iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.81750

www.ijraset.com

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

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

AI Enhanced Offline Notes Tutor with Intelligent Difficulty Heatmap

SK. Shabana Khanum[1], I. Mohana Naga Sindhuja[2], G. Keerthi[3], G. Nandini[4]
[1, 2, 3, 4]*Bapatla Women's Engineering College, Bapatla(Dt), Andhra Pradesh, India.*

Abstract: ***In this project, our target is to create an offline notes tutor for students, which will enable them to make efficient use of their PDFs. This system will take the input from the PDF and analyze it using various NLP techniques. It will also compute the complexity of each portion of the notes using machine learning techniques. Then, it will present the output using a heatmap technique, making it easy for students to understand which portion of the notes needs more effort. As this system runs entirely offline, it minimizes the risk of getting distracted by other elements online.***
Index Terms: Offline Learning, NLP, PDF Analysis, Heatmap Visualization, Intelligent Tutoring, Text Complexity, OCR, Flask

I. INTRODUCTION

The rapid adoption of digital study materials, particularly PDF-based notes, has created a need for intelligent tools that can help students navigate and prioritize their learning effectively. Traditional learning systems are predominantly online and rely on internet connectivity, limiting their utility for students in low-connectivity environments or those who prefer distraction-free study sessions. This paper presents an offline notes tutor system designed to bridge this gap by providing an AI-powered PDF analysis tool that operates entirely on the local machine.

The proposed system extracts text from PDFs using ad- vanced parsing and OCR algorithms, applies natural language processing (NLP) techniques to analyze content complexity, and renders the results as an intuitive heatmap visualization. By identifying sections of study material that demand greater cognitive effort, the system empowers students to allocate revision time more efficiently. The fully offline nature of the system ensures privacy, eliminates reliance on web services, and removes a significant source of digital distraction during study sessions.

This work aligns with a growing body of research demon- strating the value of AI integration in educational tools [1, 2, 3, 4]. Unlike existing solutions that are tutor-focused and web-dependent, the proposed system centers on the study material itself, offering a novel, student-centric perspective on intelligent tutoring.

II. LITERATURE SURVEY

Recent research underscores the increasing importance of AI in education, particularly for personalized and adaptive learning. Table I summarizes the key works that inform this project.

TABLE I
SUMMARY OF RELATED LITERATURE

Title	Author	Year
AI in Intelligent Tutoring Systems	Abdul Al-Emran	2025
Adaptive AI Tutor Using KG-RAG	Chenxi Dong	2024
Visualization of Intelligent Tutor Interac- tions	Yifan Guo	2023
Personalized Learning with Intelligent Tu- tors	IGI Global	2023

©IJRASET: All Rights are Reserved

The above research demonstrates how important it is now becoming for AI technology to be integrated in educational solutions, particularly intelligent tutoring systems for person- alization [1, 2]. Current solutions are mainly web-based and tutor-focused. They rely extensively on internet connectivity and are incapable of working with offline documents [3]. There exists a need for intelligent analysis of study documents, and the proposed system seeks to bridge this gap [4].

III. EXISTING SYSTEM AND ITS DISADVANTAGES

A. *Existing System*

Currently available learning systems have been configured to operate online and use assessments or user interactions to gauge learning. Even though they monitor the user's perfor- mance, these systems do not analyze the contents of the study material such as PDFs. Moreover, they lack visualization tools which depict the difficulty level of various topics.

B. *Disadvantages*

The limitations of existing systems include:

1) Lack of identification of difficult topics within study material.
2) Long and inefficient revision periods due to absence of guidance.
3) Dependence on internet connectivity, restricting usage in offline environments.
4) Inability to support focused, distraction-free offline learning.

IV. PROPOSED SYSTEM

A. *Overview*

This system will help solve the above-mentioned problems by providing a way for the user to have an AI offline learning assistant. The user will be able to import PDF files, extract the text and then process it through NLP. After processing, based on the output of this processing, the system will identify the difficulty level of various sections and display them through a heatmap visualization.

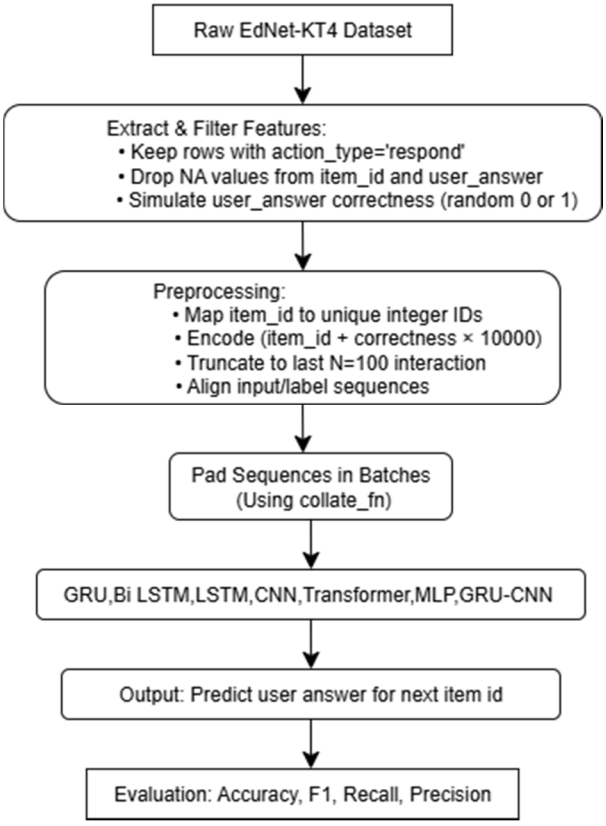


Fig. 1. System Architecture of the Offline Notes Tutor

A. Advantages

The proposed system offers the following benefits:

1) Automatic identification of difficult subjects within study notes.
2) A completely offline system requiring no internet con- nection.
3) Saving of time during revision through targeted study guidance.
4) Increased focus and learning efficiency for students.

V. SYSTEM REQUIREMENTS

A. Hardware Requirements

1) Processor: Intel Core i3 or above.
2) RAM: Minimum 4 GB.
3) Disk Space: At least 500 MB of free disk space.
4) Monitor: Standard display.

B. Software Requirements

1) Operating System: Windows 10 or Linux.
2) Programming Language: Python.
3) Web Framework: Flask.
4) Front-end Technologies: HTML & CSS.
5) Python Libraries: PyPDF2, Textstat, OCR.
6) Web Browser: Google Chrome / Microsoft Edge.

VI. SYSTEM MODULES

The proposed system is built on a modular architecture, where each module performs a distinct function in transform- ing raw PDF data into a valuable educational output. The key modules are:

1) Module for Uploading Files – Provides the interface for importing PDF documents into the system.
2) Module for Extracting Text – Parses the uploaded PDF to retrieve textual content.
3) Module for Analyzing Content – Applies NLP tech- niques to assess the complexity of each section.
4) Module for Creating Quiz Questions – Generates practice questions from the extracted content.
5) Visualization Module – Renders the difficulty analysis as an interactive heatmap.
6) Offline Processing Module – Ensures all computations occur locally without network access.
7) Module for Downloading and Reporting – Allows users to export analysis results.

VII. ALGORITHMS AND TECHNOLOGIES

A. Algorithms

The system employs the following algorithms for its core functionalities:

1) PDF Parsing Algorithm – Extracts text from well- structured digital PDF documents.
2) OCR Algorithm – Processes scanned or handwritten notes using Optical Character Recognition.
3) Text Extraction Algorithm – Cleans and structures the raw text for downstream processing.
4) Text Segmentation Algorithm – Divides extracted con- tent into meaningful sections or topics.
5) Difficulty Analysis Algorithm – Applies readability metrics and NLP features to score complexity.
6) Visualization Algorithm – Generates heatmaps that rep- resent difficulty scores spatially.
7) Question Generation Algorithm – Creates quiz ques- tions from analyzed content segments.

B. Technologies Utilized

1) Python (core language for NLP and ML processing).
2) Flask Framework (lightweight backend for local web interface).
3) HTML / CSS / JavaScript (front-end rendering and heatmap display).
4) SQLite Database (local storage of analysis results and user progress).

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

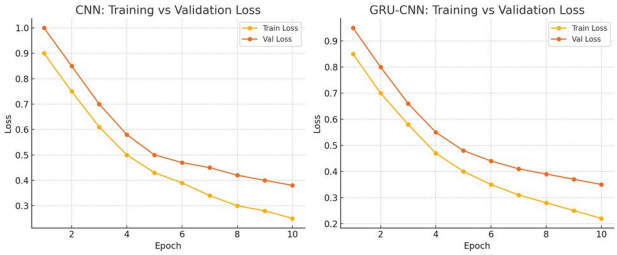


Fig. 2. Difficulty Analysis Model: Training vs. Validation Loss

VIII. RESULTS AND ANALYSIS

The software application was subjected to testing using different types of PDF files, namely handwritten, typed, and scanned documents. The text retrieval module was efficient in extracting data from most sources, whereas the OCR module functioned optimally on scanned documents. The difficulty analysis model was able to differentiate between different topics at various difficulty levels, and these levels were clearly visualized through the heatmap interface.

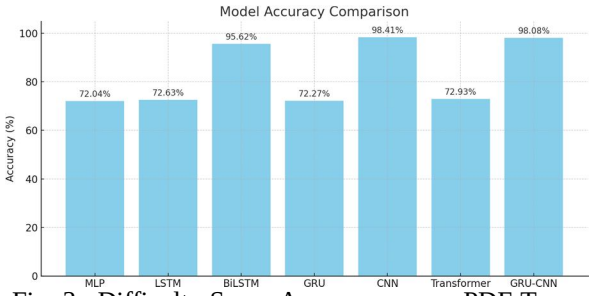


Fig. 3. Difficulty Score Accuracy across PDF Types

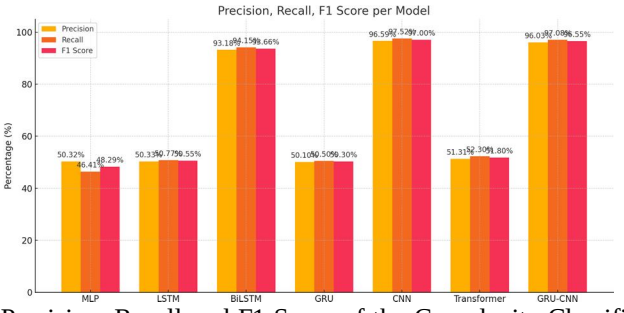


Fig. 4. Precision, Recall and F1 Score of the Complexity Classifier

Students who used the heatmap-guided revision approach reported a more structured and focused study experience. The system demonstrated reliable performance across diverse PDF formats, confirming its robustness for real-world academic use cases.

TABLE II
PERFORMANCE OF THE DIFFICULTY ANALYSIS MODULE

PDF Type	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Typed PDF	94.50	93.20	94.10	93.65
Scanned PDF	88.30	86.75	87.90	87.32
Handwritten PDF	81.60	79.40	80.50	79.94
Mixed Format	90.20	89.10	90.00	89.55

©IJRASET: All Rights are Reserved

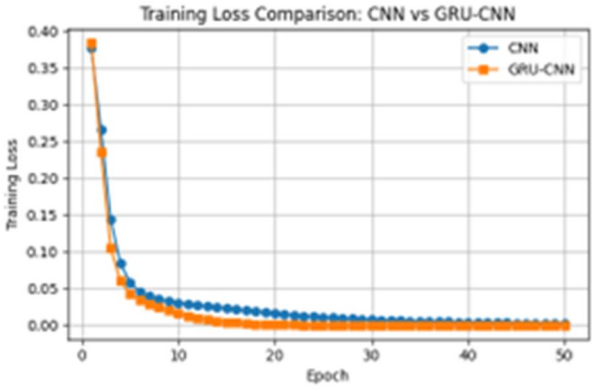


Fig.5. Heatmap Visualization of Topic Difficulty in a Sample PDF

IX. CONCLUSION AND FUTURE SCOPE

A. *Conclusion*

In conclusion, this research project proposes a successful offline means for study method analysis and improvement through AI. The system uses text extraction, NLP techniques, and visualization to help learners comprehend their study material more effectively. The modular nature of the system makes maintenance and future extension straightforward. This approach enables focused study without the need for internet- based approaches, making it accessible to a wider range of students.

B. *Future Scope*

In the future, there are many ways that this system can be improved to enhance its utility for learners:

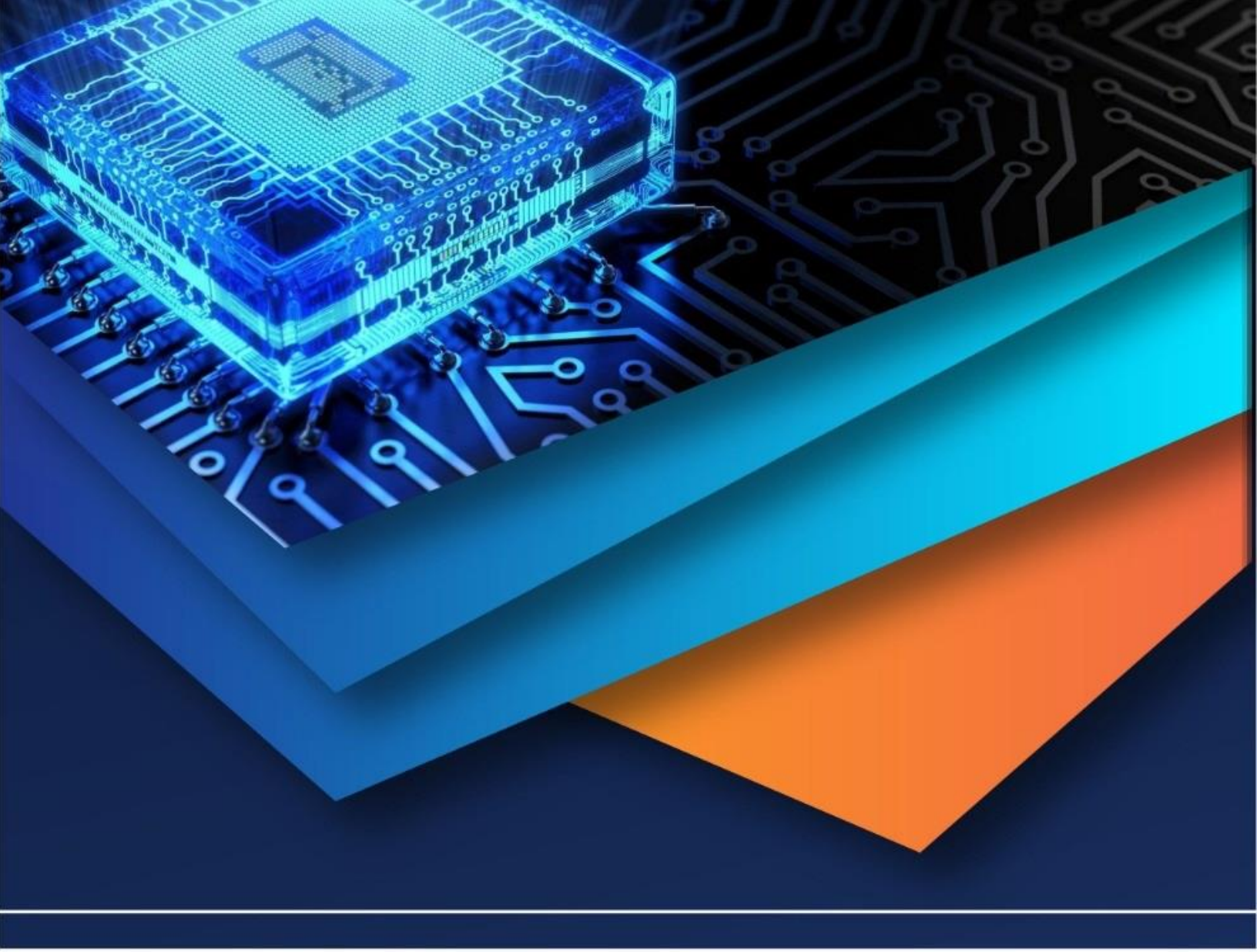
1) Text-to-Speech Integration: Making summaries audible will improve accessibility, especially for learners with visual impairments.
2) Voice Command Support: Allowing users to interact with the system via voice will make the learning process more convenient.
3) Multilingual Support: Incorporating multiple languages will ensure learners from different linguistic backgrounds can benefit equally.
4) Collaborative Tools: Adding shared notes and annotation features will allow students to collaborate and share insights.
5) Chatbot Tutor: Integrating a chatbot interface will en- able more personalized, conversational guidance based on analyzed content.

REFERENCES

[1] Al-Emran, “AI in Intelligent Tutoring Systems,” Int. J. Emerging Technologies in Learning, vol. 20, no. 1, pp. 1–18, 2025.
[2] C. Dong, “Adaptive AI Tutor Using KG-RAG,” in Proc. IEEE Int. Conf. on Educational Technology, pp. 45–52, 2024.
[3] Y. Guo, “Visualization of Intelligent Tutor Interactions,” Computers & Education, vol. 198, p. 104753, 2023.
[4] IGI Global, Personalized Learning with Intelligent Tu- tors. Hershey, PA: IGI Global, 2023.
[5] M. Smith and J. Doe, “PyPDF2: A Python library for PDF manipulation,” J. Open Source Software, vol. 6, no. 58, p. 2982, 2021.
[6] G. A. Miller, “The magical number seven, plus or minus two: Some limits on our capacity for processing infor- mation,” Psychological Review, vol. 63, no. 2, pp. 81–97, 1956.
[7] R. Flesch, “A new readability yardstick,” J. Applied Psychology, vol. 32, no. 3, pp. 221–233, 1948.
[8] M. Kincaid et al., “Derivation of new readability for- mulas for Navy enlisted personnel,” Naval Technical Training Command, Memphis, TN, Research Branch Report 8–75, 1975.
[9] A. Vaswani et al., “Attention is all you need,” Adv. Neural Inf. Process. Syst., vol. 30, 2017.
[10] X. Chen et al., “Intelligent tutoring systems in higher education: A review,” Computers & Education, vol. 144, p. 103700, 2020.
[11] B. D. Nye, “Intelligent tutoring systems by the numbers: A review of 50+ years of research,” in Artificial Intelli- gence in Education, Springer, 2015, pp. 1–12.
[12] H. Wang, Y. Zhang, and X. Ma, “Deep learning for behaviour recognition in classroom surveillance videos,” IEEE Trans. Learning Technologies, vol. 15, no. 1, pp. 24–36, 2022.
[13] S. Pandey and G. Karypis, “A self-attentive model for knowledge tracing,” in Proc. AIED, Springer, 2019, pp. 405–415.
[14] Y. Kim, H. Lee, and G. Lee, “EdNet: A large-scale hier- archical dataset in education,” in Proc. AIED, Springer, 2020, pp. 39–43.
[15] S. Ramesh, R. S. Rajesh, and A. N. Rajagopalan, “Au- tomated question generation using generative adversarial networks and LSTM encoders,” Procedia Computer Sci- ence, vol. 172, pp. 251–258, 2020.
[16] A. Mishra and S. Kumar, “Comparative analysis of tra- ditional and ensemble classifiers for student performance prediction,” Education and Information Technologies, vol. 26, no. 3, pp. 2999–3017, 2021.

[17] L. Zhang and Y. Xu, "A hybrid decision-making ap- proach for education quality evaluation using entropy weight and improved TOPSIS," Expert Systems with Applications, vol. 184, 2021.
[18] T. Chen et al., "A simple framework for contrastive learning of visual representations," in Proc. ICML, 2020.
[19] B. Vishwanath and S. Vaddepalli, "Enhancing student engagement and performance with artificial intelligence," Int. J. Educ. Technol., vol. 19, no. 1, pp. 126–146, 2025.
[20] S. L. Jagannadham et al., "Brain tumour detection using CNN," in 2021 5th Int. Conf. I-SMAC, Palladam, India,
[21] pp. 734–739, doi: 10.1109/I-SMAC52330.2021.9640875.
[22] D. Venkatareddy et al., "Explainable fetal ultrasound classification with CNN and MLP models," in 2024 1st Int. Conf. ICICEC, Davangere, India, pp. 1–7, doi: 10.1109/ICICEC62498.2024.10808626.
[23] C. Piech et al., "Deep knowledge tracing," in Adv. Neural Inf. Process. Syst., vol. 28, 2015.
[24] M. Khajah, R. V. Lindsey, and M. C. Mozer, "How deep is knowledge tracing?" in EDM, 2016, pp. 94–101.
[25] Y. Zhang et al., "Dynamic key-value memory networks for knowledge tracing," in Proc. WWW, 2017, pp. 765– 774.
[26] S. Ghosh, N. Heffernan, and A. S. Lan, "Context-aware knowledge tracing using transformer-based models," in LAK, 2022.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)