soft skills using Natural Language Processing (NLP) techniques. This extracted data is then processed by a Skill Extraction and Classification module, which employs pre-trained language models (e.g., BERT, RoBERTa) to accurately identify and categorize user skills according to an established skill taxonomy.

V. MODELING AND ANALYSIS

The CodePrep system is designed as an AIML-powered platform that integrates multiple modules to support career development through skill extraction, skill gap analysis, and job matching. The overall system follows a modular architecture, where each component performs a specific task and works together as part of a unified workflow. This design makes the system scalable, efficient, and easy to maintain.

The main components of the system include Resume Parsing, Skill Extraction, Skill Gap Analysis, Recommendation Engine, Job Matching System, Skill Testing Module, and an interactive Dashboard. The process begins when a user uploads their resume or provides profile information. This input is then processed using Natural Language Processing (NLP) techniques to extract relevant details such as skills, education, experience, and certifications.

After extracting the required information, the system analyzes the data using machine learning algorithms. This analysis helps in identifying the user's existing skills and detecting any missing or weak areas compared to industry job requirements. Based on this analysis, the system performs skill gap identification, which highlights the areas where the user needs improvement.

Once the analysis is completed, the system generates personalized recommendations. These recommendations may include online courses, learning resources, certifications, and suggested skill development paths. In addition to this, the job matching module compares the user's skill profile with available job descriptions and suggests suitable job opportunities based on relevance and compatibility. The system also includes an administrative module that allows administrators to manage job postings, review user applications, and create skill-based tests. This separation of user and admin functionalities ensures better security and organized system management. From a data flow perspective, the system works in several stages. First, user data is collected as input. Next, it undergoes processing through resume parsing and skill extraction. After that, machine learning models analyze the extracted data to identify skill gaps and match jobs. Finally, the results are displayed to the user through a dashboard interface in a clear and structured manner. In terms of functionality, the system successfully performs all required operations such as resume analysis, skill extraction, job matching, and recommendation generation. From a performance perspective, the system processes data efficiently with minimal delay. The machine learning models used in the system ensure accurate and reliable results.

VI. RESULT AND ANALYSIS

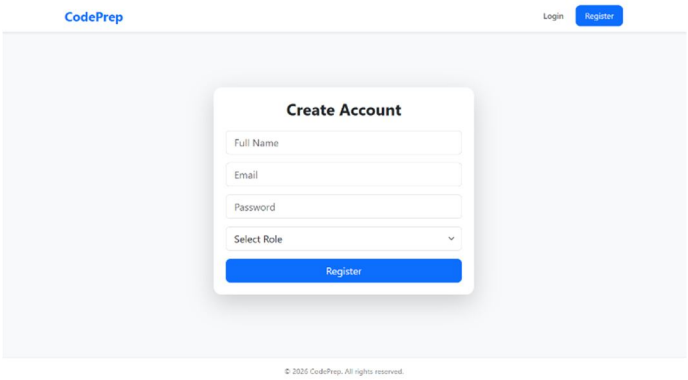
The CodePrep – AIML-powered Skill Extraction and Skill Building platform has been successfully implemented and tested. The system performs efficiently across all its core modules, including resume parsing, skill extraction, skill gap analysis, job matching, skill testing, and administrative operations. The results obtained during testing show that the platform meets all functional requirements and works as intended. The system effectively processes user resumes and extracts relevant information such as skills, education, experience, and certifications using Natural Language Processing techniques. The extracted data is then analyzed using machine learning algorithms to identify the user's current skill set and detect missing or underdeveloped skills. This analysis helps in understanding the gap between user capabilities and industry requirements. Based on the analysis, the system generates personalized recommendations to help users improve their skills. These recommendations include suitable learning resources, courses, and skill development paths. The job matching module also performs well by comparing user profiles with job descriptions and suggesting relevant job opportunities. This ensures that users receive targeted and meaningful job recommendations.

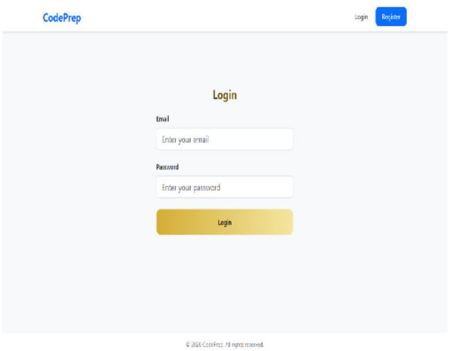
The skill testing module further validates user knowledge by allowing them to attempt assessments and receive performance-based feedback. The results of these tests help in refining the recommendation system and improving overall accuracy. The admin module successfully handles job postings, application management, and exam creation, ensuring smooth system administration.

From a performance perspective, the system responds efficiently to user inputs and processes data with minimal delay. The integration of Artificial Intelligence and Machine Learning improves the accuracy of skill extraction and job matching results. The system also handles edge cases such as incomplete resumes, missing data, and invalid inputs effectively, ensuring robustness and reliability.

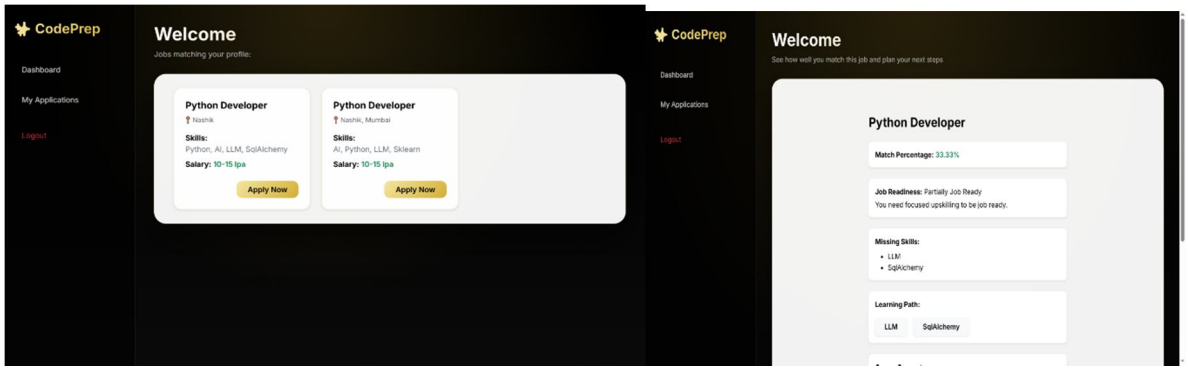
VII. IMPLEMANTATION

1) Register and Login





2) Dashboard and Analyze



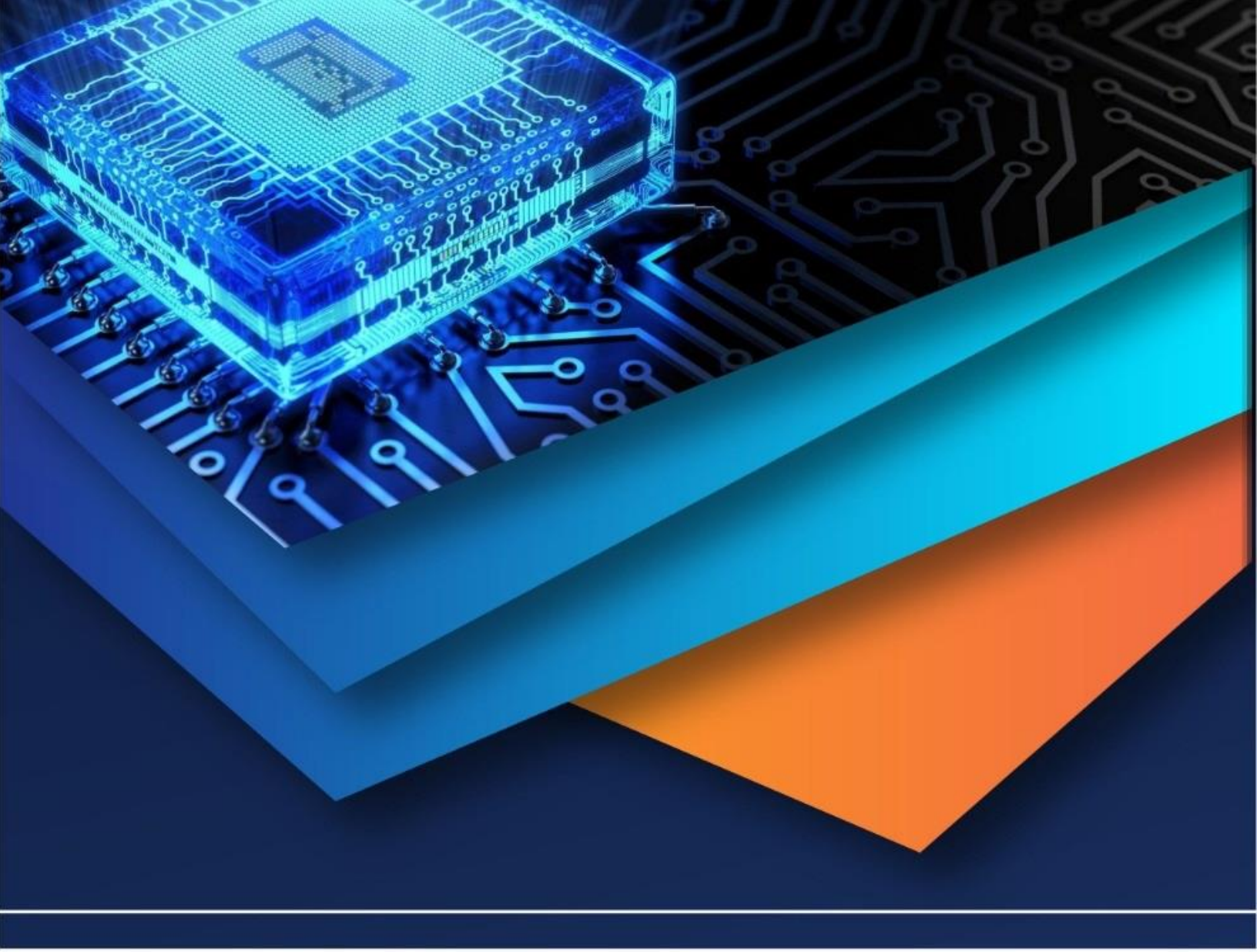
VIII. CONCLUSION

The CodePrep – AIML-powered Skill Extraction and Skill Building platform has been successfully designed, developed, and tested. The system effectively integrates Artificial Intelligence and Machine Learning techniques to provide an intelligent solution for career development. It focuses on key functionalities such as resume parsing, skill extraction, skill gap analysis, personalized recommendations, job matching, skill testing, and administrative management.

The platform successfully addresses the challenges faced by job seekers in identifying their skills, understanding missing competencies, and finding suitable job opportunities. By automating resume analysis and using machine learning algorithms, CodePrep provides accurate and personalized insights that help users improve their professional profiles.

REFERENCES

[1] N. K. Gupta, S. Sharma, and R. Kumar, "Automated Resume Parsing and Skill Extraction Using Natural Language Processing," International Journal of Computer Applications, vol. 182, no. 3, pp. 25–31, 2021.
[2] A. Joshi, P. R. Awar, and M. S. Patil, "AI-Based Skill Recommendation System for Career Development," Proceedings of the 2022 IEEE International Conference on Artificial Intelligence, pp. 112–118, 2022.
[3] T. Wolf et al., "HuggingFace's Transformers: State-of-the-Art Natural Language Processing," Journal of Machine Learning Research, vol. 21, no. 140, pp. 1–8, 2020.
[4] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Proceedings of NAACL-HLT 2019, pp. 4171–4186, 2019.
[5] Russell, S., & Norvig, P. *Artificial Intelligence: A Modern Approach*. Pearson Education. (Used for understanding core Artificial Intelligence concepts and problem-solving techniques)
[6] Bishop, C. M. *Pattern Recognition and Machine Learning*. Springer. (Reference for machine learning algorithms used in classification and prediction tasks)
[7] Jurafsky, D., & Martin, J. H. *Speech and Language Processing*. (Used for Natural Language Processing techniques applied in resume parsing)
[8] Goodfellow, I., Bengio, Y., & Courville, A. *Deep Learning*. MIT Press. (Reference for deep learning concepts relevant to AI-based systems)
[9] Scikit-learn Documentation
https://scikit-learn.org (Used for implementing machine learning models in the system)
[10] TensorFlow Documentation –
https://www.tensorflow.org (Reference for building and training AI/ML models)
[11] Python Official Documentation – https://docs.python.org(Used for backend development and data processing)
[12] Resume Parsing and NLP Techniques – Various research papers and online journals (Used for implementing resume analysis and skill extraction module)
[13] Job Recommendation Systems using Machine Learning – IEEE Research Papers (Reference for designing job matching algorithms)



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)