IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 3, May-June 2025

Impact Factor: 8.152

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203156

Embedded Handwritten OCR System for Test Paper Digitization

H. Madhuri, Dr. K. Naveen Kumar, B. Laxman, M. Sriharshitha, P. Yasmeen, S. Niveditha, T. Laxmi Pravallika

B.Tech, Department of Electronics and Communication Engineering, Pallavi Engineering College, Hyderabad, Telangana, India

Professor & HOD, Department of Electronics and Communication Engineering, Pallavi Engineering College, Hyderabad, Telangana, India

Assistant Professor, Department of Electronics and Communication Engineering, Pallavi Engineering College, Hyderabad, Telangana, India

B.Tech, Department of Electronics and Communication Engineering, Pallavi Engineering College, Hyderabad, Telangana, India

B.Tech, Department of Electronics and Communication Engineering, Pallavi Engineering College, Hyderabad, Telangana, India

B.Tech, Department of Electronics and Communication Engineering, Pallavi Engineering College, Hyderabad, Telangana, India

B.Tech, Department of Electronics and Communication Engineering, Pallavi Engineering College, Hyderabad, Telangana, India

ABSTRACT: This project develops an Embedded Handwritten OCR System using AI to digitize and evaluate handwritten test answers. It captures handwritten responses, processes them through image preprocessing and OCR, and uses Google's Gemini Vision API to compare answers with ideal responses, providing structured feedback. Built on a Raspberry Pi 3B+ with a camera module and LCD display, the system offers a portable and efficient solution for real-time processing. A user-friendly interface is developed using Python Flask, enabling teachers to evaluate descriptive answers accurately and automatically, reducing manual effort and improving grading consistency.

KEYWORDS: Handwritten OCR, Gemini Vision API, Descriptive Answer Evaluation, Multimodal AI, Image Preprocessing, Raspberry Pi 3B+, Flask Web Framework, Prompt Engineering, Vision-Language Models, Automated Grading System.

I. INTRODUCTION

GRACE is an AI-powered system that automates the assessment of descriptive student responses. It uses multimodal AI, specifically Google's Gemini Vision model, to evaluate handwritten answers based on conceptual alignment with ideal responses. The system provides structured feedback, marks, and remarks, and allows for custom rules and batch processing. GRACE's design emphasizes modularity, transparency, accessibility, and security, with features like input sanitization, API key management, and session expiration. The system aims to reduce evaluator workload, speed up assessments, and ensure uniform scoring, while providing a user-friendly and efficient experience for educators.

II. EXISTING METHODS

Traditional OCR systems struggled with handwritten text due to varying handwriting styles. However, deep learning advancements, particularly CNNs and RNNs, have improved handwritten text recognition. Models like CRNN and TrOCR have achieved significant results. Embedded OCR systems face constraints, but research on model compression

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203156

and edge computing enables efficient deployment on platforms like Raspberry Pi. Automated exam evaluation systems are evolving, with newer solutions using OCR and NLP to grade subjective answers, paving the way for more efficient and accurate assessment tools.

III. BACKGROUND AND TECHNOLOGICAL FOUNDATION

The field of automated evaluation has undergone significant evolution, from manual grading to Optical Mark Recognition (OMR) systems for multiple-choice questions. However, with the focus on deeper learning outcomes, the limitations of such systems became evident. The emergence of Natural Language Processing (NLP) and deep learning models enabled machines to process context, syntax, and intent, paving the way for evaluating descriptive answers.
Multimodal AI and Vision-Language Models

Multimodal AI combines inputs from multiple sources, such as text, images, and audio. Vision-language models, like Google's Gemini Vision, can interpret and describe visual scenes using textual descriptions. These models are valuable in educational contexts, as they can analyse handwritten answers and compare them to model answers, grading based on conceptual similarity.

Gemini Vision API and Capabilities
The Gemini Vision API is a significant leap forward in multimodal AI, allowing developers to input text and image data for analysis. It acts as a cognitive engine for the GRACE system, providing well-structured responses with marks, justification, and suggestions for improvement.

Design Approach

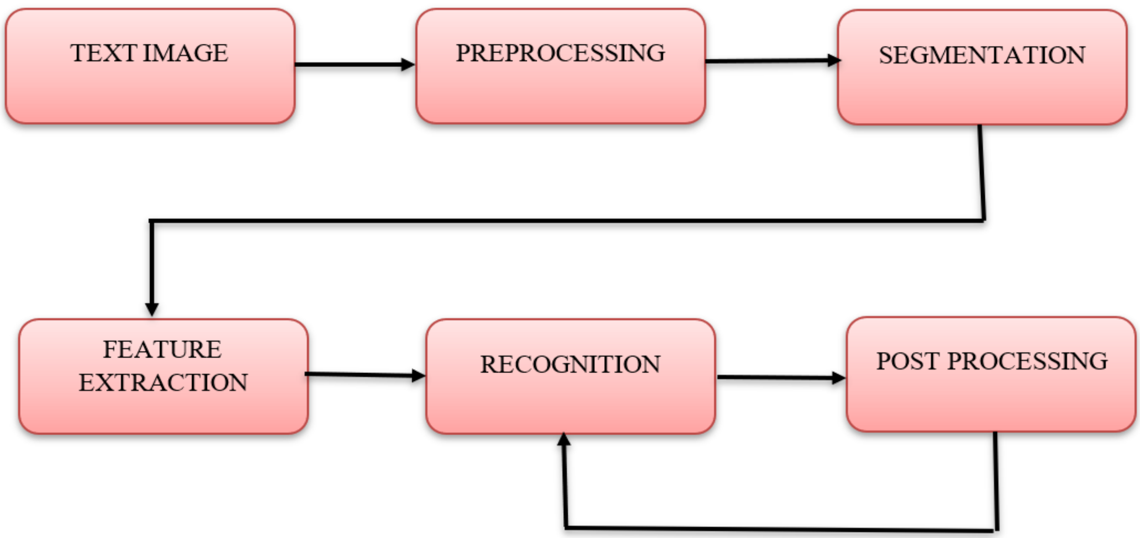

Fig: 1 Design Approach

The design approach involves several steps:
- Text Image: Input image of handwritten text.
- Preprocessing: Improve image quality and remove noise.
- Segmentation: Break image into lines, words, and characters.
- Feature Extraction: Identify key features from characters.
- Recognition: Convert features into actual characters using machine learning models.
- Post Processing: Improve accuracy and make corrections.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203156

IV. BLOCK DIAGRAM

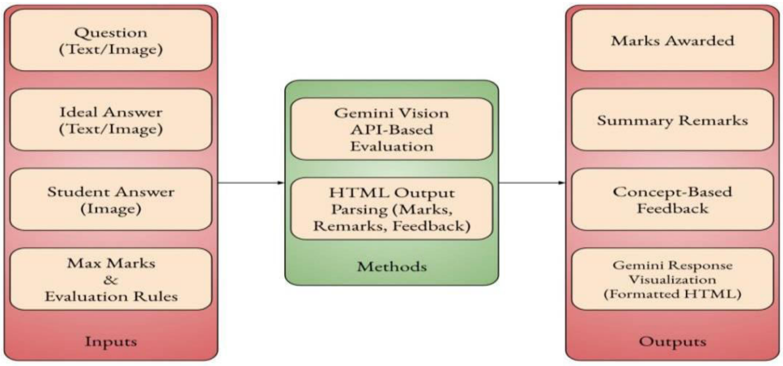

Fig: 2 Block Diagram

Inputs:	**Methods:**	**Outputs:**
- Question (text/image) - Ideal answer (text/image) - Student answer (image) - Max marks and evaluation rules	- Gemini Vision API compares student answer with ideal answer - HTML output parsing extracts marks, remarks, and feedback	- Marks awarded - Summary remarks - Concept-based feedback - Formatted HTML visualization of results

V. SYSTEM ARCHITECTURE AND DESIGN FLOW

System Overview

The GRACE system is an AI-powered tool for evaluating student responses, integrating frontend usability, backend processing, and AI inference engines for seamless and scalable evaluation.

High-Level Architecture

The GRACE architecture consists of three layers:

1. Frontend: User interaction layer built using HTML with responsive features.
2. Backend: Powered by Python and Flask, handling tasks like input processing and communication with the AI model.
3. AI Layer: Uses Google's Gemini Vision API to analyse and evaluate student responses.

Deployment and Environment Setup

GRACE is platform-independent, deployable on Windows, Linux, and Raspberry Pi environments with a lightweight setup requiring basic Python packages. The system is designed for plug-and-play portability, making it suitable for educational institutions with varying technical resources.

Fig: 3 Grace Setup

International Journal of Advanced Research in Education and TechnologY(IJARETY)

IJARETY | ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203156

VI. FRONTEND DESIGN AND INTERFACE

The GRACE system features a user-friendly interface with three main pages:
1. Welcome Page: A clean and mobile-optimized landing page with a call-to-action button.
2. Evaluation Form: A structured form for collecting question, ideal answer, and student answer, with dynamic image previews and validation.
3. Result Display Page: A page displaying marks, feedback, and full Gemini output in HTML format, with a retry button.

Module-Wise Implementation
The system consists of several modules:
1. Image Acquisition and Preprocessing: Captures and preprocesses student answer sheets.
2. OCR and Handwriting Recognition: Uses Gemini Vision API for handwritten text recognition.
3. Input Management and Prompt Generation: Handles input collection and prompt construction.
4. AI Evaluation and Response Parsing: Sends prompts to Gemini API and parses responses.
5. Result Display and User Interface: Displays results in a clean and readable layout.

Backend Logic and AI Integration
The backend is built on Flask, with features like:
1. Flask Routing: Manages user interactions and routes.
2. File Handling and Secure Upload Pipeline: Handles file uploads securely.
3. Prompt Engineering and Instructional Formatting: Constructs detailed prompts for AI evaluation.
4. Gemini API Integration and Session Management: Integrates with Gemini Vision API for evaluation.
5. Output Parsing and HTML Processing: Parses and sanitizes AI responses.
6. Result Rendering and Feedback Integration: Renders results with dynamic content.

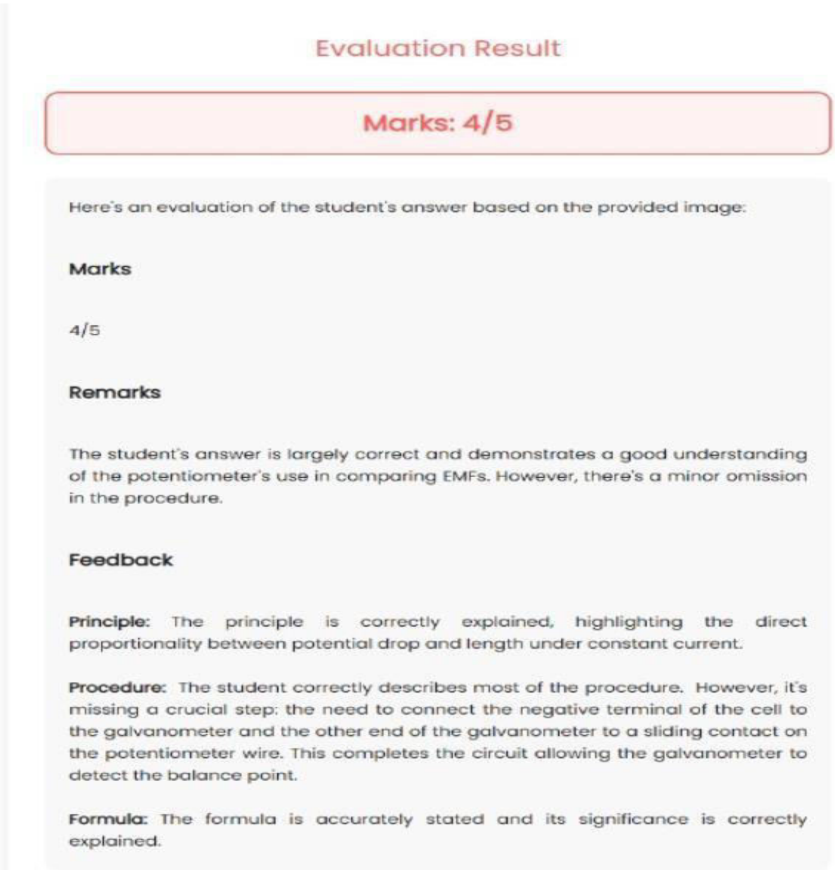

Fig: 4 Evaluation Result

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203156

VII. FLOW CHART

The system follows a structured flow:
1. Upload question, ideal answer, and student answer.
2. Generate prompt and send to Gemini API.
3. Receive and parse response.
4. Render results with marks, feedback, and comments.

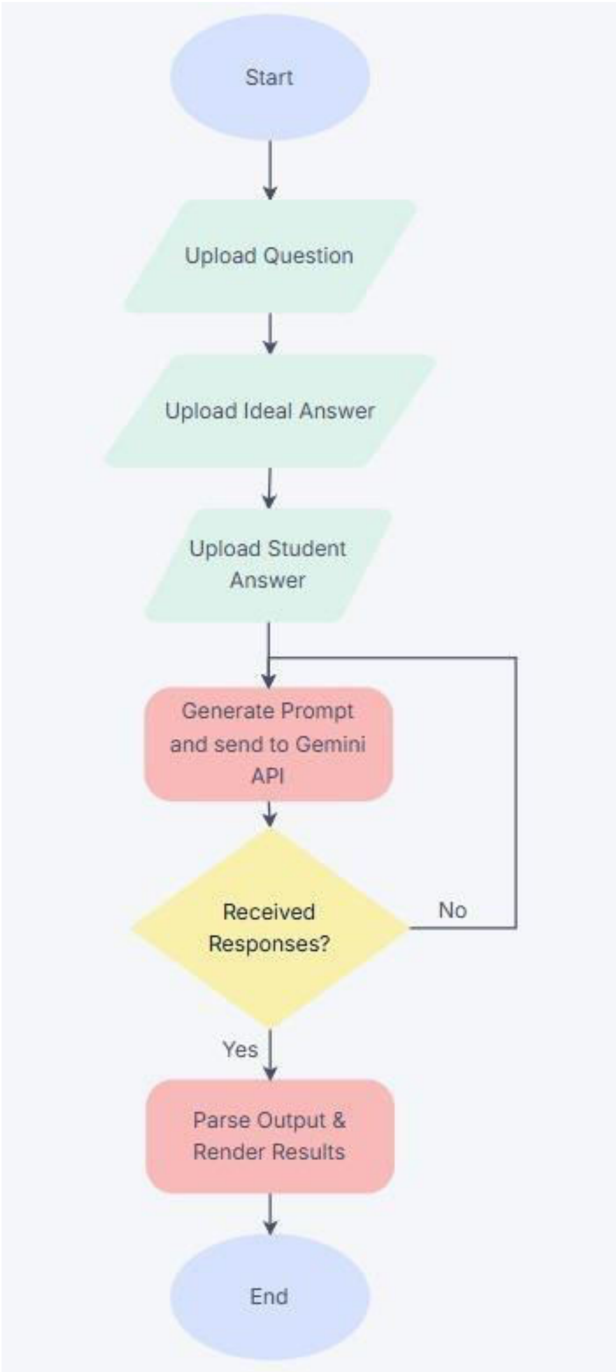

Fig: 5. Flow chart

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May - June 2025 ||

DOI:10.15680/IJARETY.2025.1203156

VIII. HARDWARE COMPONENTS

The Raspberry Pi 3B+ is a low-cost, single-board computer designed for education and embedded system applications. It features a 1.4 GHz quad-core processor, 1 GB RAM, dual-band Wi-Fi, and Bluetooth 4.2. The board is compact, affordable, and supports Linux-based operating systems, making it ideal for projects like handwritten paper digitization**.**

History of Raspberry Pi

The Raspberry Pi Foundation was established in 2008 to promote computer science education. The first prototypes were developed in 2011, and the first commercial Raspberry Pi, Model B, was launched in 2012 with an initial price of $35. The Raspberry Pi quickly gained popularity, and its recommended operating system, Raspbian, was optimized for the device. Today, the Raspberry Pi is widely used in education and embedded system

Fig 9.1: Raspberry Pi 3B+ Circuit

IX. IMPLEMENTATION AND RESULTS

1. System Integration: Seamless integration of UI, backend logic, and AI evaluation engine using Flask.
2. Multi-Platform Deployment: Responsive design ensures smooth functionality across devices.
3. Evaluation Flow: Structured evaluation process with prompt generation and AI-driven assessment.
4. Result Rendering and UI: Clear and readable result display with structured feedback.
5. Testing and Robustness: Extensive testing demonstrated resilience and accuracy in handling various input scenarios.

X. CONCLUSION

The backend logic of GRACE brings together lightweight web routing, secure file handling, advanced prompt engineering, and generative AI evaluation to form a cohesive academic grading tool. Each part of the system — from Flask routes to Gemini API interaction has been modularly designed for clarity, extensibility, and stability.Through its seamless integration of text and image-based inputs, GRACE is capable of delivering meaningful, structured feedback that mimics the reasoning of a human evaluator. The choice of a controlled HTML structure, combined with dynamic prompt generation and rigorous post-processing, ensures consistent results that are easy to interpret.This chapter has outlined how these backend components work in unison to turn user inputs into intelligent evaluations. The architecture not only serves its current purpose but is scalable for features like batch processing, rubric-based scoring, and integration with learning management systems in future versions of GRACE.

XI. FUTURE ENHANCEMENTS

Despite its current capabilities, GRACE has immense scope for future upgrades that can make it even more versatile and powerful.In many real-world academic settings, evaluators work with scanned PDF documents — especially in online exams or centralized assessments. Introducing support for PDF-based input would allow users to upload full answer sheets in a standard format. GRACE could then extract each page or segment as individual responses using a pre-processing step. This would reduce manual segmentation and enhance usability.Currently, GRACE supports one student evaluation per run. A future enhancement involves allowing users to upload folders containing multiple answer sheets. GRACE could automatically loop through the input folder, perform evaluations on each file, and generate a cumulative report. This batch mode would be ideal for school or college-level evaluations and save time during high-volume assessments.To go beyond surface-level image analysis, future versions of GRACE could employ deep learning models for layout parsing. This includes identifying zones of interest in handwritten pages, detecting answer headers, separating multiple answers, and even tracking corrections orhighlighted margins. Such structured understanding of the page can improve evaluation accuracy and mimic the granularity of human assessment.

Adding OCR capabilities will allow GRACE to extract text directly from images or PDF files, enabling it to compare content with rubrics dynamically. For example, it could identify question numbers, keywords, and structure without needing manually entered ideal answers. This opens the door to rubric-based evaluation workflows, automated tagging, and further objective scoring mechanisms.In the long term, GRACE can evolve into a Software-as-a-Service (SaaS) solution or a standalone mobile application. A cloud-based deployment will allow integration with Learning Management Systems (LMS), auto-login for teachers and students, and centralized analytics dashboards. A mobile version can bring the platform to users in rural or low-resource environments, making real-time AI evaluation accessible anytime, anywhere — especially during educational outreach or teacher training program

REFERENCES

[1] Dr. KOMMU NAVEEN, Dr. R.M.S PARVATHI, on “A Review of Deep Learning Applications for Speech Processing Improvement & Applications” “Journal of Telecommunication, Switching Systems and Networks (JOTSSN)” | ISSN: 2454-637, p-ISSN: 2455-638| www.jotssn.com | Impact Factor: 7.569| || Volume 9, Issue 2, September 2022 || | DOI: DOI (Journal): 10.37591/JoTSSN/|http://engineeringjournals.stmjournals.in/index.php/JoTSSN/index

[2] Dr. KOMMU NAVEEN, Dr. R.M.S Parvathi, on “A Comprehensive Review On Machine Learning Applications of Convolutional Neural Networks to Medical Image Analysis” at ‘International Conference on Robotics and Communication Technology-ICRCT-2022)’ Page No: 50-57, ISSN : 2708-1079; IC Value: 45.98; Impact Factor: 7.379; VOLUME-03, ISSUE-1; 4th January, 2022. URL:https://doi.org/10.46379/jscer.2022.030104

[3] Hatice Cinar Akakin and Metin N. Gurcan, “Content Based Microscopic Image Retrieval System for Multi-Image Queries”, IEEE Transactions On Information Technology In Biomedicine, VOL. 16, NO. 4, JULY 2012

[4] Yong-Hwan Lee and Sang-Burm Rhee, Bonam Kim, “Content- based Image Retrieval Using Wavelet Spatial-Color and Gabor Normalized Texture in Multi-resolution Database”, 978-0- 7695- 4684-1/12 $26.00 © 2012 IEEE, DOI 10.1109/IMIS.2012.98

[5] Dr. KOMMU NAVEEN, Dr. R.M.S Parvathi, on “Contract and Feature Extraction of CBIR Method Using Soft Computing Techniques in Machine Learning” at ‘Journal of Instrumentation Technology and Innovations)’ Page No: 17-27; ISSN (PRINT): 2249-4731; ISSN (ONLINE): 2347-7261 VOLUME-12; ISSUE-01; April-May -2022. DOI(journal):10.37591/JoITI, URL: http:www.engineeringjournals.stmjournals.in/index.php/JoITI/index.

[6] S. Mangijao Singh , K. Hemachandran , “Content-Based Image Retrieval using Color Moment and Gabor Texture Feature”, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1, September 2012; ISSN (Online): 1694- 0814

[7] Dr. K. Naveen Kumar, Dr. M. B. Raju, Prof. A. Sridhar, IJRAR August 2024, Volume 11, Issue 3 www.ijrar.org (E-ISSN 2348-1269, P- ISSN 2349-5138) IJRAR24C2120 , International Journal of Research and Analytical Reviews (IJRAR) 2024 IJRAR August 2024, Volume 11, Issue 3 www.ijrar.org (E-ISSN 2348-1269, P- ISSN 2349-5138).

[8] Kanchan Saxena, Vineet Richaria, Vijay Trivedi,“A Survey on Content Based Image Retrieval using BDIP, BVLC AND DCD”, Journal of Global Research in Computer Science , Vol.3, No. 9, September 2012 ,ISSN-2229-371X.

[9] Dr. K. Naveen Kumar, Dr. M. B. Raju, Prof. A. Sridhar on International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 12 Issue VIII Aug 2024- Available at www.ijraset.com.

[10] Dr. KOMMU NAVEEN, Dr. R.M.S Parvathi, on “A Review On Content Based Image Retrieval Systems Features derived by Deep Learning Models” at ‘International Journal For Research In Applied Science And Engineering Technology (IJRASET)’ Page No: 42-57, ISSN : 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429; VOLUME-09, ISSUE-XII; November- 2021. URL:www.ijraset.com/index.html/doi.org.

IJARETY

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com