

## COREGAP: Concept Gap Detection System for Personalized Learning

Dr. K. Prem Kumar<sup>1</sup>, Shaik Noureesh<sup>2\*</sup>, Tanisha Gundu<sup>3</sup>, Suniganti Venkateshwar Reddy<sup>4</sup>, Satai Vishal Kumar<sup>5</sup>

<sup>1</sup>Head of the Department, Department of Artificial Intelligence and Machine Learning, ACE Engineering College, India

<sup>2,3,4,5</sup>UG Scholar, Department of Artificial Intelligence and Machine Learning, ACE Engineering College, India

Email: [noureesh2302@gmail.com](mailto:noureesh2302@gmail.com)

### ARTICLE INFO

#### Article history:

Received 15 May 2026

Accepted 21 May 2026

Available online 29 May 2026

#### Keywords:

Concept Gap Detection, Personalized Learning, Machine Learning, Educational Technology, Diagnostic Systems, Concept Mapping

#### Indexed in:



INDEX COPERNICUS  
INTERNATIONAL



and in major libraries

### ABSTRACT

Traditional educational systems primarily evaluate students based on examination scores, which often fail to measure actual conceptual understanding. Many students rely on memorization techniques rather than developing a strong foundation of concepts, leading to poor understanding of prerequisite topics and difficulties in advanced learning stages. Existing evaluation methods provide limited personalized feedback and are unable to identify specific conceptual weaknesses effectively. This paper proposes COREGAP, a Concept Gap Detection System designed to analyze student understanding at the concept level using diagnostic assessments, concept mapping, and prerequisite relationship analysis. The proposed system integrates machine learning techniques, rule-based analysis, and performance tracking to identify weak concepts, missing prerequisite knowledge, and learning patterns in real time. Based on the analysis, the system generates personalized feedback and adaptive learning recommendations to improve conceptual clarity and learning efficiency. Experimental observations indicate that the proposed approach provides deeper insights into student learning behavior compared to traditional evaluation methods. The system can play a significant role in intelligent educational platforms and personalized learning environments.

© 2026 International Journal of Advanced Research in Science and Technology (IJARST).

All rights reserved.

## I. INTRODUCTION

In modern educational environments, evaluating student performance solely through examination scores does not accurately represent conceptual understanding. Many students rely on memorization techniques rather than understanding the core concepts and prerequisite relationships between topics. This often results in weak foundational knowledge, learning gaps, and difficulty in understanding advanced subjects.

Traditional assessment methods mainly focus on marks and overall performance rather than identifying specific conceptual weaknesses. While such methods are useful for grading, they fail to provide detailed insights into individual learning behavior, weak concepts, and prerequisite deficiencies. Instructors also face challenges in monitoring the conceptual progress of a large number of students effectively.

Recent advancements in Artificial Intelligence (AI), Machine Learning (ML), and Educational Technology have enabled the development of

[www.ijarst.com](http://www.ijarst.com)

intelligent learning systems capable of personalized analysis and adaptive feedback. Concept-based learning systems can analyze student responses, detect weak areas, track learning patterns, and recommend targeted improvements based on prerequisite dependencies.

This paper presents the design, architecture, and operational workflow of COREGAP, a Concept Gap Detection System for Personalized Learning. The proposed system uses diagnostic assessments, concept mapping, rule-based analysis, and machine learning techniques to identify conceptual gaps and missing prerequisite knowledge. The system provides personalized feedback and adaptive learning recommendations to improve conceptual clarity, learning efficiency, and student performance.

## II. LITERATURE REVIEW

Recent research in intelligent educational systems has focused on concept-level learning analysis and personalized assessment techniques. Early learning evaluation methods primarily relied on examination scores and overall grading systems to measure student performance. However, these approaches failed to identify specific conceptual weaknesses and prerequisite knowledge gaps among students.

Modern machine learning and educational data mining techniques have significantly improved personalized learning systems. Concept mapping approaches are widely used to establish relationships between topics and prerequisite dependencies. Knowledge tracing models help analyze student learning behavior and track conceptual understanding over time. Diagnostic assessment systems enable detailed evaluation of student performance at the concept level rather than relying solely on final scores.

Researchers have also explored adaptive learning methods that provide personalized recommendations based on individual performance patterns. Machine learning models trained on student interaction data can identify weak concepts, predict learning difficulties, and recommend suitable learning paths. Although existing systems demonstrate effective concept analysis capabilities, many lack real-time feedback mechanisms and integrated prerequisite gap detection. The proposed COREGAP system addresses these limitations by combining diagnostic assessments, concept mapping, rule-based analysis, and personalized recommendation mechanisms to improve conceptual understanding and learning efficiency.

### III. METHODOLOGY

The methodology of the proposed COREGAP system is designed to perform intelligent concept extraction, diagnostic assessment, and personalized learning analysis. The system follows a structured pipeline that transforms uploaded study material and student responses into actionable insights regarding conceptual understanding and prerequisite knowledge gaps. The workflow is shown in Fig. 1.

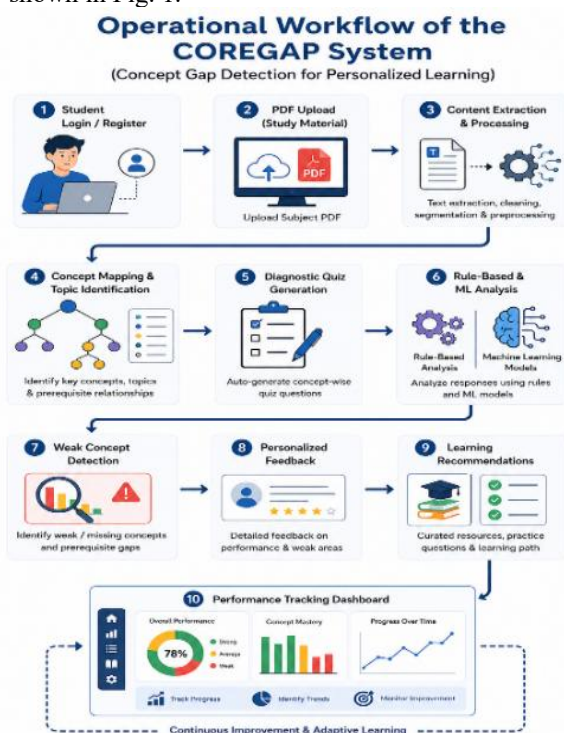


Fig. 1. Operational workflow of the COREGAP System

#### A. PDF Upload and Data Acquisition

The first stage involves uploading subject-related study material in PDF format. The uploaded documents may contain theoretical concepts, definitions, examples, and topic explanations. The system collects educational content for further processing and concept extraction.

Mathematically, the uploaded document can be represented as:

$$D = \{P_1, P_2, P_3, \dots, P_n\}$$

where each  $P_i$  represents a page extracted from the uploaded PDF document.

#### B. Content Extraction and Preprocessing

Raw PDF content often contains unnecessary symbols, formatting inconsistencies, and redundant text. Preprocessing improves content quality and prepares the data for concept analysis. The following operations are performed:

1. Text extraction from PDF documents
2. Removal of special characters and stop words
3. Sentence segmentation and tokenization
4. Text normalization and preprocessing

These preprocessing steps improve concept identification accuracy and reduce processing complexity.

#### C. Concept Mapping and Topic Identification

The extracted content is analyzed to identify key concepts and prerequisite relationships between topics. Natural Language Processing (NLP) and rule-based techniques are used to detect important educational terms and establish concept dependency structures.

If  $C$  represents the set of extracted concepts:

$$C = \{c_1, c_2, c_3, \dots, c_n\}$$

then prerequisite relationships are mapped between related concepts to create a structured learning graph.

#### D. Diagnostic Quiz Generation

The system automatically generates concept-based quiz questions from the extracted educational content. Each question is linked to a specific concept and its prerequisite topics.

Quiz generation helps evaluate conceptual understanding rather than memorization ability. The generated assessment is adaptive and focuses on identifying weak areas in student learning.

#### E. Rule-Based and Machine Learning Analysis

Student responses are analyzed using rule-based logic and machine learning models. The analysis evaluates:

- Correct and incorrect responses
- Concept mastery level
- Prerequisite knowledge gaps
- Learning behavior patterns

The conceptual understanding score can be represented as:

$$S = \text{Correct Concept Responses} / \text{Total Concept Questions}$$

Machine learning models analyze performance trends and improve recommendation accuracy over time.

### F. Weak Concept Detection

The system identifies weak concepts and missing prerequisite knowledge by analyzing assessment performance. Concepts with low mastery scores are marked as weak areas.

If:

$$S < \text{Threshold}$$

then the concept is classified as weak and requires additional learning support.

This stage helps students focus on concepts that require improvement.

### G. Personalized Feedback and Recommendations

Based on detected learning gaps, the system generates personalized feedback and adaptive learning recommendations. Recommendations may include:

- Suggested prerequisite topics
- Additional study material
- Practice questions
- Revision recommendations

This ensures targeted learning and improves conceptual clarity.

### H. Performance Tracking Dashboard

The system continuously tracks student progress and learning improvements through a performance dashboard. The dashboard provides:

- Concept-wise performance analysis
- Progress tracking graphs
- Weak and strong concept identification
- Learning trend visualization

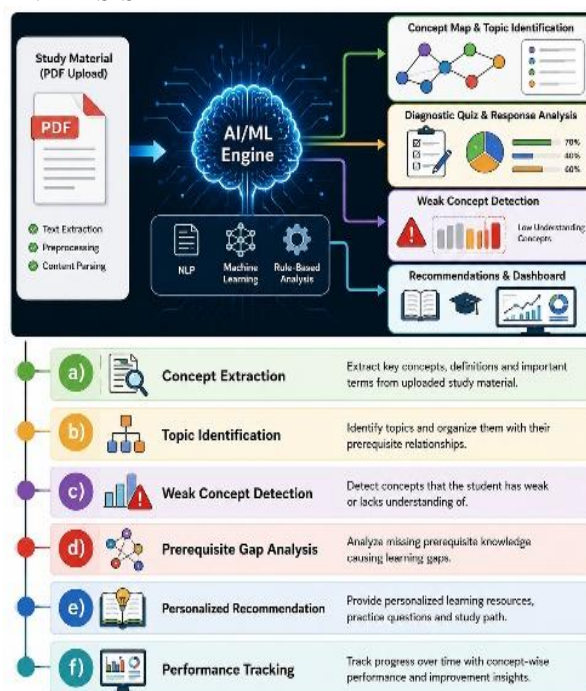
This stage helps students and educators monitor conceptual development effectively.

### I. System Integration

All modules operate within a unified architecture using Flutter and Firebase technologies. Real-time data flows between the frontend interface, concept analysis engine, and database services to ensure efficient processing and adaptive learning support.

The integrated methodology ensures intelligent concept analysis, personalized feedback generation, and continuous learning improvement.

## EXTRACTED FEATURES FOR CROWD SAFETY ANALYSIS



## DATASET DISTRIBUTION FOR CONCEPT GAP DETECTION

Category	Description	Percentage
Strong Concepts	High conceptual understanding and accurate responses	40%
Moderate Understanding	Partial understanding with minor conceptual errors	30%
Weak Concepts	Poor conceptual clarity and incorrect responses	20%
Missing Prerequisite Knowledge	Lack of foundational concepts affecting learning	10%

## IV. PROPOSED SYSTEM

### A. System Overview

The proposed COREGAP system integrates PDF-based study material processing, concept extraction modules, diagnostic assessment generation, and intelligent learning analysis techniques. The system continuously analyzes uploaded educational content, identifies important concepts and prerequisite relationships, evaluates student responses, and generates personalized feedback based on conceptual understanding.

The system uses Natural Language Processing (NLP), rule-based analysis, and machine learning techniques to detect weak concepts and missing prerequisite knowledge. Personalized recommendations are generated to improve conceptual clarity and learning efficiency.

### B. Objectives

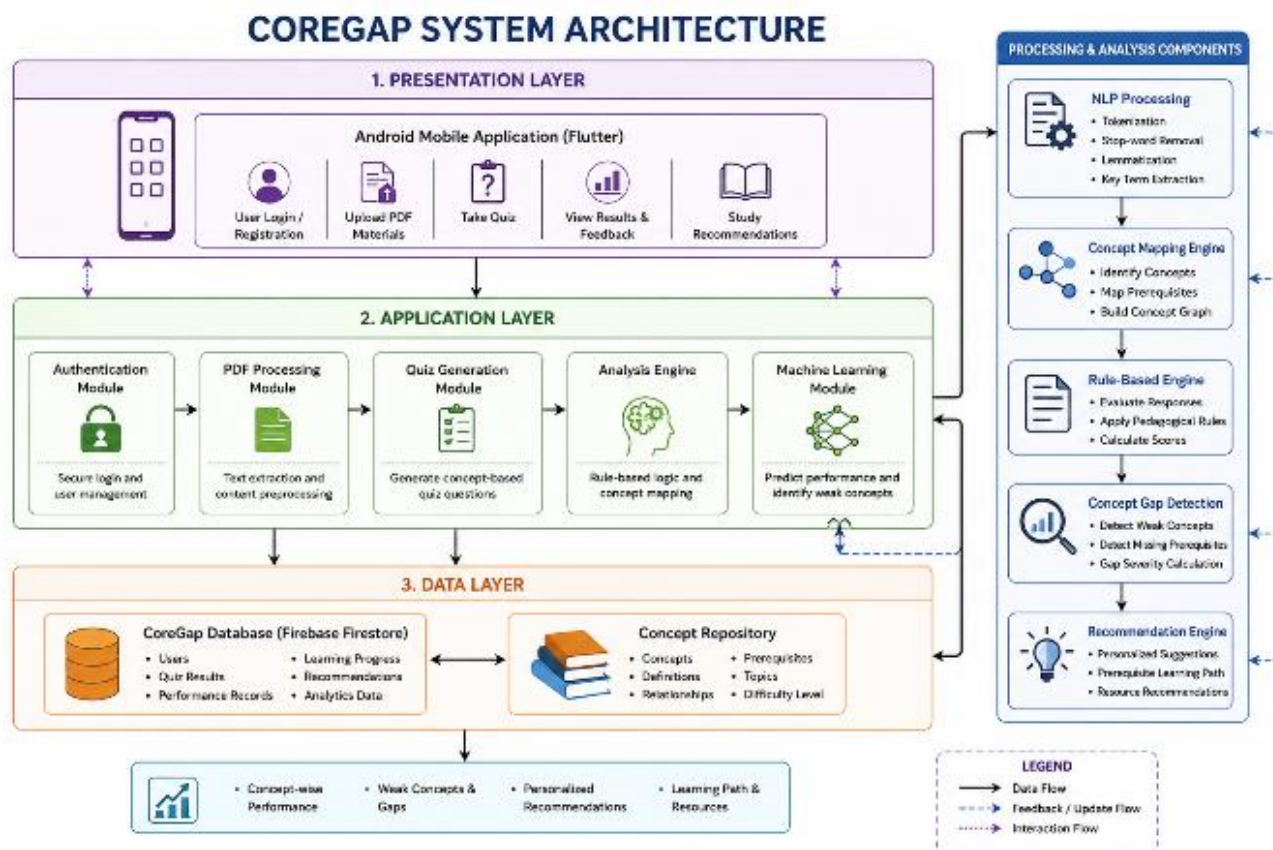
The primary objectives of the system include:

1. Automated concept extraction from uploaded study material
2. Detection of weak concepts and prerequisite knowledge gaps
3. Personalized diagnostic assessment generation
4. Adaptive learning recommendations and performance tracking.

**PROPOSED SYSTEM MODULES**

Module	Function
PDF Upload Module	Upload subject study material
Content Preprocessing	Text extraction and cleaning
Concept Extraction	Identify important concepts and keywords
Concept Mapping	Build prerequisite relationships between topics
Quiz Generation	Generate concept-based diagnostic assessments
Rule-Based & ML Analysis	Analyze student responses and learning patterns
Weak Concept Detection	Identify low-understanding concepts
Recommendation System	Provide personalized learning suggestions

**SYSTEM ARCHITECTURE**



**VI. IMPLEMENTATION AND EXPERIMENTAL OBSERVATIONS**

The COREGAP system was tested using multiple subject PDF documents and concept-based diagnostic assessments generated from uploaded study materials. The system performance was evaluated based on concept extraction accuracy, weak concept detection, prerequisite gap identification, and personalized recommendation generation.

Results indicate that the proposed system effectively identifies conceptual weaknesses and missing prerequisite knowledge among students. Compared to traditional examination-based evaluation methods, the system provides deeper insights into student learning behavior and conceptual understanding.

**Key observations include:**

- Accurate identification of weak and strong concepts

based on quiz performance

- Reliable detection of missing prerequisite knowledge affecting advanced topic understanding
- Personalized feedback and learning recommendations generated in real time
- Improved concept-wise performance tracking and learning analysis
- Reduced manual effort for educators in monitoring conceptual progress
- Enhanced learning efficiency through adaptive and targeted revision support

**VIII. CONCLUSION**

The COREGAP system represents a significant advancement in intelligent educational technology and personalized learning. By combining concept mapping, machine learning, rule-based analysis, and diagnostic assessments, the proposed system enables accurate identification of weak concepts and missing prerequisite

knowledge. The system enhances conceptual understanding and supports adaptive learning through personalized feedback and recommendation mechanisms.

As digital learning environments continue to evolve, intelligent concept gap detection systems will play an important role in improving educational quality and student performance. The adoption of AI-driven personalized learning systems can improve learning efficiency, reduce dependency on rote memorization, and support the development of smarter and more effective educational platforms.

## **IX. REFERENCE**

- [1] C. Liang, J. Ye, Z. Wu, B. Pursel and C. L. Giles, "Recovering Concept Prerequisite Relations from University Course Dependencies," in *Proc. AAAI Conf. Artificial Intelligence*, 2017, pp. 4786–4791.
- [2] C. Liang, Z. Wu, W. Huang and C. L. Giles, "Measuring Prerequisite Relations among Concepts," in *Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP)*, Lisbon, Portugal, 2015, pp. 1668–1674.
- [3] L. Pan, C. Li, J. Li, J. Tang and P. S. Yu, "Prerequisite Relation Learning for Concepts in MOOCs," in *Proc. 55th Annual Meeting of the Association for Computational Linguistics (ACL)*, Vancouver, Canada, 2017, pp. 1447–1456.
- [4] S. Roy, M. Madhyastha, S. Lawrence and V. Rajan, "Inferring Concept Prerequisite Relations from Online Educational Resources," in *Proc. IEEE International Conference on Data Mining Workshops (ICDMW)*, Singapore, 2018, pp. 349–356.
- [5] H. Liu, W. Ma, Y. Yang and J. Carbonell, "Learning Concept Graphs from Online Educational Data," *IEEE Transactions on Learning Technologies*, vol. 10, no. 3, pp. 282–295, Jul.–Sep. 2017.
- [6] C. Liang, J. Ye, H. Zhao, B. Pursel and C. L. Giles, "Active Learning of Strict Partial Orders: A Case Study on Concept Prerequisite Relations," in *Proc. IEEE International Conference on Data Mining (ICDM)*, 2018, pp. 1125–1130.
- [7] X. Qu, X. Shang and Y. Zhang, "Concept Prerequisite Relation Prediction Using Graph Neural Networks," in *Proc. IEEE International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, 2023, pp. 145–152.