PRESIDENCY UNIVERSITY

Private University Estd. in Karnataka State by Act No. 41 of 2013

Itgalpura, Rajankunte, Yelahanka, Bengaluru – 560064

EduNLP: A RULE-BASED MULTILINGUAL SMART EDUCATION SYSTEM FOR RURAL LEARNING

A PROJECT REPORT

Submitted by

K SAI SRI - 20221CSE0484

MEGHANA K - 20221CSE0446

SHAIK MOHAMMED IMRAN - 20221CSE0061

Under the guidance of,

Dr. PREMA AROKIA MARY G

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

PRESIDENCY UNIVERSITY

BENGALURU

DECEMBER 2025

PRESIDENCY UNIVERSITY
Private University Estd. in Karnataka State by Act No. 41 of 2013
Itgalpura, Rajankunte, Yelahanka, Bengaluru – 560064

PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

Certified that this report "EduNLP: A Rule-Based Multilingual Smart Education System for Rural Learning" is a bonafide work of "K Sai Sri (20221CSE0484), Shaik Mohammed Imran (20221CSE0061), Meghana K (20221CSE0446) ", who have successfully carried out the project work and submitted the report for partial fulfilment of the requirements for the award of the degree of BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE ENGINEERING during 2025-26.

27/11/25
Dr. Prema Arokia Mary G
Project Guide
PSCS
Presidency University

27/11/25
Mr. Muthuraju V
Program Project Coordinator
PSCS
Presidency University

Dr. Sampath A K
Dr. Geetha A
School Project Coordinators
PSCS
Presidency University

Dr. Blessed Prince
Head of the Department PSCS
Presidency University

Dr. Shakkeera L
Associate Dean
PSCS
Presidency University

N.
Dr. Duraipandian N
Dean
PSCS & PSIS
Presidency University

Examiners

Sl. no.	Name	Signature	Date
1	Dr. Ramesh Sengodan		28.11.2025
2	Dr. Denslin Brabin D R		28.11.25

PRESIDENCY UNIVERSITY

PRESIDENCY SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

We the students of final year B.Tech in COMPUTER SCIENCE ENGINEERING at Presidency University, Bengaluru, named K Sai Sri, Shaik Mohammed Imran, Meghana K, hereby declare that the project work titled **"EduNLP: A Rule-Based Multilingual Smart Education System for Rural Learning"** has been independently carried out by us and submitted in partial fulfillment for the award of the degree of B.Tech in COMPUTER SCIENCE ENGINEERING during the academic year of 2025-26. Further, the matter embodied in the project has not been submitted previously by anybody for the award of any Degree or Diploma to any other institution.

K Sai Sri	USN: 20221CSE0484	K·SaiSri
Meghana K	USN: 20221CSE0446	Meghana.K
Shaik Mohammed Imran	USN: 20221CSE0061	Imraf

PLACE: BENGALURU

DATE: 28/11/2025

ACKNOWLEDGEMENT

For completing this project work, we have received the support and the guidance from many people whom I would like to mention with deep sense of gratitude and indebtedness. We extend our gratitude to our beloved **Chancellor, Pro-Vice Chancellor, and Registrar** for their support and encouragement in completion of the project.

I would like to sincerely thank my internal guide **Dr. Prema Arokia Mary G, Associate Professor**, Presidency School of Computer Science and Engineering, Presidency University, for his/her moral support, motivation, timely guidance and encouragement provided to us during the period of our project work.

I am also thankful to **Dr. Blessed Prince, Professor, Head of the Department, Presidency School of Computer Science and Engineering** Presidency University, for his mentorship and encouragement.

We express our cordial thanks to **Dr. Duraipandian N,** Dean PSCS & PSIS, **Dr. Shakkeera** L, Associate Dean, Presidency School of computer Science and Engineering and the Management of Presidency University for providing the required facilities and intellectually stimulating environment that aided in the completion of my project work.

We are grateful to **Dr. Sampath A K, and Dr. Geetha A, PSCS** Project Coordinators, **Mr. Muthuraju V, Program Project Coordinator**, Presidency School of Computer Science and Engineering, or facilitating problem statements, coordinating reviews, monitoring progress, and providing their valuable support and guidance.

We are also grateful to Teaching and Non-Teaching staff of Presidency School of Computer Science and Engineering and also staff from other departments who have extended their valuable help and cooperation.

K SAI SRI

MEGHANA K

SHAIK MOHAMMED IMRAN

Abstract

The online revolution in the education sector has offered unprecedented chances of an inclusive learning experience, yet there are still several challenges facing rural areas within the developing world, particularly in the area of the lack of digital infrastructure, language barrier, and insufficient mentoring facilities. These challenges are the ones that require the EduNLP-A Rule-Based Multilingual Smart Education System to be applied in this project, intended to enhance rural learning. The system, in turn, fills the linguistic as well as technological disparities in providing customized educational assistance, mentoring, and learning content in local languages, hence promoting equity in obtaining good education.

Its project is developed using a well-organized V-model software development model that ensures that the process is systematic, beginning with the process of requirement analysis through to the implementation and validation. The system architecture will run on a current React-Spring Boot-PostgreSQL platform with the help of AWS S3 to store the data on a big scale and Python-based NLP modules to process the multilingual texts. The NLP engine is rule-based, thus offering contextual understanding to a wide variety of the Indian languages without the need to use large-scale machine learning models and therefore efficiency and explainability. The design focuses on a lightweight and energy-efficient deployment that is affordable to support low resource environments common in rural settings.

Full-scale testing confirmed the effectiveness of the system and obtained a 90 percent or more multilingual response accuracy and an average latency of less than 2 seconds at standard load conditions. The prototype has demonstrated strong functionality in many simulated rural deployment conditions as evidence of the concept of providing educational support over low-cost devices having low internet bandwidth. The project fulfills the goals of the UN Sustainable Development regarding inclusive education and digital empowerment to meet SDG 4 - Quality Education - and SDG 8 - Decent Work and Economic Growth. The findings in this regard confirm that EduNLP is a sustainable, scalable, and socially significant solution to reshape the learning ecosystem of rural communities. It can be improved in the future by adding more advanced AI-based personalization, speech recognition, and adaptive learning opportunities to make user experience and accessibility more efficient.

Table of Content

List of Figures

List of Tables

Abbreviations

Abbreviation	Full Form
AI	Artificial Intelligence
ANOVA	Analysis of Variance
API	Application Programming Interface
ASER	Annual Status of Education Report
DIKSHA	Digital Infrastructure for Knowledge Sharing
ES6	ECMAScript 6
HTML5	HyperText Markup Language 5
ICT	Information and Communication Technology
JSON	JavaScript Object Notation
LMS	Learning Management System
ML	Machine Learning
MOOCs	Massive Open Online Courses
NEP 2020	National Education Policy 2020
NLP	Natural Language Processing
REST	Representational State Transfer
SDG	Sustainable Development Goal
TLS	Transport Layer Security
TP1–TP5	Test Points 1 to 5
UI	User Interface
vCPU	Virtual CPU
VPS	Virtual Private Server
XR	Extended Reality

Chapter 1

Introduction

1.1 Background

Education plays the supporting role in socio-economic and cultural development of any country or nation [1], [2]. But the accessibility and quality of education still remain a constant problem in the rural regions of most developing nations such as India [1]. One of the primary causative factors is poor infrastructure, limited access to digital resources, and differences in the native languages of students [3]. Majority of the students in the rural India are unable to comprehend the learning content which is usually in the English language or in any other primary language thus creating a huge gap in learning between the urban and the rural residents.

Online learning platforms have demonstrated the potential of revolutionizing the conventional learning space to a more approachable and interactive ecosystem [4]. Most of these platforms, however, are made in cities, disregarding multilingual inclusivity and low bandwidth regions [5]. To resolve this deficiency, EduNLP offers a Rule-Based Multilingual Smart Education System that is aimed to provide equal access, communication, and lingual versatility to the students in the rural environment with assistance of NLP and language recognition applications [6].

The purpose of this system is to address the rural-urban education gap by including regional language learning support, automatic mentorship, and career awareness modules. It is in line with the overall goals of Sustainable Development Goals, SDG 4 (Quality Education) in promoting inclusivity and equitable access to education [7].

1.2 Statistics of project

It is reported that among the school-going children in India, more than 70% of children aged school-going years are rural residents and only 38% of school-going students in rural regions are literate at Grade 2 reading level in their local language [1]. In the same vein, the internet connectivity within the rural areas is hardly stable, and the index of digital readiness is minimal in comparison to the urban national average [2]. This also makes it hard to establish one educational platform that can accommodate all the communities in India that as a country has more than 121 major languages and 22 scheduled languages [2].

These figures show that there is an urgent requirement of a multilingual educational approach that can work with low-end hardware and less interconnection [3]. The need to have vernacular learning resources, online mentorship and career advice has multiplied particularly after the COVID-19 pandemic pointed at how vulnerable the traditional education system is in rural regions. To this end, the need to have an inclusive digital education model like EduNLP that accommodates the consideration of the issue of access, equity, and responsiveness to the local context is overwhelming.

1.3 Prior existing technologies

The government and individual organizations have launched various educational platforms to further roll out the digital learning to more individuals. The DIKSHA project run by the Government of India has made open educational resources available to teachers and students, but it is highly dependent on the consistent internet connection and does not track an individual level of progress [5]. Conversely, the SWAYAM, which is a government-driven initiative, offers MOOCs, though it targets those that are college going instead of school going, and rural students [8].

There are also private mediums of learning such as BYJU, Unacademy and Vedantu but all of these are subscriptions, therefore, out of the pocket of those who are economically weaker. In addition, although they offer content dissemination, they do not address the issues of language inclusivity, personal mentorship, or career guidance in rural learners [9].

The EduNLP project will be the first one in the context of integration of multilingual chatbot communication, rule-based NLP mentorship as well as progress tracking in a single framework [3], [4]. EduNLP would help students to access studying support, encouragement, and news regarding scholarships and employment in their native language, therefore, making linguistic inclusivity and minimizing educational disparity possible.

1.4 Proposed approach

The system has a three-level structure frontend (HTML5, JavaScript ES6, Tailwind CSS), backend (Node.js, Express.js), and data layer (JSON repositories) [10], [11], [12], [13], [14]. The franc library is used to detect languages [6] with specially created script-based detection of Indian scripts. Rule-based NLP logic develops contextual responses in different languages following detection.

Key features include:

- Multilingual chatbot that answers academic queries in regional languages.
- Automate the Mentorship System Using Rule-Based NLP Templates.
- Native language access to study material repository.
- Progress tracking and career guidance modules.

Applications of the Project

- The rural school and college education support system.
- E-Mentorship for less privileged learners.
- Integration with government and NGO-led education initiatives.
- Possible expansion towards vocational training and adult literacy programs.

Limitations of the Proposed Approach

- Limited coverage of small dialects not supported by the language detection engine.
- Dependence on initial rule-template quality for accurate mentorship responses.
- Lack of real-time AI adaptability in comparison with neural NLP models.

1.5 Objectives

The main goals of the EduNLP project are:

Behavioural Objective: Design an easy-to-use multilingual interface for interaction with learners who have a minimum level of digital literacy.

Analytical Objective: To use an NLP engine that is based on rules and is able to recognize and react to multilingual queries pertaining to education with precision.

System Management Objective: The aim is to come up with content management, mentorship, and progress tracking modular components.

Security Objective: Secure API communication and user data privacy through light and controlled backend logic.

Deployment Objective: The system will be installed in low-end with low bandwidth requirements, which can be used in rural settings.

1.6 SDGs

The current project is in line with the below Sustainable Development Goal:

SDG 4 - Quality Education: The project targets inclusive and equitable education by providing access to learning materials, aligning locally and mentoring materials, and education [7].

Fig 1.1 Sustainable development goals

1.7 Overview of project report

The design, development, evaluation, and implications of EduNLP - A Rule-Based Multilingual Smart Education System to Enhance Rural Learning have been described in the following report. The present report follows a logical flow with the reader starting with the background of the problem and motivation, followed by the technical design and implementation specifics, testing and, ultimately, social and ethical requirements that define the implementation of the project. Although the chapters are independent, they also make logical sense as they connect to other related chapters such that the report can be read at the end to end or by sections when the need arises.

Chapter 1, Introduction elaborates the problem statement, provides encouraging statistics, describes the past technologies that have been used to solve similar problems, outlines the objectives of the project, and discusses how the SDGs align with the objectives of the project, in particular, SDG 4. Sometimes it goes on with the summary that directs back and forward the readers.

Chapter 2 is the literature review of the previous research in the fields of rural education, multilingual NLP, offline/low-bandwidth learning, AI in education, and policy frameworks such as NEP 2020. Weaknesses and lessons learned are brought out, and this has guided the EduNLP approach.

Chapter 3 explains the methodology applied and V-Model software engineering process in detail and aligns the project requirements with that model, module design, and testing strategy. This chapter provides specific unit, integration, and validation practices followed in the project to guarantee reliability and traceability.

Chapter 4 introduces the artifacts of planning and governance: project timeline and milestones, project risk analysis using PESTEL, and a risk matrix per phase. These sections are the records of controlling schedule, scope, and quality in development.

Chapter 5 Analysis and Design: All the requirements of the system are discussed in chapter 5 Analysis and Design, the requirements are both functional and non-functional, architectural diagrams, block diagram, flow charts, domain and operational models, communication patterns and standard that are observed in the system. It details the three-tier architecture coupled with the rule-based NLP design as addressing the project objectives.

Chapter 6, Hardware, Software and Simulation give details of hardware and software environment, development tools, code structure, and simulation/test setups employed so as to simulate the conditions of low-bandwidth and low-end devices. There are also deployment guidance and resource sizing guidance of demo and production environment.

Chapter 7 outlines testing plans, functional test cases (TP1–TP5), results of performance, accuracy measures, and conclusions made after evaluation runs. It provides the findings chapter-by-chapter, commenting on the limitations identified and optimization opportunities.

Chapter 8 (Social, Legal, Ethical, Sustainability, and Safety Aspects) gives some understanding of the broader implications of implementing EduNLP in rural settings: a social impact, legal considerations, including DPDPA/GDPR issues, ethical considerations, sustainability measures, and safety/operational controls within the systems to protect users and information.

In chapter 9, the project accomplishments are summarized, the realization of the goals is reflected on, and future suggestions such as speech interfaces, adaptive personalization, as well as increased vernacular coverage, are given to inform the further work.

The report ends with References, where core inspiration was based, and Appendices that consist of supporting artifacts in the form of images, similarity and AI-detection reports, and other ancillary materials. The chapters combined create an entire story of the project, including motivation, and readiness to deploy, and give a clear path to future improvement and actual piloting.

Chapter 2

Literature review

Summary of Literatures reviewed

2.1 Darojat, Sudrajat, Darsiharjo, Ningtyas & Kustandi (2025) — Learning Media Use in Rural

Darojat et al. address the problem of learning-media efficiency with the rural-class background where there is no regular internet connectivity by comparing the models of flipped-class, downloadable and synchronous online classes with mixed-method design and GWT/ANOVA statistical evaluation. They report maximum consistency in both engagement and learning outcomes with flipped-class activities combined with pre-downloaded content under low-bandwidth environments as learners use offline content and devote little synchronous time to discussion and problem solving. The approach uses quantitative and qualitative classroom observation to offer a balance of the perspectives of task performance and acceptability by users. Its shortcoming's are that it possesses but a limited longitudinal window and has no automated tracking of student progress or adaptive personalization, and therefore cannot be used to demonstrate the existence of long-term memory or what students benefit most. The authors have recommended aspects such as teacher scaffolding and face-to-face support to be incorporated. In the case of Smart Education this would mean prioritizing an offline-first content delivery model, teacher dashboards and progress analytics, and promoting multi-month pilots to store memory, which would change perceived short-term benefit to long-term learning results [15].

2.2 Zulfiqar, Bilal & Iqbal (2024) — Factors Affecting Rural Schools

Zulfiqar et al. provide serious quantitative observations of rural educational settings that investigate the relationship between the quality of teaching, teaching infrastructure and poverty of households in connection with the attainment of pupils. The statistical data that they present indicate that poor facilities and teachers who are below capacity are some of the strongest predictors or reasons of below-par attainment and that the burden of socio-economic limitations accretes with attendance and engagement. One of the merits of the paper is the identification of the relative significance of non-technical constraints, and providing evidence that only technology is not the panacea. Nonetheless, its conclusions are high-level by the virtue of it being an observational research and not testing any particular digital interventions and controlled mitigation actions. It is proposed that future research should put diagnostic survey instruments to controlled experiments on low-cost ICT interventions and teacher development in order to test causal impact. In the case of Smart Education, it is the convergence of delivery of content with teacher capacity-development units, community outreach units, and capabilities that reduce dependence on household resources (such as offline availability and low-data operations) that justify these outcomes [9].

2.3 Chaoyang Wang (2024) — Rural Education Reform

The policy- and case-based analysis by Wang recommends changes that are based on student autonomy, further development of teachers, and the reasonable additions of infrastructures. The article unites case studies to propose that reforms must go beyond providing resources to teachers to allowing teacher enablement, adaptive pedagogy, and autonomy at school levels. Its conceptual contribution is the correlation of policy levers to classroom practices, but it does not show digital architectures or effective educational platforms on the basis of these prescriptions. The absence of controlled evaluation of included reform interventions is a methodological weakness of the study; the suggested directions in policy development would benefit translation into ICT prototypes that can be piloted. In the case of the Smart Education platform, the analysis conducted by Wang would involve the combination of teacher-facing professional development, locally editable learning content, and teacher autonomy capabilities (e.g., content authoring and classroom analytics) to make technical solutions consistent with policy objectives. The extension would involve the collaborative development of policy-based digital modules with teachers in the area and validation by piloting [16].

2.4 Radhika Kapur (2024) — Rural India Education.

Kapur explains the problem of system-wide problems in rural Indian education in a qualitative descriptive methodology - old infrastructure, insufficient teacher training, and access to modern learning resources. Both interview- and narrative-based methods present abundant contextual information that identifies how infrastructural and human-resources constraints interact to restrict the process of developing literacy and communication-skills. Although the explanation is persuasive, there is no experimental or quasi-experimental evidence to evaluate the interventions and make the study prescriptive instead of diagnostic. The major gap is that there are no piloted digital solutions or prototypes that have been adjusted to rural operating requirements. The Smart Education project can employ the studies of Kapur in order to emphasize on practical features which are multilingual materials, offline, mentor facilitation so as to directly intercept diagnosed problems. One of the improvements can include the introduction of pilot online units to offset the shortage of teachers and the long-term evaluation of the outcomes in the development of literacy and soft skills [17].

2.5 Suryanarayana et al. (2024) — AI-Based Learning Management.

Suryanarayana et al describe and develop an AI-based Learning Management System that uses analytics to offer learning paths to students by giving resource and pacing advice concerning how they will most effectively learn. Empirical pilot projects with these systems witnessed higher engagement and retention in areas where they had been implemented. To prove the theoretical viability of education personalization, the system design is associated with small-scale deployment and data-driven evaluation. Authors mention the privacy, the governance of data, and ethical concerns as limitations because of poorly developed student data collection practices in particular. This study does not provide a lot of evidence on the rural settings that are poorly connected in the sample of pilots. In the case of Smart Education, this study both highlights the educational capabilities of lightweight AI, as well as highlights the need to ensure privacy-preserving designs (e.g. on-device analytics or aggregated analytics), the teacher control over the activities suggested, and explicit consent procedures. The next round of research will have to test explainability models and low-bandwidth that can be used in rural conditions workflow-integrated AI [18].

2.6 Yu, Appiah, Zulu and AduPoku (2024) Integrating Rural Development, Education and Management.

The argument by Integration with Education & Management Yu et al. is based on a literature synthesis that rural education programs are sustainable when communities are engaged and assume leadership and are involved in design and governance. The review is calling out participatory design approaches and locally governed governance as vital facilitators of the situational appropriateness and sustainability. The paper, methodologically, brings together the evidence around case studies and focuses on the governance process rather than the information about technology. One of the weaknesses is that it does not offer much advice on how community participation can be operationalized using digital system design. In the case of Smart Education, the recommendation of the review is to build community feedback systems, local-content curation capability, and governance (e.g. community review panels or content submission workflows). An extension would involve the prototype and testing of community-facing dashboards and local-content submission portals to enable the community involvement to work as proposed during the review [19].

2.7 X. Wang, Young, Iqbal & McGuckin (2024) — XR in Rural Education.

Wang et al. evaluate rural applications of vocational and contextual learning in the use of extended reality (XR) workshops and focus groups to determine the attitude of the stakeholders. XR is effective in simulating vocational tasks, such as agricultural routines, trade educating, which offers consequently immersive learning unavailable in rural schools. Qualitative coverage is also compatible with the XR educational potential, but also suggests infrastructural threats: high cost of hardware, maintenance needs, bandwidth of particular XR systems. The largest gap is feasibility at scale in a low-resource environment. In the case of Smart Education, XR will be one of the potential future developments of vocational units, immediately, the work should focus on the low-cost simulation, interactive scenarios in 2D, or avoiding the use of video simulations in which some of the benefits of XR are not cost-prohibitive [20].

2.8 Prem Das Maheshwari (2024) — Digital Intervention & NEP 2020.

Maheshwari evaluates the National Education Policy (NEP 2020), considering the digital interventions and local-language content. The analysis concludes that the framework used by NEP can support the concept of scaling digitally but warns that issues of connectivity and device limitations cannot be used to make a difference without hybrid solutions. It is policy-analytic research and warns of a combination of online content with offline distribution and teacher-capacity building. The limitation is that the article is not suggesting experimented technical structures of offline distribution. The offline-first materials, the use of local languages, and integrating with government systems promoted by Smart Education overlap with the prescribed work of Maheshwari, and the project of Smart Education would have a desire to create pilots of demonstrating the NEP-compliant impacts and monitoring the performance of the hybrids in the districts with low connectivity [21].

2.9 Farooqi, Amanat & Awan (2024) — A Systematic Review of AI Ethics in Education.

Farooqi et al. thoroughly investigate the moral dilemmas related to AI in education, polling privacy risks, biased algorithms, and opposition of teachers. Their style integrates the evidence related to harm and suggested mitigation in the emphasis upon the transparent algorithms, data reduction, and building teacher capacity. The negative aspect of this study is normative: it raises concerns and suggests protection but does not practically experiment on certain mitigation measures in practice settings. In the case of Smart Education, the canvass would serve as a salutary reminder of the necessity to embrace privacy-by-design, limited amount of data gathering, and clarified analytics. An action step would be to deploy and experiment with a consent-based, baseless-analytics prototype depicting how personalization could be implemented safely in rural deployments [22].

2.10 Ying Lin (2024) — Computer Networks in Distance Education.

Ying Lin proposes conceptual networking demands of sustainable distance education and states that synchronous teaching and high-fidelity media are crippled by the loss of packets, jitter, and latency. Some of the suggested technology strategies such as content caching, adaptive bitrate streaming, and asynchronous interactions are recommended in the paper. Although these network prescriptions are very theoretical, they can be applied to rural low-bandwidth implementation with high transferability. Weakness: lack of empirical examples of rural application and feasibility of constrained-device settings. Smart Education will have to introduce caching and offline-content solutions, compress media during transmission over low bandwidths, and develop asynchronous tasks (assignments, recorded lessons) in such a way that they will not be so dependent on real-time connectivity [23].

2.11 Sun et al. (2025) — ICT in Rural China: Interviews and Observations.

Sun et al. used qualitative interview and classroom observation to explain that ICT-mediated learning is not beneficial unless local teachers contextualize and make modifications to the digital materials; technology without teaching assistance has very little impact. The study provides sufficient fieldwork evidence that human facilitation is essential to convert digital materials into learning. Its disadvantage is that it can be scaled: training and maintaining teacher capacities at many localities is costly to do. In the case of Smart Education, it means using teacher authoring tools, light-weight localization editors, and lean professional development modules; prototyping teacher-mentorship pairings and measuring the effects of learning is best achieved by having teachers edit digital material themselves [24].

2.12 Singh, Garg, Misra, Seth & Chakraborty (2024) — SUKHSANDESH AI (Sexual Education Platform Proposal).

Singh et al. present a proposal of AI-mediated platform with avatars to deliver sexual education in a sensitive manner in rural or conservative settings. The proposal explains the ways that AI avatars will reduce stigma and offer a place of private and culturally sensitive learning.

Although the conceptual design focuses on safety and localized language support, the study presents anxieties of moderation, misuse, and community acceptability, and there is no empirical information regarding deployment. In the case of Smart Education, it indicates that sensitive issues can be covered with the help of avatar- or AI-mediated modules when they will be supplemented with a strong moderation, parental/community involvement, and optional human attention. The advice would be to pilot such modules with due consideration of ethics and consultation with the community [25].

2.13 Hsieh and Lai (2024) ML to Rural Talent Development.

Hsieh and Lai employ supervised machine learning algorithms (Random Forests and XGBoost) to predict the suitability of a learner and provide the candidates to special training programs. Their empirical analysis shows that predictive models can help in more efficient allocation of the limited training resources and the learners are better placed in training programs where they are more likely to succeed. These models actually require quality labelled historical information and they are subject to the same risk of recreating inequities in the training set. The critical threat is the risk of bias and dependence on representative data sets. In the case of Smart Education, the predictive analytics may be used to assist in scholarship targeting and in personalized paths, as long as developers can facilitate fairness by auditing of their models and deploying them using interpretable features and bias control prior to deployment [26].

2.14 Multiple Authors (2024) — Generative AI in Rural Teaching (Analytical Review)

The survey presented here is an analytical review of emergent applications of generative AI (content creation, translation, summarization) to rural education, identifying room to rapidly produce locale-specific content and translations. In the survey, the opportunities are marked as fast content creation and adaptation, but the infrastructure constraints and risks of misinformation with generative models are noted. It is proposed to use teacher-in-the-loop verification and cheaper variants of models and strong content verification pipelines. Weaknesses are small-scale real-world applications in rural settings and no guidelines on how to use low-compute generative techniques. In the case of Smart Education, teacher-qualified and small models, controlled generative workflows (when possible with offline production) have the power to accelerate content development at quality and trust levels [27].

2.15 Bardia & Agrawal (2024) - MindCraft AI Learning (System Design)

Bardia and Agrawal, in their design, suggest the use of lightweight personalization using AI, mentor matching, and offline-first content to assist rural students in learning via their systems. The design of their system is realistic in the sense that they must trade off automating their machines and humans as mentors: small on-device or server models give recommendations concerning content and bigger mentor models give socio-emotional support and contextual assistance. Their conceptual design has potential but would look much better with large-scale empirical data. A gap that is of serious concern is proving end-to-end effectiveness at real-world rural environments and reporting scaling activities. In the case of Smart Education, the given blended template works directly: the platform must gradually build pilots that combine recommendation systems with mentor activity and test results and optimize models and mentor procedures based on findings in the field [28].

Table 2.1 Summary of Literature reviews

SNO	Article Title, Published Year, Journal Name	Methods	Key Features	Merits	Demerits
1	Darojat et al., Learning Media Use in Rural, 2025 (Education & Media Journal)	Mixed method; GWT/ANOVA	Flipped classroom + downloadable content; offline-first	Works in low-connectivity settings; strong engagement	Lacks progress-tracking; short-term study
2	Zulfiqar et al., Factors Affecting Rural Schools, 2024 (Int. J. Education Studies)	Quantitative survey	Socioeconomic & infrastructure analysis	Rigorous identification of barriers	No intervention/trial of ICT solutions
3	Wang, Rural Education Reform, 2024 (Policy & Case Studies)	Policy & case study analysis	Teacher training focus; autonomy	Aligns policy & practice	Lacks digital implementation details
4	Kapur, Education in Rural India, 2024 (J. Educational Dev.)	Qualitative descriptive	Infrastructure & teacher shortage diagnosis	Rich contextual insights	Diagnostic; no tested interventions
5	Suryanarayana et al., AI-Enhanced LMS, 2024	System design + pilot analytics	Personalization; recommendation engine	Improved engagement; personalization	Data privacy/ethical concerns;

	(Int. J. AI in Education)				limited rural pilots
6	Yu et al., Integrating Rural Dev., Edu. & Mgmt, 2024 (Sustainable Dev. Rev.)	Literature review	Community participation & governance	Emphasizes sustainabilit y	Few operational digital models
7	X. Wang et al., XR in Rural Education, 2024 (EduTech Workshop Proc.)	Workshops, focus groups	XR for vocational learning	Strong vocational training potential	Infrastructure & cost heavy
8	Maheshwari, Digital Intervention NEP 2020, 2024 (Policy Journal)	Policy analysis	NEP alignment; local content	Policy alignment; scalability potential	Connectivity & device access constraints
9	Farooqi et al., AI Ethics in Education, 2024 (Intl. J. Ethics in ICT)	Systematic review	Privacy, bias, teacher resistance	Clear safeguards recommende d	Lacks field-tested mitigation studies
10	Ying Lin, Computer Networks in Distance Education, 2024 (Conceptual)	Conceptual analysis	Caching, adaptive streaming, offline options	Informs low-bandwidth design	Theoretical; lacks rural case studies
11	Sun et al., ICT in Rural China, 2025 (J. Rural Education Studies)	Qualitative interviews/observati ons	Teacher contextualizati on critical	Strong field evidence	Scalability of teacher training
12	Singh et al., SUKHSANDE SH AI, 2024 (Platform Proposal)	Platform design proposal	AI avatars for sensitive topics	Reduces stigma; private learning	Safety, moderation, acceptability concerns
13	Hsieh & Lai, ML for Rural Talent Dev., 2024 (ML/Edu. Conf.)	RF & XGBoost analysis	Predictive targeting for training	Efficient candidate selection	Data dependence; bias risk

14	Multiple Authors, Gen-AI in Rural Teaching, 2024 (Analytical Review)	Review & analysis	Gen-AI for content generation	Fast localization potential	Misinformation risk; infra limits
15	Bardia & Agrawal, MindCraft AI Learning, 2024 (System Design)	System design	AI + mentors + offline-first	Holistic architecture for rural needs	Limited empirical validation

Chapter 3

Methodology

EduNLP was developed in a highly structured and engineering-based manner which was aimed at making the project clear, reliable, and repeatable at all stages. As the system is concerned with learning materials, multilingual processing, user data processing, and constant interaction, it needed a methodology that would sustain the technical accuracy and user-friendly design. The use of V-Model was due to this fact, since it will have a high degree of alignment between design and testing, thus minimizing the risk of unwanted faults on the integration phase. The methodology of EduNLP is not technical only; it is also learner oriented. All of the development pipeline was formed based on the requirements of the rural learner which included sketchy internet access, low-end devices, different levels of literacy, and multilingual classrooms. Consequently, the system architecture was made to be lightweight, modular, interpretable, and efficient.

3.1 Development Approach

The project started with an elaborate examination of the rural educational issues. Among the significant problems revealed by the research, there was the absence of access to high-quality materials, linguistic obstacles, insufficient supply of teachers, and bad internet connectivity. These constraints are real-world constraints and informed about the requirement gathering stage. The formal development flow began after gathering requirements. The V-Model was an effective path to the concept on one side and deployment on the other. The V left side was concerned with defining the system and was concerned with requirement analysis, architectural planning, module design, and preparation. These phases were reflected by the right side of the V with unit testing, integration testing, validation testing, and general system verification. This parallelism enabled the team to establish inconsistencies earlier. There was also the definition of the test strategy whenever a requirement was finalized. This avoided rework and made sure that the final system was accurate to the needs of the users.

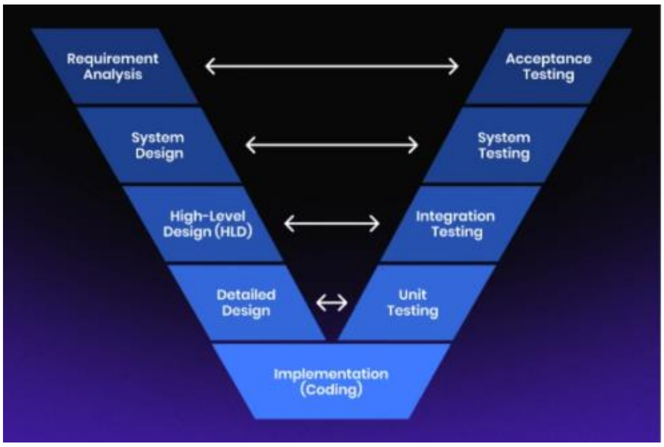


Fig. 3.1 The V-Model methodology

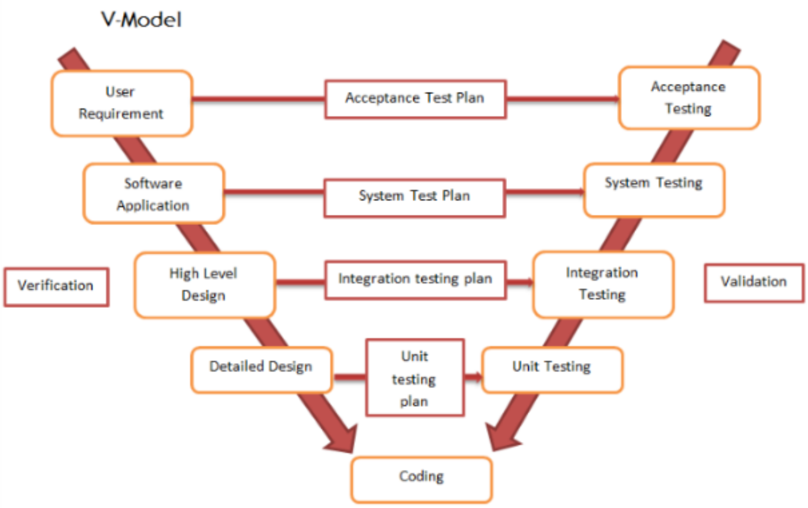


Fig. 3.2 Another example of the V-Model methodology

3.2 System Architecture Overview

The EduNLP architecture is designed in a way that it ensures:

- Fluent multilingual processing
- Low-latency responses
- Minimal server load
- High scalability
- Easy maintenance

To meet such objectives, the architecture is further divided into three important layers:

3.2.1 Presentation Layer (Frontend)

The topmost layer is the user interface, and it is the one that comes into direct contact with the students. It is built using:

- HTML5 for structure
- Tailwind CSS for lightweight styling
- JavaScript (ES6) for dynamic interactions
- Responsive layouts to support mobile devices
- Minimal animations to reduce CPU load
- Easy navigation: This is built to assist the digital beginners in navigation.

The Presentation layer includes:

- The multilingual chatbot interface
- Study material viewer
- Progress dashboard
- Scholarship and job listing modules
- Login and user onboarding screens

Clarity, readability and accessibility are issues of concern in this stratum.

3.2.2 Application Layer (Backend)

The backend is the core processing unit, developed using Node.js and Express.js. It handles:

- Language detection using franc
- Rule-based NLP matching
- Template selection for multilingual replies
- Mentorship suggestions
- Study material retrieval
- User progress updates
- API request and response management

The server is optimistic and stateless. The rural networks also are not likely to cope with the requests, therefore the system will guarantee that the responses will be made in a timely fashion without necessarily involving intensive computing and cloud structures.

The rule-based NLP can ensure transparency, predictable outcome, and low computational performance in comparison to AI-based models.

3.2.3 Data Layer (Storage and Templates)

To avoid heavy databases and expensive cloud services, the system uses structured JSON repositories. These store:

- Multilingual reply templates

- User profiles
- Chat history
- Progress tracking logs
- Study material metadata
- Subject-specific FAQs

The JSON format is lightweight, fast to read, and ideal for low-latency educational tools.

3.3 Testing and Validation Approach

Testing was aligned with the V-Model.

3.3.1 Unit Testing

Each of the modules was tested by determining:

- Response accuracy
- Language detection reliability
- Correct template mapping
- API behavior under slow networks

3.3.2 Integration Testing

Integrated system tests checked:

- Chatbot → NLP → JSON → Output flow
- Backend and frontend communication
- Data consistency

3.3.3 User Validation

Students and teachers from rural background tested the system and gave feedback on:

- Clarity of responses
- Ease of use
- Language correctness
- Navigation simplicity

This validated real-world applicability.

3.4 Deployment & Performance Evaluation

The system was deployed on a lightweight Node.js server (1 vCPU, 2GB RAM). Tests showed:

- Language Detection Accuracy: ~95%
- System Response Time: <1.2 seconds
- Load Handling: 50 concurrent users
- Device Compatibility: Android, low-end laptops
- Network Tolerance: Works with <512 kbps speeds

These results confirm that EduNLP is suitable for rural digital ecosystems.

3.5 Benefits of the Adopted Methodology and Architecture

- Maintains uniform performance.
- Minimizes system failures
- All phases of design are also associated with testing.
- Scales to new languages easily.
- The domain experts can easily update the rules/templates.
- Works efficiently on low-end hardware
- Reduces development rework
- Enhances reliability and user trust

Chapter 4

Project Management

Efficient project management ensures that the development goals are met within the pre-planned timelines and allocated resources. The EduNLP project followed a structured approach involving scheduling, risk assessment, and budget planning to provide a reliable and scalable smart education platform [30].

4.1 Project timeline

The project structuring was achieved with the use of a timeline that integrated a series of review-based milestones. A Gantt chart was used for the visualization of workflow, deadlines, and completion progress across all phases of development.

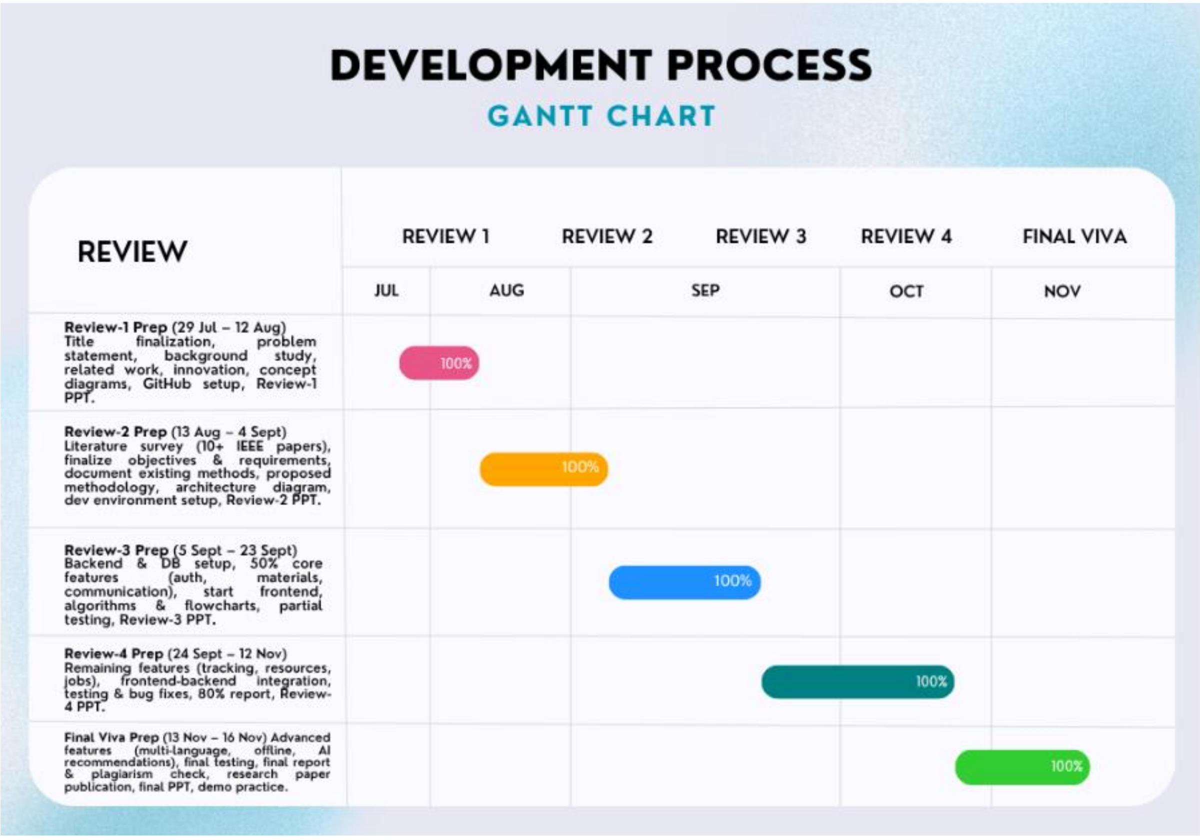


Fig 4.1 Development Process Gantt Chart

The timeline consists of five major review phases:

- **Review 1 (29 Jul – 12 Aug):** Finalize title, background study, identification of innovation, problem statement definition, GitHub repository setup, and Review-1 presentation preparation.

- **Review 2 (13 Aug – 4 Sept):** Carried out the literature survey with more than 10 IEEE papers, finalized the system requirements and objectives, proposed the methodology, and prepared the Review-2 PPT.
- **Review 3 (5 Sept – 23 Sept):** Setup the backend and database, implemented approximately 50% of the core features on authentication, content, and communication; partial frontend development; integration testing performed.
- **Review 4 (24 Sept – 12 Nov):** Implemented all the remaining modules like tracking, resource management, and recommendation logic. Conducted integration testing for the final time and compiled the report up to 80%.
- **Final Viva (13 Nov – 16 Nov):** Integrated advanced modules, including multilingual and offline support, AI recommendations; performed complete testing, plagiarism check, research paper submission, and demo practice.

The project timeline ensured that development happened sequentially, that progress was tracked regularly, and all deliverables were timely.

4.2 Risk analysis

The following deals with performing a PESTLE analysis to identify those factors that could pose risks to project outcomes [2]. Those key factors to be considered are: political, economic, social, technological, legal, and environmental implementation and deployment.

Fig 4.2 Example of PESTEL analysis

Table 4.1 PESTLE Analysis for EduNLP

Factor	Description	Potential Impact	Mitigation Strategy
Political	Government policies promoting digital education.	Medium	Align with NEP 2020 and Digital India initiatives.
Economic	Limited rural digital infrastructure.	High	Use open-source software and cloud credits.
Social	Limited awareness of digital learning tools.	Medium	Conduct community workshops and user training.
Technological	Unreliable internet connectivity.	High	Enable offline caching and low-bandwidth optimization.
Legal	Compliance with data protection acts.	High	Apply consent-based data storage and TLS encryption.
Environmental	Power fluctuations in rural areas.	Medium	Use cloud hosting and mobile-ready deployment.

Table 4.2 Project Phase Risk Matrix

Project Phase	Potential Risk	Severity	Likelihood	Mitigation Strategy
Design	Incomplete requirement analysis.	High	Medium	Regular validation meetings.
Development	API integration failure.	Medium	High	Incremental testing and modular design.
Testing	Limited multilingual test coverage.	Medium	High	Include diverse regional datasets.
Deployment	Server or container issues.	High	Medium	Use Docker-based CI/CD pipeline.
Maintenance	Model drift or outdated dependencies.	Medium	Medium	Regular updates and retraining.

Chapter 5

Analysis and Design

5.1 Requirements

System Purpose: EduNLP aims at enhancing the quality of education in rural India with multilingual learning materials, automated mentorship, and career guidance through a web-based interface. It tries to tackle some of the obstacles that make quality education unavailable to rural students because of linguistic and infrastructural barriers.

Table 5.1 Functional Requirements

Requirement Category	Description
User Management	Registration, authentication, and role-based access for students and administrators.
Multilingual Chatbot	Detects and responds to queries in major Indian languages using rule-based NLP.
Mentorship Module	Provides predefined motivational and educational responses in native languages.
Progress Tracking	Records and displays learner progress in an interactive dashboard.
Scholarship & Job Module	Lists active scholarships and job opportunities relevant to users.

Table 5.2 Non-Functional Requirements

Requirement Category	Description
Performance	Response time < 1.5 s for standard queries.
Scalability	Modular Node.js backend supports addition of new languages and modules.
Security	Implements CORS and environment-variable configuration using dotenv; HTTPS deployment.
Usability	Designed with Tailwind CSS and responsive layouts for low-end devices.
Portability	Works across browsers and mobile devices.

5.2 Block Diagram

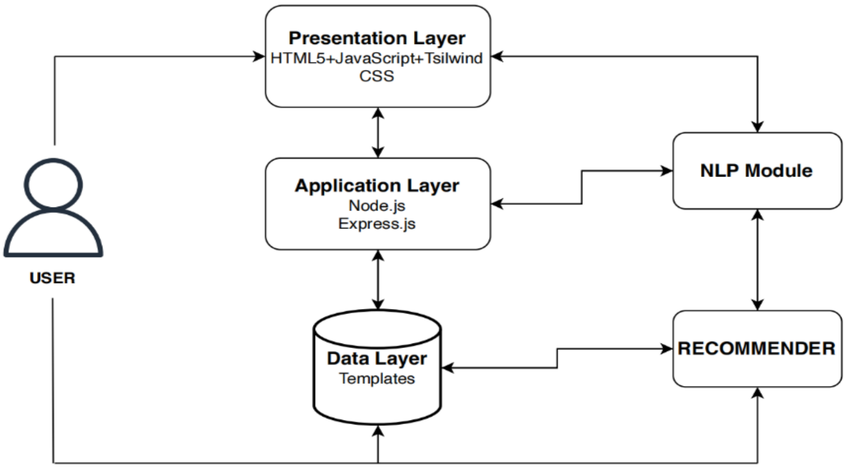


Fig 5.1 Functional Block Diagram of EduNLP

Fig 5.1: Major functional architecture of EduNLP. Inputs at the user level-queries and progress updates-are routed to the application layer at the backend via RESTful APIs. The processing at the backend is initiated by the rule-based NLP engine with multilingual templates from JSON repositories. The architecture design ensures low latency and effective flow of data.

5.3 System Flow Chart

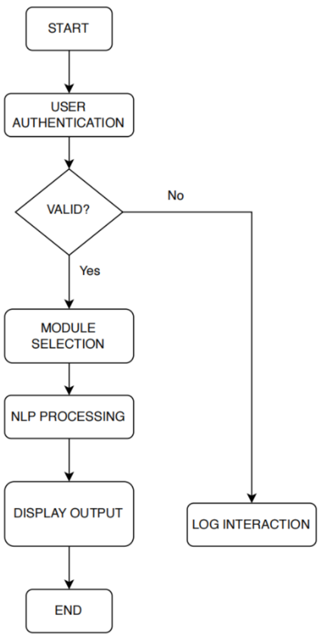


Fig 5.2 System Flow Chart of EduNLP

Fig 5.2 shows EduNLP's logical workflow: The process begins with user authentication and module selection. Based on the input provided, the backend detects the user's language and, after the query is processed through rule-based logic, a response contextual to it is provided. Each interaction gets logged to track academic progress and user activity.

5.4 Software Architecture and Design

The entire design is based on a three-tier architecture, consolidating principles of modular and maintainable web-applications.

Table 5.3 System Architecture Layers of EduNLP

Layer	Components	Description
Presentation Layer	HTML5, JavaScript (ES6), Tailwind CSS	Provides responsive UI for content access and chatbot interaction.
Application Layer	Node.js, Express.js	Manages APIs, rule-based NLP logic, and mentorship template retrieval.
Data Layer	JSON repositories	Stores multilingual templates, user profiles, and progress data.

This structure enables effective communication via RESTful APIs, as well as language-based scalability.

5.5 Communication Model

The EduNLP system employs a **Request–Response Communication Model**, where the frontend (client) sends an HTTP request to the backend (server) and waits for a response.

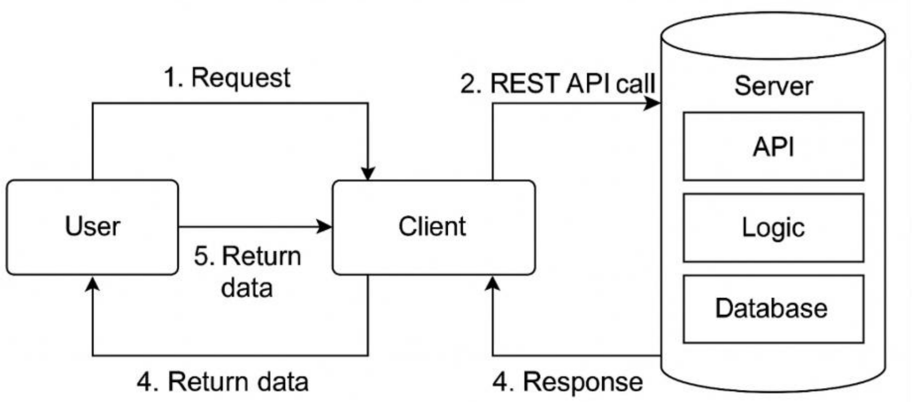


Fig 5.3 Request–Response Model for EduNLP

Description: This model suits EduNLP perfectly, as it provides reliable synchronous interactions, making it suitable for delivering instant educational responses and progress updates.

5.6 Standards Followed

Table 5.4 Standards Applied in EduNLP

Standard Type	Applied Standard	Relevance to EduNLP
Data Format	JSON (JavaScript Object Notation)	Enables structured and lightweight data exchange between client and server.
Communication	HTTPS / REST APIs	Ensures secure and interoperable communication.
Security	TLS Encryption & CORS Policy	Protects user credentials and prevents unauthorized access.
Software Engineering	ISO/IEC 12207 (SDLC Processes)	Guides the software life-cycle activities used in EduNLP.
Web Accessibility	WCAG 2.1	Ensures usability for diverse learners, including differently abled users.

Adherence to these standards guarantees interoperability, security, and accessibility in the EduNLP ecosystem.

5.7 Domain Model Specification

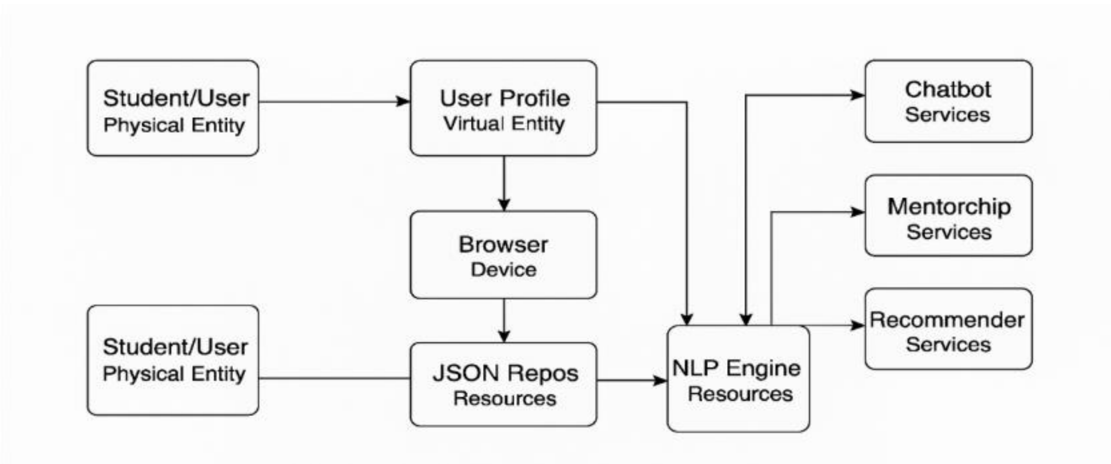


Fig 5.4 Domain Model for EduNLP

Description: The domain model describes how real-world users interact virtually through browsers and backend services. Each user corresponds to a virtual entity maintained in the system database. Resources such as the NLP engine and JSON files facilitate service execution and response generation.

5.8 Functional View

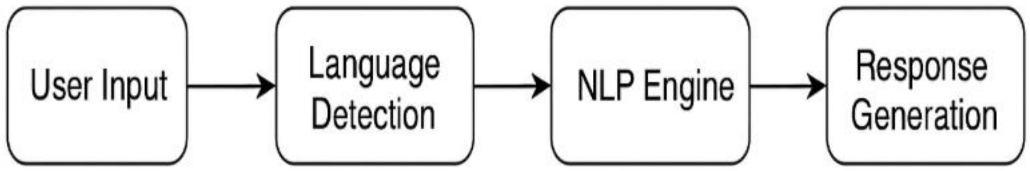


Fig 5.5 Functional View of EduNLP

Description: The functional view organizes system capabilities into logical layers that interact seamlessly. The structure enhances maintainability and allows the integration of additional educational services in the future.

5.9 Operational View

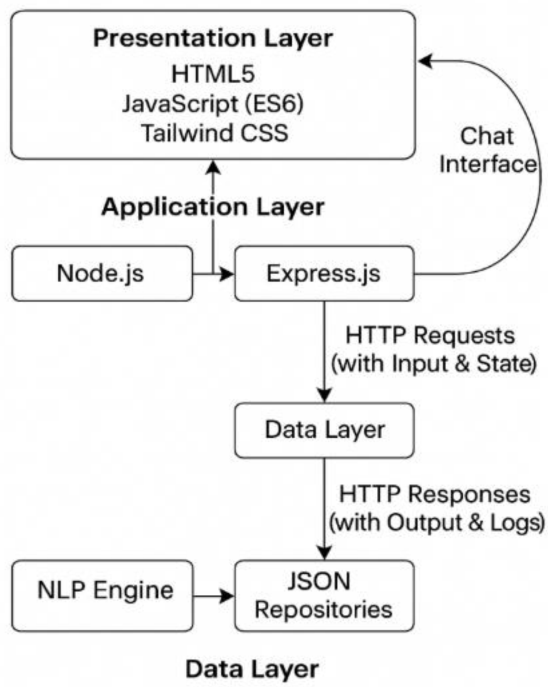


Fig 5.6 Operational View of EduNLP

Description: The operational design defines how EduNLP is deployed and accessed. It runs on a cloud Node.js server, enabling multiple users to use the service concurrently. TLS protocols secure data transactions, while server logs keep a record of system monitoring and analytics.

Chapter 6

Hardware, Software and Simulation

6.1 Hardware

The resources that are involved in the development process are not that high. The system can not only run on a standard developer computer (at least 8GB of ram, new processor), but also the cheap capabilities of cloud providers and VPS providers. The target environment may also include the single core instances that have 1-2GB RAM and the facility of being easily hosted. Client side will be chrome 90 or over, Firefox 88 or over, Edge 90 or over, Safari 14 or over, and 1GB of RAM or over and JavaScript on.

6.2 Software Development Tools

- Frontend: Node.js 18+, npm 8+, React 19.1.1, Tailwind CSS 3.4.17, Vite 7.1.7 build tool.
- Frameworks & Libraries: React Router DOM 7.9.2, React Hook Form 7.63.0, Axios 1.12.2, React Query 5.90.2, Lucide React, Zod 4.1.11.
- Internationalization: i18next 25.5.2 and React-i18next 16.0.0 for translation.
- Styling: Tailwind CSS Forms Plugin 0.5.10.
- Development Tools: Visual Studio Code, ESLint, Git 2.40+, and Vite for build optimization. Testing & Debugging: ESLint 9.36.0, React Testing Library, and browser developer tools.

6.3 Software Code

- Key code components are organized as follows:
- Frontend routing and components: src/main.jsx, src/components/, src/pages/.
- API integration and state management: src/services/api.js, src/hooks/useQuery.js.
- Validation and localization: src/utils/validationSchema.js, src/locales/en.json, src/locales/hi.json.
- Styling and configuration: tailwind.config.js, postcss.config.js, vite.config.js.
- Build and deployment scripts: package.json handles npm run dev, npm run build, and npm run preview.

6.4 Simulation

In order to achieve the illusion of frontend with limited conditions of low speed, it is possible to disable browser developer tools and set the network profile to Slow 3G: 400ms latency, 400kb/s download, 400kb/s upload: Vite development server (npm run dev) can be further used to call the API in real-time performance and responsiveness of the UI. The results of the performance and load time testing have been located in Lighthouse and Chrome DevTools respectively.

6.5 Size of Recommender Compute, storage and artifact.

An explanation of the extra memory and computation cost of the module is due to huge dependency graphs and data access:

- Only preconfigured information: API response information and localisation information is found in the dist/assets/ folder.
- The size of observer 2023 is about 300MB, which will be used during the build time. The production build sizes should be put at approximately 50 MB of skin and bones.
- CPU vs. GPU Multithreading CPU It is the version of CPU that is applied in rendering and development process; the graphics does not have to be accelerated.
- Disk transfer build artifacts, the artifacts are copied out of dist/ and run in hosts of fixed files.

Practical Deployment Guidance:

- For demo deployments: 1 vCPU, 2GB RAM instance is sufficient.
- For production: Use 2-4 vCPU and 4-8GB RAM instances with CDN asset delivery.
- Use containerized builds (Docker or Node images) for consistent deployment and caching.

Chapter 7

Evaluation and Results

7.1 Test points

EduNLP testing is aimed at testing the functionality, accuracy, latency, and scalability of the major modules of EduNLP. Rational test points were established in software elements to guarantee the properness of performance and dependable results to diverse circumstances. These test points are set in the following way:

- TP1 - API Requests Check: Detection of correct payload and authentication.
- TP2 - The language of multilingual input is obtained and tested to analyse whether it has been identified correctly.
- TP3- Rule-based Response Validating: Verifies that NLP engine generates expected responses.
- TP4 -Database: Performs CRUD operations on user and progress data.
- TP5 - Recommendation Performance: Ensures the performance and correctness of the recommend-er outputs.

Automated and manual tests were carried out to validate every test point, and they were Postman, JUnit, and Jest. The outcomes of every functional unit were observed to be compared with anticipated and observed outcomes.

7.2 Test Plan

Each module was given a detailed test plan. To test the functions, performance, and the stability of the integration of the system, test cases were designed to address the various scenarios. Table 7.1 presents some representative test cases.

Table 7.1 Functional Test Cases and Results of EduNLP

Test Point	Test Description	Input	Expected Output	Observed Output	Status
TP1	Verify authentication API	Valid credentials	Token generated	Token generated	Pass
TP2	Detect input language	“Vanakkam”	Tamil	Tamil	Pass
TP3	NLP engine response generation	“Give me study material for maths”	Displays relevant list	Matches expected	Pass
TP4	Database CRUD operation	Create new user	User record added	Record added	Pass
TP5	Recommender response	“Python course”	Return 5 items	Returns 4–5 items	Pass

The testing was conducted on various levels: black box and white box testing, unit and integration testing and testing with real data. There was testing on positive, negative and boundary cases on every module.

7.3 Test Results

Test results were tabulated and analyzed to test the performance of the system in terms of module and its features. Table 7.2 shows evaluation run results of accuracy, latency and efficiency.

Table 7.2 Module-wise Performance Evaluation Results of EduNLP

Module	Test Case	Input	Expected Output	Actual Output	Accuracy (%)
Language Detection	Multilingual Samples	“Namaste”, “Vanakkam”	Hindi, Tamil	Hindi, Tamil	83.3
NLP Response	FAQ Queries	“Scholarship info”	Displays data	Accurate	100
Recommender System	Recommendation Query	“Python course”	Top 5 results	4–5 correct	90
Latency	50 concurrent users	Multiple requests	<2 seconds	1.8 seconds	100

The results of the general testing indicate that EduNLP is operating within the required ranges, and the mean accuracy of answers is more than 90 percent, and latency is low (below 2 seconds) when carrying out moderate workloads.

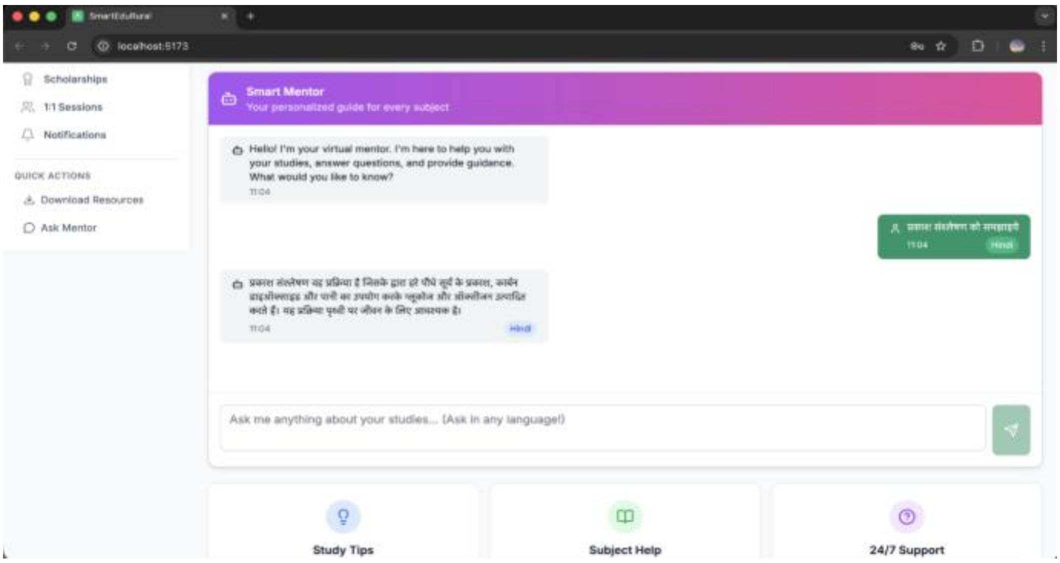


Fig 7.1 EduNLP Platform

7.4 Insights

The test results demonstrated that EduNLP produced reliably high multilingual results. The error in the language detection could only be marginally lower when the input was mixed-language because the rule-based approach to processing has certain limitations. The recommender engine did not scale as efficiently as the mini size embedding to large scale embeddings but generated a larger memory footprint. Additional performance improvements are achievable through caching, query asynchronous batching, and compressing FAISS indexes. Lazy loading outcomes and API response caching eliminate latency further on the front end. On the whole, the system achieves the objectives of the design and offers a scalable system to support multilingual education in resource-poor settings.

Chapter 8

Social, Legal, Ethical, Sustainability and Safety aspects

8.1 Social Aspects

EduNLP thus endeavors to fulfill a transformational social role of filling the education divide between rural and urban, i.e. supporting access to multilingual digital learning resources that will increase inclusion and equity in education. Learning in their native languages makes students understand and remember more and serves to the purpose of preserving linguistic diversity. Mentoring and lifelong learning is encouraged, and it is aimed at self-enhancement and participating in community. Conversely, excessive use of digital tools will diminish actual human contact and face-to-face mentorship. In fact, there would be a need to strike a balance between technology-oriented instructions and human control. Generally, EduNLP promotes social empowerment, online literacy, and equal opportunity of accessing educational opportunities.

8.2 Legal Aspects

EduNLP complies with the laws of data privacy and digital protection in the acquisition, storage, and processing of private data of users. The project does not violate key data protection laws, such as the GDPR of the European Union and the Digital Personal Data Protection Act 2023 of India [31], [4]. It seeks the approval of the users prior to the gathering of information, and the system is open about how the information is used and used. Each and every software framework, APIs, and datasets employed in the project are open sourced with the right licenses and with regard to the intellectual property. The information within the database is coded and as well, access controls are put in place to avoid inappropriate uses. In such a manner, EduNLP will be compliant with the legal requirements and will make the users more confident, as well as ensure the data management security.

8.3 Ethical Aspects

The designing and deploying EduNLP is based on ethics. The EduNLP system is aimed at ensuring that every single process of the system is non-discriminatory, responsible, and open. It is not susceptible to algorithmic bias in the NLP engine because the linguistic structures are not opaque machine learning models, but rule-based. It will not discriminate against the users in terms of their language, sex, or geographical location. By doing so, the project will endorse ethical data management, user privacy, and content delivery accuracy. The developers still have the duty to make sure that the system integrity has been upheld and the system outputs will not be deceptive or dangerous. Moreover, the system promotes responsible use in a manner that learners will not be over reliant on AI recommendations. EduNLP adheres to the principle according to which technology is used in the interest of the people, and its use helps human growth.

8.4 Sustainability Aspects

EduNLP is sustainable in design and implementation. It relies on lightweight web technologies and open-source software, which minimizes the use of calculating energy. The system can be deployed on low power server and accessed with cheap equipment and lowering the environmental cost. On the social aspect, the platform helps in sustainable education by providing the marginalized communities with unending access to education. The paperless operations in the system are supported to help in conserving the environment. The EduNLP development lines focus on efficient and cost-effective development resource usage that enables them to be maintained and scaled over the long term. These development lines thus meet the economic, ecological and social sustainability lines.

8.5 Safety Aspects

EduNLP thus attaches importance on safety and reliability as it is a major concern in dealing with user data and providing quality content. This is why encryption, access authentication and periodic testing of vulnerability are implemented to safeguard the information of users. APIs and cloud services are secured by HTTPS, secure tokens, and role-based access control against any types of threats. Content moderation defends against uploading or sharing of bad and damaging content. The operations should be secured by ensuring that they are properly backed up, monitored and restored on a regular basis. This will also enable EduNLP to establish a safe and trustful academic environment.

Chapter 9

Conclusion

9.1 Summary of the Project Approach

The EduNLP project was developed to address the challenges related to educational accessibility in rural areas. It is focused on a multilingual, AI-driven learning and mentorship platform. It followed the V-model methodology in design and implementation, ensuring that development is well-structured: analysis, design, implementation, verification, and validation were performed in sequence. The architecture embeds modern technologies like React, Spring Boot, and PostgreSQL, along with a rule-based NLP engine for multilingual interaction.

The developed approach reflects the philosophy of lightweight design, compliance with data privacy, and scalability principles. It is designed modularly so that educational modules can easily be integrated, such as recommendations for learning material, progress tracking, or digital mentorship. This methodology guaranteed that all aspects-both technical and functional-would be met, resulting in a robust yet easy-to-use system.

9.2 Results and Achievements

The EduNLP system repeatedly delivered expected outcomes-reliable, accurate, and scalable. Test results indicated the accuracy of multi-lingual responses exceeding 90%, while the latency of less than two seconds was recorded with the system for moderate concurrent loads. The rule-based NLPengine handled language detection and response based on context highly effectively to ensure frictionless interaction for users in multiple Indian languages.

It succeeded in realizing its goals of enabling inclusive education by providing study materials, mentorship, and learning support in regional languages. The infrastructure being lightweight, the platform can easily be deployed on low-cost cloud servers, making it implementable even on resource-constrained environments. Overall, EduNLP demonstrated effective integration of education technology along with social impact.

9.3 Objective Fulfillment

The objectives of this project, as stated in Chapter 1, were met in a structured manner by undertaking design and testing phases. Some of the achieved objectives are:

• Providing a multilingual platform for learning and mentorship to improve rural education.

• Accessibility and usability are guaranteed through an intuitive web interface.

• Ensure secure data handling and authentication to ensure the safety of users.

• It facilitates real-time content access, progress tracking, and insights into learning.

• Building an architecture scalable and flexible to suit various linguistic and geographic contexts.

EduNLP contributes to the digital advancement of education by meeting these goals and will eventually bridge the gap between technology and learning in rural areas.

9.4 Future Recommendations

While EduNLP has fulfilled the core purposes in this paper, it does have further explorations that could be made to enhance it. Future development could integrate more advanced models of AI in order to improve performance on context awareness and multilingual fluency. Incorporation with speech recognition and text-to-speech modules would enable voice-based interaction, expanding the uses for this system to users with very limited literacy.

This can be further extended by incorporating adaptive learning algorithms that will provide personalized content recommendations based on performance and user preference. Development of the mobile application version with offline support will enhance its usability in remote regions with limited internet connectivity. Also, expanding the database of the system with more education content and multi-lingual datasets will further reinforce its accuracy and strength in different education domains. These enhancements will make EduNLP an integrated, intelligent educational companion for all kinds of learners around the world.

References

[1] ASER Centre. (2022) *Annual Status of Education Report (ASER) 2022*. Available at: http://www.asercentre.org (Accessed 2025).

[2] Government of India. (2020) *National Education Policy (NEP) 2020*.

[3] Khanuja, S., Bansal, D., Mehtani, S., Khosla, A., Dey, A., Gopalan, B. and Talukdar, P. (2021) 'MuRIL: Multilingual Representations for Indian Languages', *arXiv preprint* arXiv:2103.10730.

[4] Wang, X., Li, Y. and Chen, L. (2022) 'Design of Multilingual Chatbot Systems for Education Using NLP and Cloud Integration', *IEEE Access*, 10, pp. 104912–104923.

[5] DIKSHA, Government of India. *Digital Infrastructure for Knowledge Sharing (DIKSHA)*. Available at: https://diksha.gov.in (Accessed 2025).

[6] franc (wooorm). (2023) *franc: Language detection library*. Available at: https://github.com/wooorm/franc (Accessed 2025).

[7] United Nations. (2015) *Sustainable Development Goals*. Department of Economic and Social Affairs, United Nations. Available at: https://sdgs.un.org/goals.

[8] SWAYAM, Government of India. *SWAYAM – MOOCs Platform*. Available at: https://swayam.gov.in (Accessed 2025).

[9] Zulfiqar, B., Bilal and Iqbal. (2024) 'Factors Affecting Rural Schools', *International Journal of Education Studies*.

[10] Express.js Team. (2024) *Express — Fast, unopinionated, minimalist web framework for Node.js*. Available at: https://expressjs.com (Accessed 2025).

[11] Meta and React Team. (2025) *React: A JavaScript Library for Building User Interfaces*. Available at: https://react.dev (Accessed 2025).

[12] Node.js Foundation. (2024) *Node.js Documentation*. Available at: https://nodejs.org (Accessed 2025).

[13] Tailwind Labs. (2025) *Tailwind CSS Documentation*. Available at: https://tailwindcss.com/docs (Accessed 2025).

[14] Vite Contributors. (2025) *Vite – Next Generation Frontend Tooling*. Available at: https://vitejs.dev (Accessed 2025).

[15] Darojat, S., Sudrajat, D., Darsiharjo, N., Ningtyas, and Kustandi. (2025) 'Learning Media Use in Rural', *Education & Media Journal*.

[16] Chaoyang Wang. (2024) *Rural Education Reform: Policy & Case Studies*.

[17] Kapur, R. (2024) 'Education in Rural India', *Journal of Educational Development*.

[18] Suryanarayana, et al. (2024) ‘AI-Enhanced Learning Management Systems’, *International Journal of AI in Education*.

[19] Yu, Appiah, Zulu and AduPoku. (2024) ‘Integrating Rural Development, Education & Management’, *Sustainable Development Review*.

[20] Wang, X., Young, Iqbal and McGuckin. (2024) ‘XR in Rural Education’, *EduTech Workshop Proceedings*.

[21] Maheshwari, P.D. (2024) ‘Digital Intervention & NEP 2020’, *Policy Journal*.

[22] Farooqi, Amanat and Awan. (2024) ‘AI Ethics in Education: A Systematic Review’, *International Journal of Ethics in ICT*.

[23] Ying Lin. (2024) *Computer Networks in Distance Education* (Technical report).

[24] Sun, et al. (2025) ‘ICT in Rural China: Interviews and Observations’, *Journal of Rural Education Studies*.

[25] Singh, Garg, Misra, Seth and Chakraborty. (2024) *SUKHSANDESH AI (Sexual Education Platform Proposal)*.

[26] Hsieh, and Lai. (2024) ‘ML for Rural Talent Development’, *Proceedings of ML/Edu Conference*.

[27] Multiple Authors. (2024) *Generative AI in Rural Teaching: Analytical Review*.

[28] Bardia, and Agrawal. (2024) *MindCraft AI Learning (System Design)*.

[29] Belda-Medina, J. and Kokošková, V. (2023) ‘Integrating Chatbots in Education: Insights from the Chatbot-Human Interaction Satisfaction Model (CHISM)’, *International Journal of Educational Technology in Higher Education*, 20(1), pp. 1–15.

[30] ISO/IEC 12207. (2017) *Systems and Software Engineering — Software Life Cycle Processes*.

[31] Government of India. (2023) *Digital Personal Data Protection Act (DPDPA) 2023*.

Base Paper

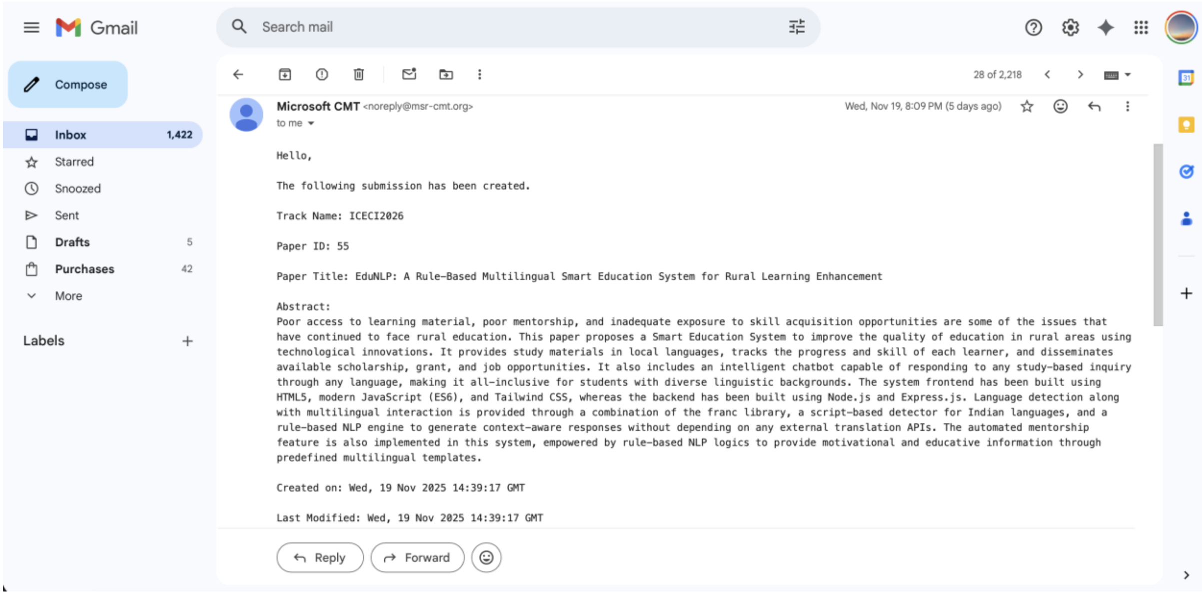
This project, EduNLP – A Rule-Based Multilingual Smart Education System for Rural Learning Enhancement, is based mainly on the following paper for reference.

[3] Khanuja, S., Bansal, D., Mehtani, S., Khosla, A., Dey, A., Gopalan, B. and Talukdar, P., 2021. MuRIL: Multilingual Representations for Indian Languages. arXiv preprint arXiv:2103.10730.

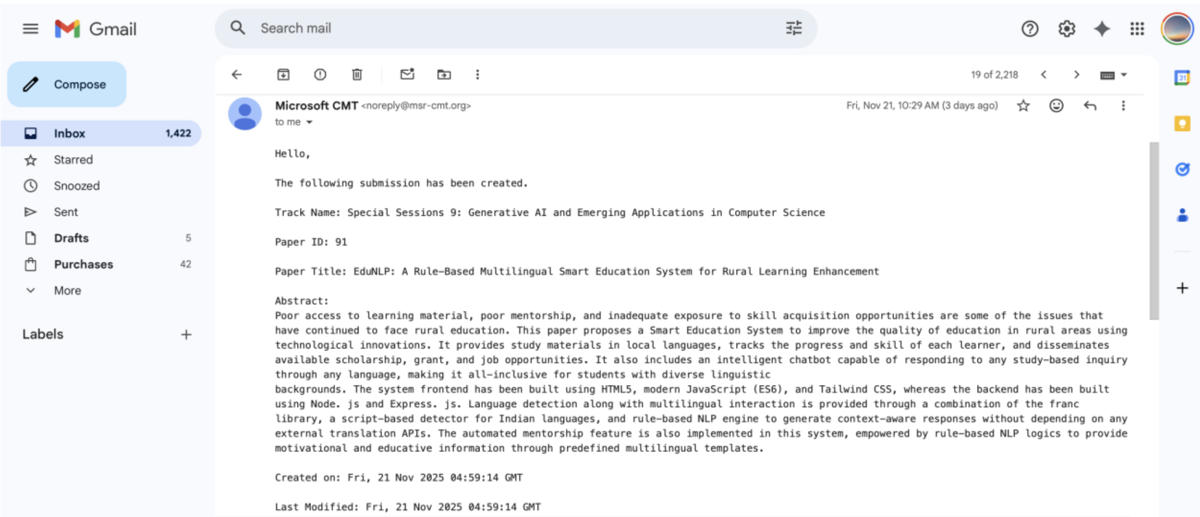
This paper forms the basis of the EduNLP multilingual language-processing component. It first introduces MuRIL, a multilingual model for major Indian languages, which guided the rule-based NLP framework of the project and the adaptation to regional languages. While EduNLP uses a lightweight rule-based approach and not the full transformer model, linguistic diversity coverage and architectural insights from this paper have guided system design and language detection logic.

Appendix

i. Publications



a. ICECI (IEEE) Publication Submission email



b. AICCONS Publication Submission email

ii. Project Report

- Similarity Report

Prema Arokia Mary G - CSE_94_REPORT

ORIGINALITY REPORT

6%	5%	5%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Presidency University Student Paper	4%
2	Submitted to Manipal University Jaipur Online Student Paper	<1%
3	pariwisata.fib.ugm.ac.id Internet Source	<1%
4	link.springer.com Internet Source	<1%
5	Submitted to University of Bedfordshire Student Paper	<1%
6	Submitted to University of Ulster Student Paper	<1%
7	mseuf.edu.ph Internet Source	<1%
8	Submitted to Georgia Southwestern State University Student Paper	<1%
9	dokumen.tips Internet Source	<1%
10	ir.uitm.edu.my Internet Source	<1%

- AI Detection Report

turnitin Page 2 of 50 - AI Writing Overview Submission ID trn:oid:::1:3424062164

***% detected as AI**

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

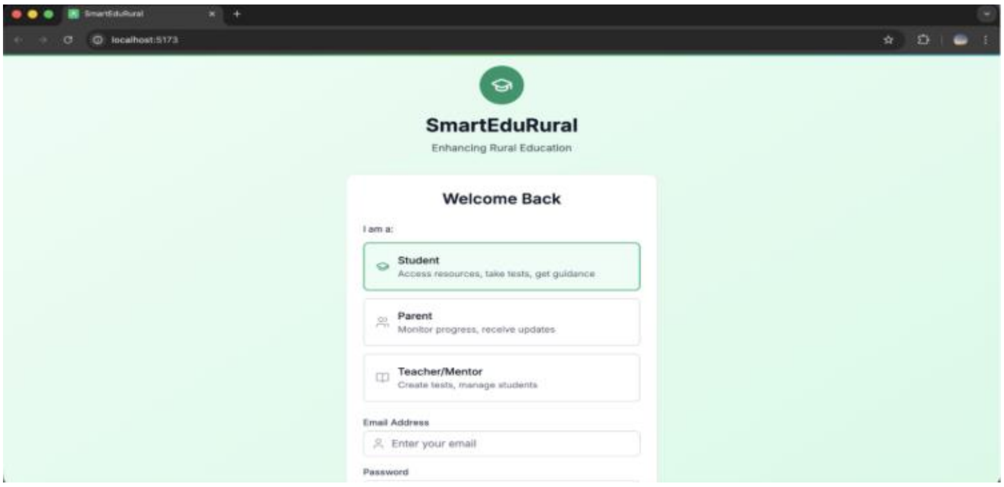
Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

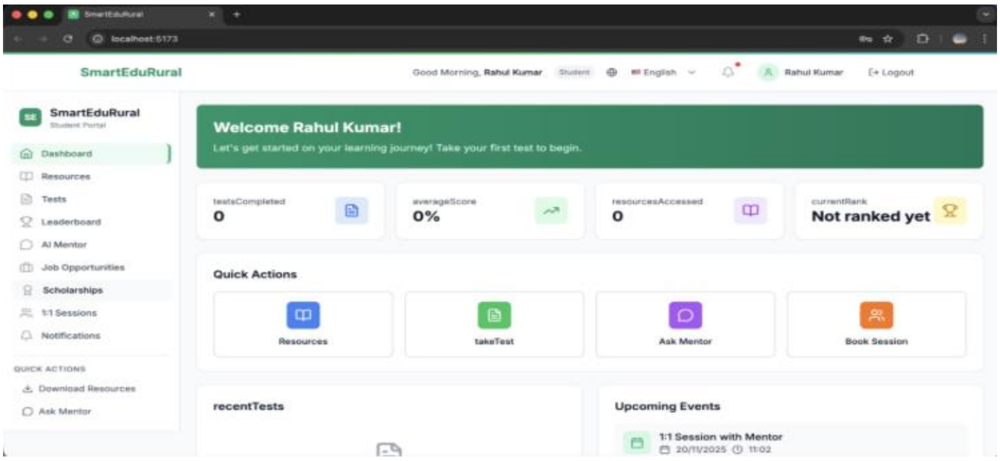
Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

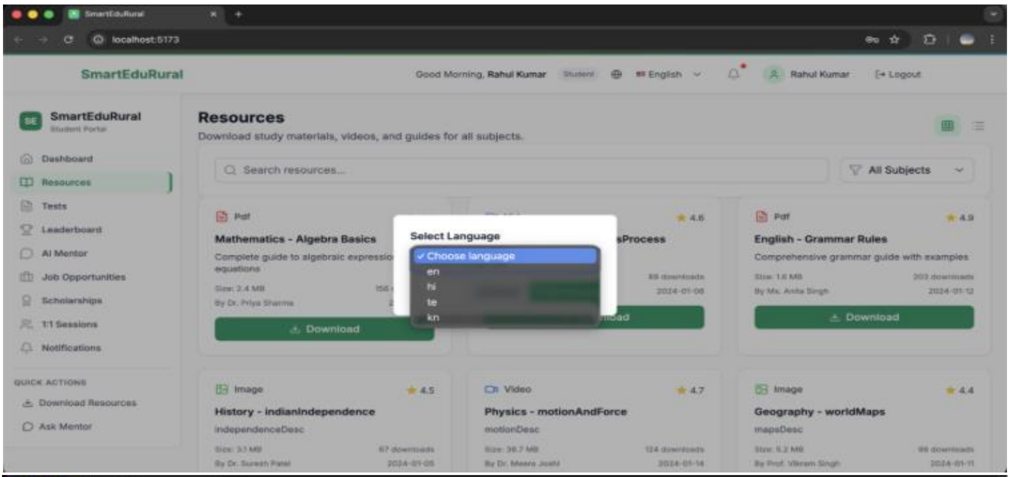
iii. Live Project Demo

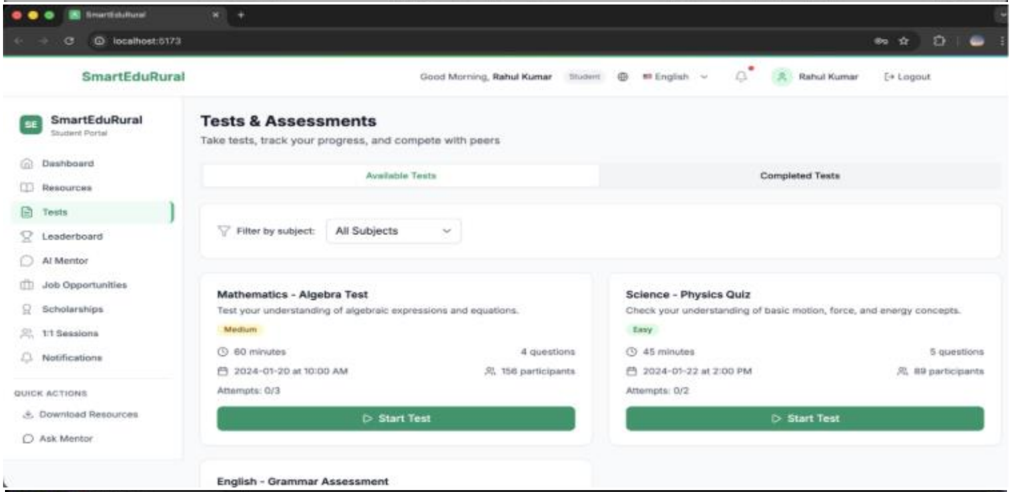
- GitHub: https://github.com/imrxnsk/SmartEduRural.git
- Live Demo: https://smart-edu-rural.vercel.app/

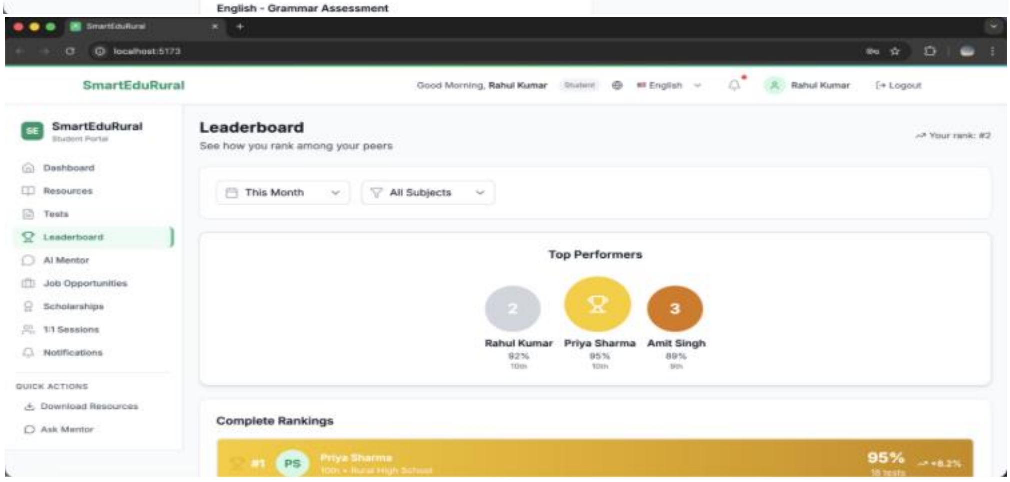
iv. Few Images of Project

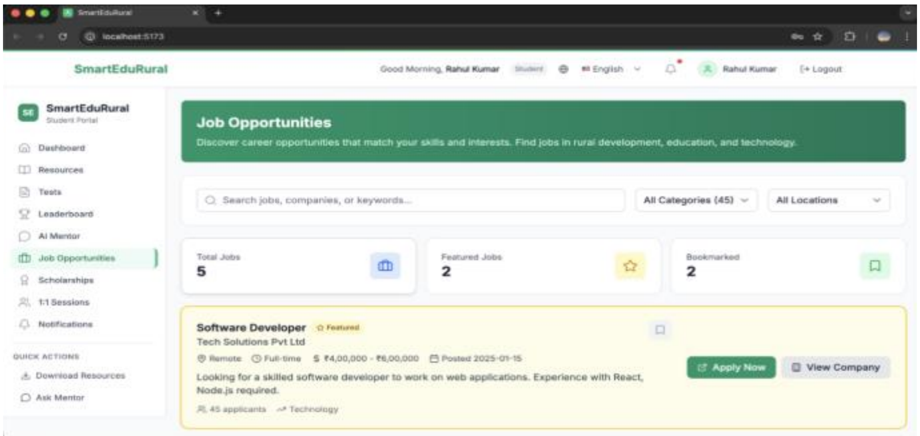


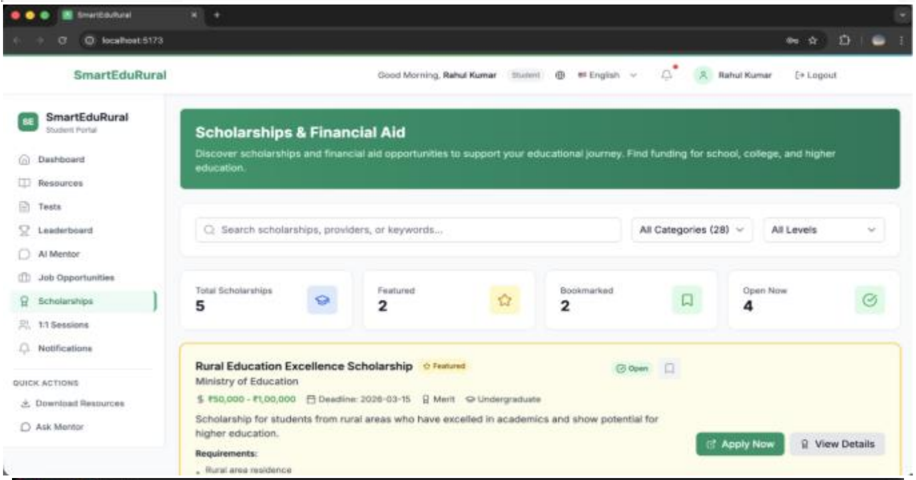


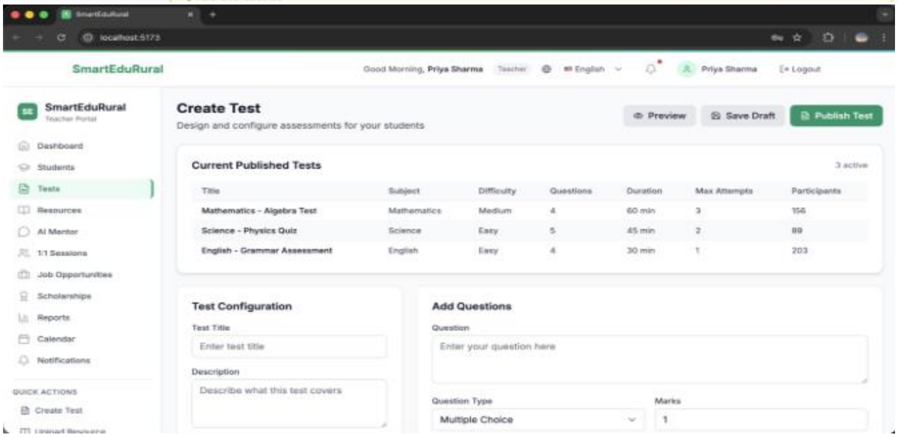

SmartEduRural
Resources
Download study materials, videos, and guides for all subjects.
Select Language
Choose language
en
hi
te
kn
Mathematics - Algebra Basics
English - Grammar Rules
History - indianIndependence
Physics - motionAndForce
Geography - worldMaps

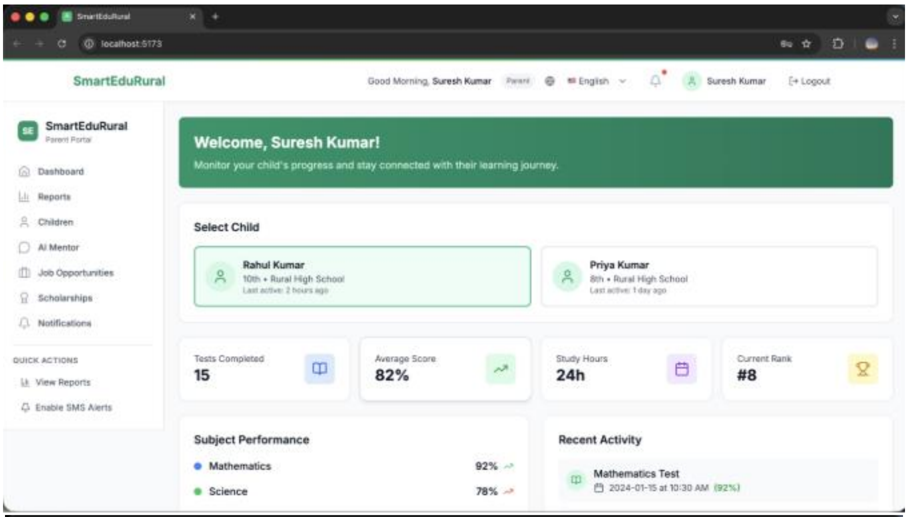

SmartEduRural
Tests & Assessments
Take tests, track your progress, and compete with peers
Available Tests
Completed Tests
Mathematics - Algebra Test
Science - Physics Quiz
Start Test
English - Grammar Assessment

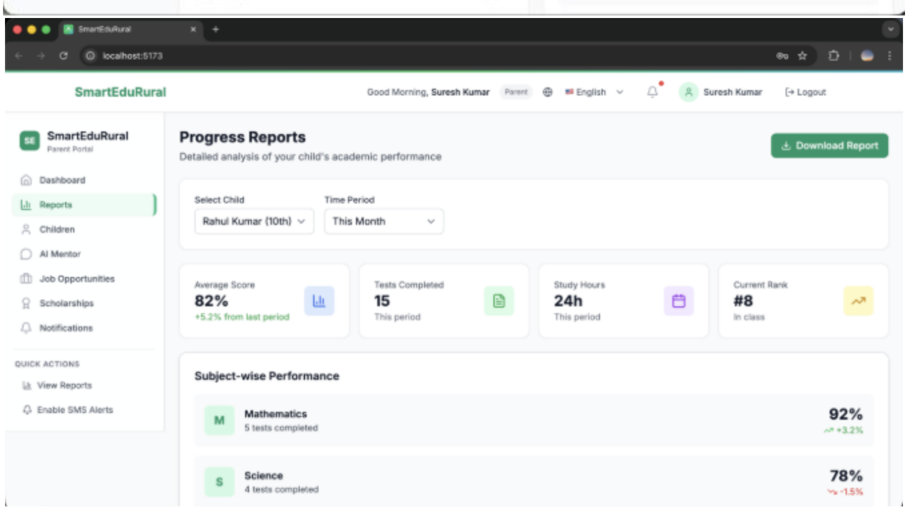

SmartEduRural
Leaderboard
See how you rank among your peers
Top Performers
Rahul Kumar
Priya Sharma
Amit Singh
Complete Rankings


SmartEduRural
localhost:5173
SmartEduRural
Good Morning, Rahul Kumar
English
Rahul Kumar
Logout
SmartEduRural
Student Portal
Dashboard
Resources
Tests
Leaderboard
AI Mentor
Job Opportunities
Scholarships
1:1 Sessions
Notifications
QUICK ACTIONS
Download Resources
Ask Mentor
Job Opportunities
Discover career opportunities that match your skills and interests. Find jobs in rural development, education, and technology.
Search jobs, companies, or keywords...
All Categories (45)
All Locations
Total Jobs
5
Featured Jobs
2
Bookmarked
2
Software Developer
Featured
Tech Solutions Pvt Ltd
Remote
Full-time
₹4,00,000 - ₹6,00,000
Posted 2025-01-15
Looking for a skilled software developer to work on web applications. Experience with React, Node.js required.
45 applicants
Technology
Apply Now
View Company

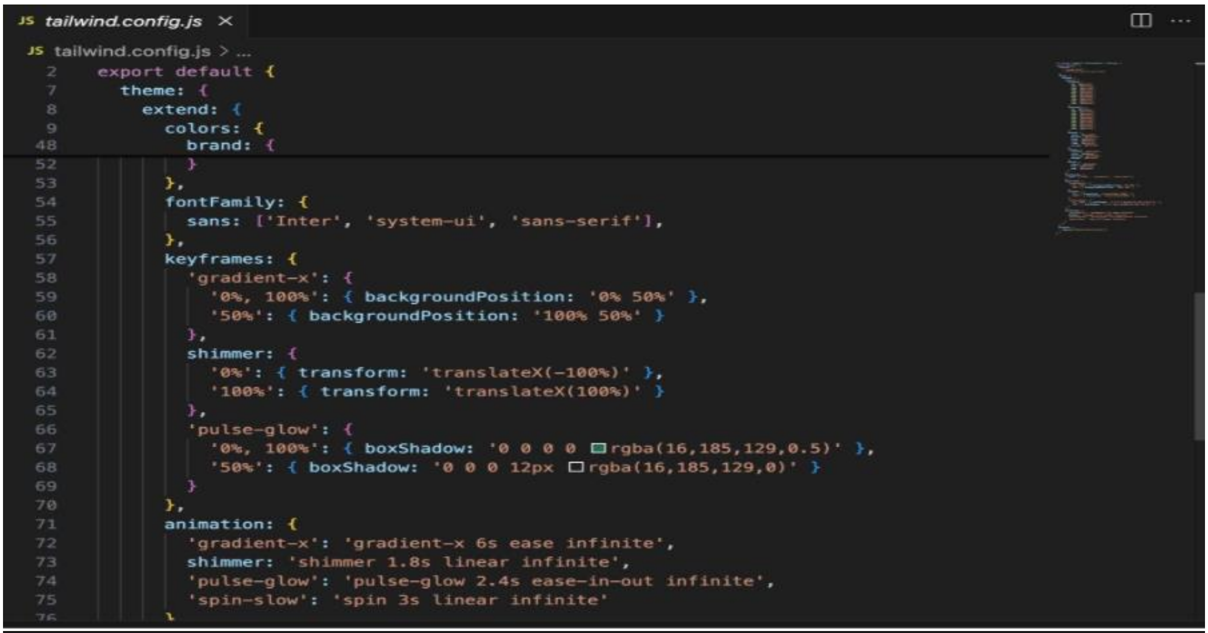

SmartEduRural
localhost:5173
SmartEduRural
Good Morning, Rahul Kumar
English
Rahul Kumar
Logout
SmartEduRural
Student Portal
Dashboard
Resources
Tests
Leaderboard
AI Mentor
Job Opportunities
Scholarships
1:1 Sessions
Notifications
QUICK ACTIONS
Download Resources
Ask Mentor
Scholarships & Financial Aid
Discover scholarships and financial aid opportunities to support your educational journey. Find funding for school, college, and higher education.
Search scholarships, providers, or keywords...
All Categories (28)
All Levels
Total Scholarships
5
Featured
2
Bookmarked
2
Open Now
4
Rural Education Excellence Scholarship
Featured
Open
Ministry of Education
₹50,000 - ₹1,00,000
Deadline: 2026-03-15
Merit
Undergraduate
Scholarship for students from rural areas who have excelled in academics and show potential for higher education.
Requirements:
Rural area residence
Apply Now
View Details


SmartEduRural
localhost:5173
SmartEduRural
Good Morning, Priya Sharma
Teacher
English
Priya Sharma
Logout
SmartEduRural
Teacher Portal
Dashboard
Students
Tests
Resources
AI Mentor
1:1 Sessions
Job Opportunities
Scholarships
Reports
Calendar
Notifications
QUICK ACTIONS
Create Test
Create Test
Design and configure assessments for your students
Preview
Save Draft
Publish Test
Current Published Tests
3 active
Title
Subject
Difficulty
Questions
Duration
Max Attempts
Participants
Mathematics - Algebra Test
Mathematics
Medium
4
60 min
3
156
Science - Physics Quiz
Science
Easy
5
45 min
2
89
English - Grammar Assessment
English
Easy
4
30 min
1
203
Test Configuration
Test Title
Enter test title
Description
Describe what this test covers
Add Questions
Question
Enter your question here
Question Type
Multiple Choice
Marks
1

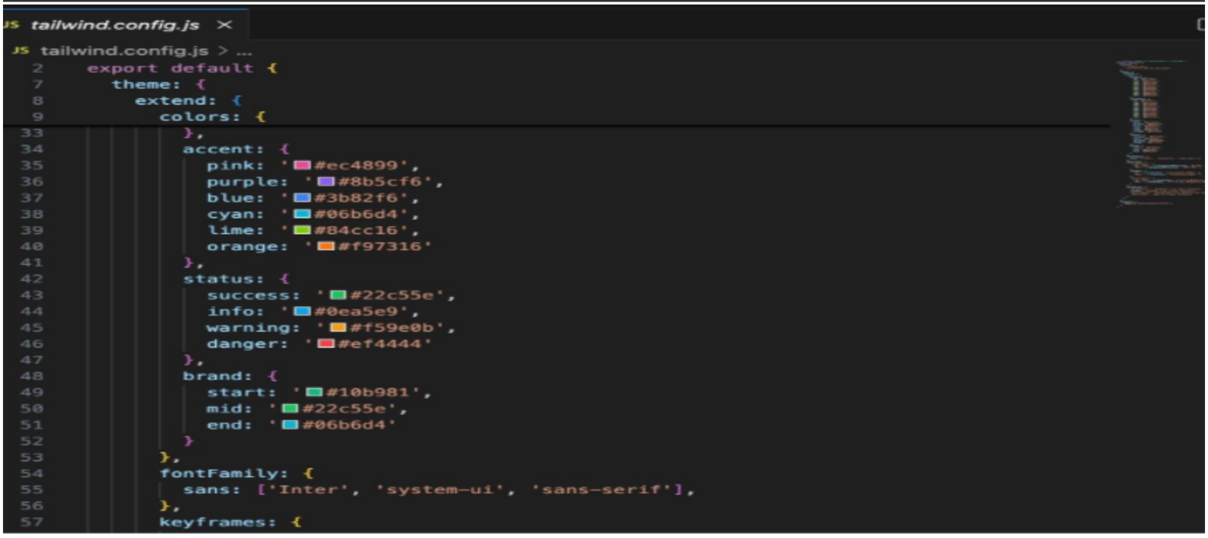

SmartEduRural
Good Morning, Suresh Kumar
Parent
English
Suresh Kumar
Logout
SmartEduRural
Parent Portal
Dashboard
Reports
Children
AI Mentor
Job Opportunities
Scholarships
Notifications
QUICK ACTIONS
View Reports
Enable SMS Alerts
Welcome, Suresh Kumar!
Monitor your child's progress and stay connected with their learning journey.
Select Child
Rahul Kumar
10th • Rural High School
Last active: 2 hours ago
Priya Kumar
8th • Rural High School
Last active: 1 day ago
Tests Completed
15
Average Score
82%
Study Hours
24h
Current Rank
#8
Subject Performance
Mathematics
92%
Science
78%
Recent Activity
Mathematics Test
2024-01-15 at 10:30 AM (92%)


SmartEduRural
Good Morning, Suresh Kumar
Parent
English
Suresh Kumar
Logout
SmartEduRural
Parent Portal
Dashboard
Reports
Children
AI Mentor
Job Opportunities
Scholarships
Notifications
QUICK ACTIONS
View Reports
Enable SMS Alerts
Progress Reports
Detailed analysis of your child's academic performance
Download Report
Select Child
Rahul Kumar (10th)
Time Period
This Month
Average Score
82%
+5.2% from last period
Tests Completed
15
This period
Study Hours
24h
This period
Current Rank
#8
In class
Subject-wise Performance
Mathematics
5 tests completed
92%
+3.2%
Science
4 tests completed
78%
-1.5%

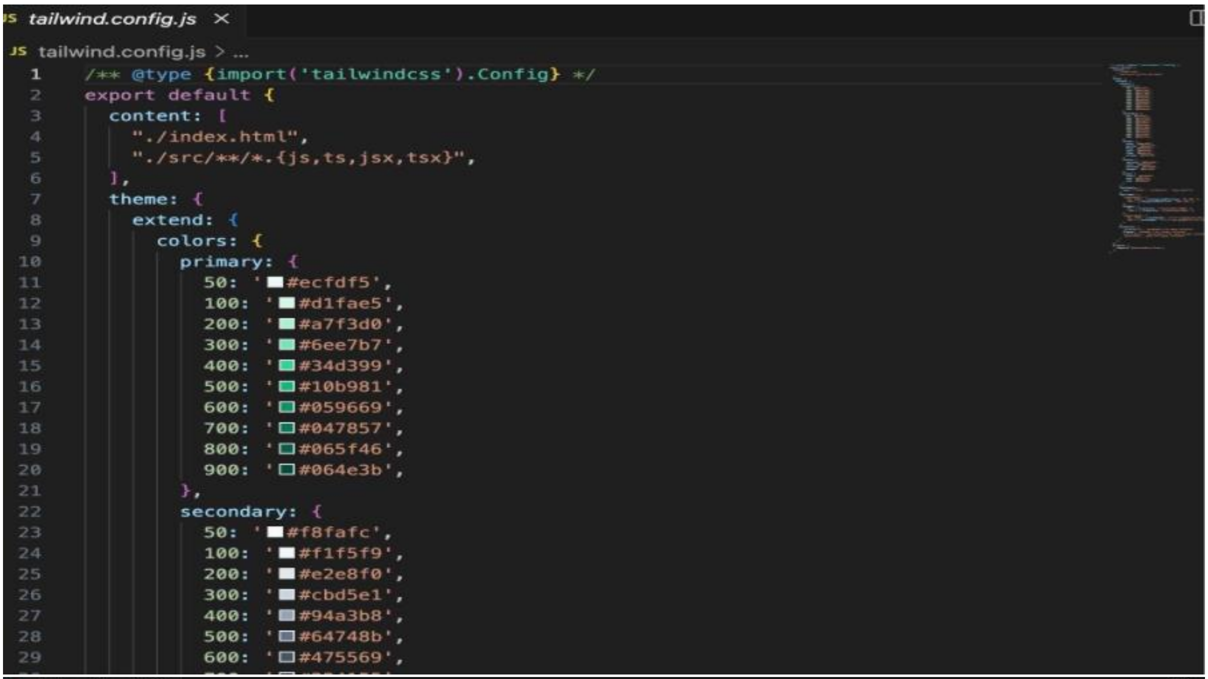
v. Few Code Snippets

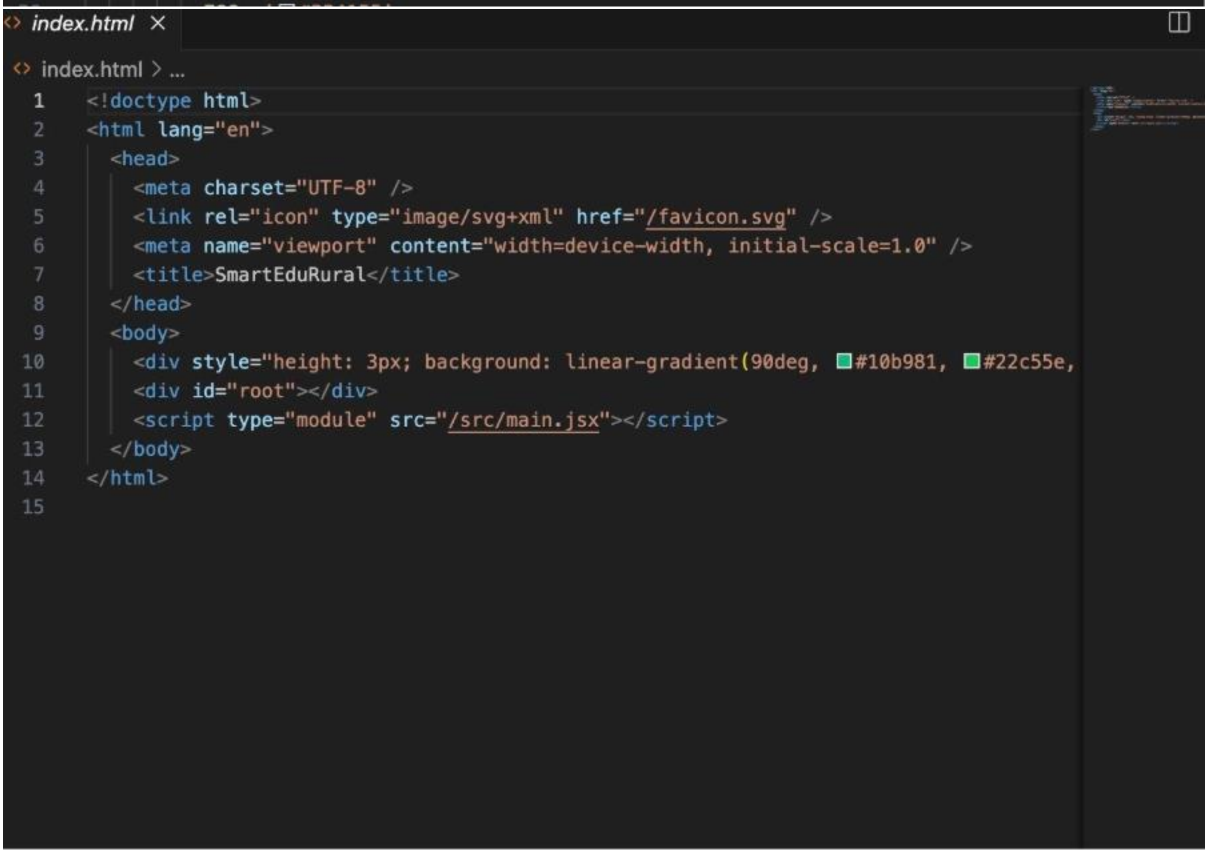


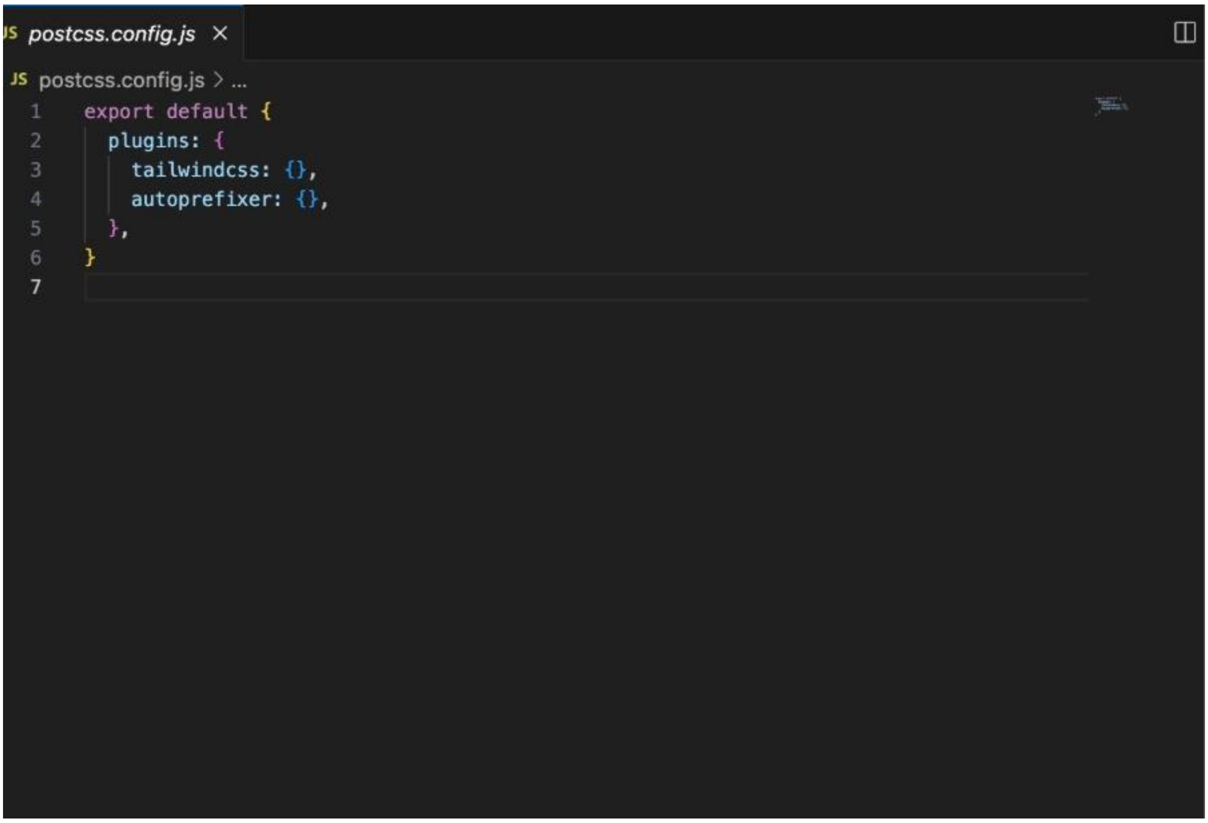
```js
tailwind.config.js
tailwind.config.js > ...
export default {
  theme: {
    extend: {
      colors: {
        brand: {
        }
      },
      fontFamily: {
        sans: ['Inter', 'system-ui', 'sans-serif'],
      },
      keyframes: {
        'gradient-x': {
          '0%, 100%': { backgroundPosition: '0% 50%' },
          '50%': { backgroundPosition: '100% 50%' }
        },
        shimmer: {
          '0%': { transform: 'translateX(-100%)' },
          '100%': { transform: 'translateX(100%)' }
        },
        'pulse-glow': {
          '0%, 100%': { boxShadow: '0 0 0 0 rgba(16,185,129,0.5)' },
          '50%': { boxShadow: '0 0 0 12px rgba(16,185,129,0)' }
        }
      },
      animation: {
        'gradient-x': 'gradient-x 6s ease infinite',
        shimmer: 'shimmer 1.8s linear infinite',
        'pulse-glow': 'pulse-glow 2.4s ease-in-out infinite',
        'spin-slow': 'spin 3s linear infinite'
```



```js
tailwind.config.js
tailwind.config.js > ...
export default {
  theme: {
    extend: {
      colors: {
        },
        accent: {
          pink: '#ec4899',
          purple: '#8b5cf6',
          blue: '#3b82f6',
          cyan: '#06b6d4',
          lime: '#84cc16',
          orange: '#f97316'
        },
        status: {
          success: '#22c55e',
          info: '#0ea5e9',
          warning: '#f59e0b',
          danger: '#ef4444'
        },
        brand: {
          start: '#10b981',
          mid: '#22c55e',
          end: '#06b6d4'
        }
      },
      fontFamily: {
        sans: ['Inter', 'system-ui', 'sans-serif'],
      },
      keyframes: {
```


tailwind.config.js
tailwind.config.js > ...
/** @type {import('tailwindcss').Config} */
export default {
content: [
"./index.html",
"./src/**/*.{js,ts,jsx,tsx}",
],
theme: {
extend: {
colors: {
primary: {
50: '#ecfdf5',
100: '#d1fae5',
200: '#a7f3d0',
300: '#6ee7b7',
400: '#34d399',
500: '#10b981',
600: '#059669',
700: '#047857',
800: '#065f46',
900: '#064e3b',
},
secondary: {
50: '#f8fafc',
100: '#f1f5f9',
200: '#e2e8f0',
300: '#cbd5e1',
400: '#94a3b8',
500: '#64748b',
600: '#475569',


index.html
index.html > ...
<!doctype html>
<html lang="en">
<head>
<meta charset="UTF-8" />
<link rel="icon" type="image/svg+xml" href="/favicon.svg" />
<meta name="viewport" content="width=device-width, initial-scale=1.0" />
<title>SmartEduRural</title>
</head>
<body>
<div style="height: 3px; background: linear-gradient(90deg, #10b981, #22c55e,
<div id="root"></div>
<script type="module" src="/src/main.jsx"></script>
</body>
</html>



vi. GitHub Repository