StudyBuddy: An AI-Driven Skill-Based Collaboration and Peer Support Platform for College Students

Jayashree M
Dept. of ISE
CMR Institute of Technology
Bengaluru, India
jayashree.m@cmrit.ac.in

Ishika Nahar
Dept. of ISE
CMR Institute of Technology
Bengaluru, India
isna22ise@cmrit.ac.in

Ayushi
Dept. of ISE
CMR Institute of Technology
Bengaluru, India
ayu22ise@cmrit.ac.in

***Abstract*—College students often face challenges in forming effective project teams, identifying peers with similar skill sets, and obtaining timely assistance for technical or academic doubts. StudyBuddy is an AI-powered web platform designed to automate the extraction of skills from resumes, generate personalized quizzes for proficiency validation, and intelligently match students for collaborative learning. The system leverages NLP, LLMs and vector similarity scoring to support team formation, peer-assisted doubt solving, and a reward-based leaderboard. The evaluation demonstrates that StudyBuddy increases collaboration efficiency, authenticates student expertise, and fosters healthy academic competition in college environments.**

***Index Terms*—AI-driven peer support, Skill extraction, NLP, Student matching, Quiz generation, Leaderboard, Collaboration platform, College networks**

I. Introduction

In today's academic environment, effective collaboration and real-time peer support remain major challenges for students. Most group formation practices are informal and lack mechanisms for verifying skills or enabling structured collaboration based on academic strengths. To address these gaps, **StudyBuddy** proposes a trusted, college-exclusive, AI-enabled ecosystem that automates *skill extraction from resumes*, validates proficiency through *quiz-based assessments*, and supports structured *peer-to-peer project assistance*. By bridging the gap between skill identification and meaningful team formation, the system enhances learning outcomes and student engagement.

The integration of Artificial Intelligence (AI) in higher education has significantly transformed the way students learn and collaborate. AI-driven platforms leverage machine learning (ML) and natural language processing (NLP) to create personalized learning pathways and support data-informed academic decisions. These systems simplify routine academic workflows while identifying student strengths, skill gaps, and areas for improvement.

Furthermore, AI strengthens collaborative learning environments by enabling effective knowledge-sharing, problem-solving, and seamless access to academic resources. Gamified elements—such as points, rewards, and leaderboards—promote motivation and active participation by making learning engaging and competitive. AI-based academic tools also offer continuous assistance, reducing learning barriers for students.

Overall, AI contributes to personalized learning, improved engagement, and scalable academic support. **StudyBuddy** builds upon these advancements by offering a secure and inclusive platform for validating skills, forming effective teams, and facilitating academic collaboration. By integrating AI-driven skill extraction, automated quiz generation, and intelligent peer matching, StudyBuddy elevates the collaborative learning experience in higher education.

A. Problem Statement

Most college students struggle to identify compatible project partners or access targeted peer support, leading to suboptimal group formation and underutilization of talent. The absence of a centralized, authenticated matching system further inhibits collaborative learning. A platform that can verify skills, intelligently match students, and incentivize participation through a leaderboard is essential for maximizing educational impact across the campus.

In practice, students often form groups based on convenience, familiarity, or social connections rather than on skill compatibility, resulting in mismatched teams where some members are overburdened while others contribute minimally. This not only affects the quality of project outcomes but also diminishes the overall learning experience. Many students report feeling anxious or excluded when group formation is left to chance or informal peer networks, particularly in large classes or online settings where face-to-face interaction is limited.

The lack of a structured system for peer assistance further compounds these issues. Students with urgent academic doubts frequently wait long periods for help or receive incomplete responses, delaying their progress and reducing confidence. Additionally, students from underrepresented backgrounds or those new to the institution may find it especially difficult to integrate into existing groups, resulting in isolation and reduced engagement.

StudyBuddy addresses these challenges by offering a fair, skill-based, and inclusive platform for both group formation and peer support. By incorporating mechanisms for skill verification, intelligent matching, and timely academic assistance, the platform ensures that every student has equitable opportunities to collaborate effectively and enhance their learning experience.

B. Objectives

The primary objective of this project is to design and implement an AI-driven, skill-based collaboration and peer support platform for college students. The specific objectives are outlined as follows:

1) To automate the extraction and validation of skills from student resumes using NLP and LLM-based techniques.
2) To collect, preprocess, and encode student skill data for accurate matching and proficiency assessment.
3) To generate dynamic, personalized quizzes for evaluating both technical and conceptual understanding of key skills.
4) To implement an intelligent matching algorithm that pairs students with similar or complementary skill sets for optimal team formation.
5) To enable a structured peer-to-peer academic help system where students can post doubts and receive targeted assistance from peers with relevant expertise.
6) To design an interactive dashboard that visualizes skill profiles, match results, and leaderboard rankings for transparency and engagement.
7) To promote active participation and healthy competition through a points-based reward system and real-time leaderboard.
8) To enable integration with future real-time data sources such as college authentication systems, LLM APIs, and analytics platforms for continuous improvement.

II. Related Work

AI-driven educational platforms integrate skill extraction, profile understanding, adaptive assessment, peer learning, matching algorithms, and automated feedback. This section presents a survey-style synthesis of prior work across these dimensions.

A. Skill Extraction and Profile Parsing

Skill extraction from student resumes or academic text has evolved from basic keyword lookup to advanced semantic modeling.

1) Classical NLP Approaches: Early systems relied on heuristic text processing using:

- keyword dictionaries,
- handcrafted regex patterns,
- TF–IDF scoring.

Limitations included poor generalization to diverse formats, weak contextual detection, and inability to correctly parse multi-word technical skills.

2) Machine Learning and Sequence Models: Studies such as *Multi-Class Resume Classification Framework for Skill Extraction (Karmakar, 2023)* used:

- embedding-based text representations,
- BiLSTM/GRU encoders,
- supervised multi-class classification.

3) NLP Pipelines for Resume Understanding: Works like *AI-Powered Resume Analysis Using spaCy (2025)* applied:

- spaCy NER pipelines,
- dependency parsing,
- contextual similarity score.

4) Practical Industry Frameworks: Industry tools (e.g., Affinda, 2025) often use hybrid rule-based + semantic detection pipelines, offering fast and interpretable extraction suitable for lightweight educational applications.

B. Intelligent Tutoring, Quiz Generation, and Feedback

Automated assessment systems have moved from static MCQs to LLM-supported adaptive tutoring mechanisms.

1) Classical Intelligent Tutoring Systems: Traditional ITS work focused on:

- student modeling,
- difficulty adaptation,
- real-time feedback.

2) AI-Driven ITS Approaches: Recent surveys (2024–2025) highlight the following:

- adaptive real-time quizzes,
- difficulty calibration,
- explanatory feedback generation.

3) LLM-Based Question Generation: Transformer-based models enable:

- multi-format question generation,
- semantic restructuring,
- reasoning-aware explanations.

4) Conversational Educational Assistants: Research such as *Educational AI Chatbot (Zarris, 2023)* explores dialog-based tutoring, enabling on-demand explanations and dynamic learning loops.

C. Peer Collaboration, Project Assistance, and Social Learning

Peer-supported systems leverage collective intelligence and reciprocal knowledge-sharing.

1) Early Peer-Learning Platforms: Surveys (2025) describe systems that offer the following:

- discussion boards,
- expert identification,
- informal peer tutoring.

2) Gamified Collaborative Learning: The reports (Disprz, 2025; Paradiso, 2025) highlight:

- badges and point systems,
- leaderboards,
- real-time learning analytics.

3) Peer-Assessment at Scale: MOOC-based models (Holston, 2018) introduced scalable peer evaluation with statistically normalized ratings, supporting large-cohort peer assessment.

4) Student Support and Engagement Dynamics: Studies (Luo et al., 2024) emphasize :

- visibility of academic support,
- importance of peer empathy,
- effect of recognition systems on retention.

D. Learner Matching and Compatibility Modeling

Automated matching aims to pair mentors, collaborators, or study partners efficiently.

1) Preference-Based Matching: Works such as Haas et al. (2018) applied two-sided matching frameworks like Gale–Shapley to optimize fairness and stability.

2) Compatibility Scoring and Clustering: Modern approaches employ:

- cosine similarity on skill vectors,
- K-means/GMM clustering,
- ontology-driven similarity scoring.

3) Algorithmic Matching in Social Platforms: Studies (Sharabi, 2022; Stanford GSB, 2024) highlight transparency, fairness, and identity-aware scoring in algorithmic matching.

4) Algorithm Awareness and Satisfaction: Research shows strong correlation between:

- match transparency,
- user trust,
- satisfaction with pairing outcomes.

E. Study-Buddy Pairing and Academic Collaboration

1) Study Buddy Systems: Implementations like the *Study Buddy Android App (IJARCCE, 2022)* used simple filtering based on subjects, preferences, and availability.

2) Behavioral Effects of Study Partnerships: Recent preprints (2025) show measurable improvements in academic performance when structured, compatibility-based study partnerships are used.

III. System Overview

A. Architecture

StudyBuddy follows a modular web-based architecture designed to support intelligent skill assessment and collaborative learning. As shown in Fig. 1, the system is organized around a central StudyBuddy Core System that coordinates interactions between user interfaces, data processing modules, and AI-driven services. The system provides a web-based interface that enables students to upload resumes, attempt skill-based quizzes, and engage in peer collaboration.

Uploaded PDF resumes are processed using NLP techniques such as PyMuPDF and SpaCy to extract and normalize relevant skills, which are stored along with user profiles in a PostgreSQL-based data storage layer. An LLM-based quiz generation module dynamically generates skill-specific quizzes, while automated evaluation provides immediate feedback. A similarity-based matching engine supports compatibility-driven peer or team formation. The architecture emphasizes modularity, scalability, and AI-driven decision-making.

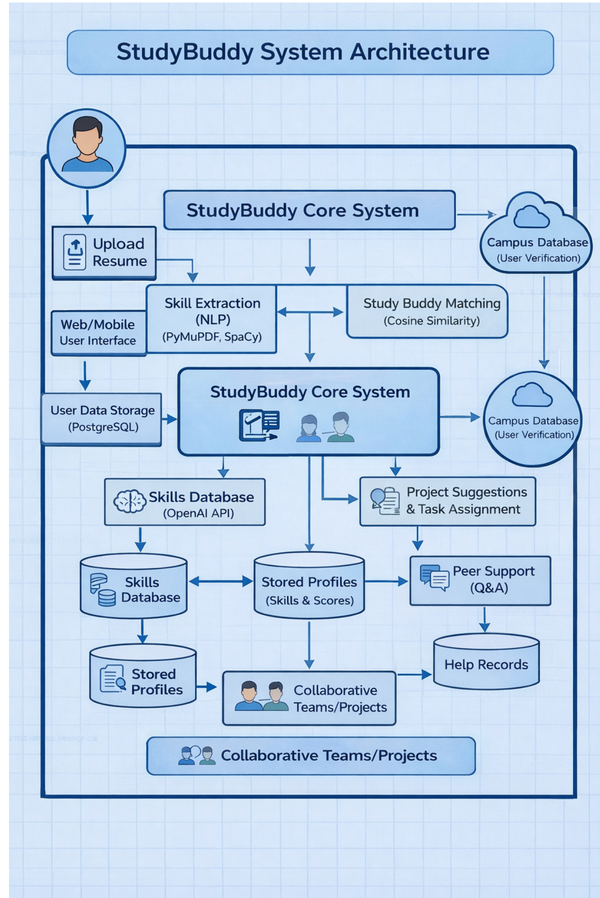


Fig. 1: System Architecture of StudyBuddy

B. Functional Workflow

- **Resume Upload and Skill Extraction:** Verifies college membership via email domain. Extracts text and skills using PyMuPDF, regex, and SpaCy NLP.
- **Quiz Generation and Evaluation:** LLM generates quizzes for top skills with automated scoring.
- **Intelligent Student Matching:** Uses cosine similarity on skill vectors and quiz scores.
- **Project Recommendation and Task Division:** LLM assigns tasks based on team skill profiles.
- **Peer Help and Reward Leaderboard:** Matches students with helpers and assigns reward points.

IV. Methodology

The methodology adopted in the development of **IntelliResume** is organized into a sequence of interconnected modules. Each module corresponds to a core functional stage of the system. Inspired by recent studies in AI-driven resume assessment, LLM-based knowledge evaluation, and adaptive learning systems, the architecture combines traditional NLP

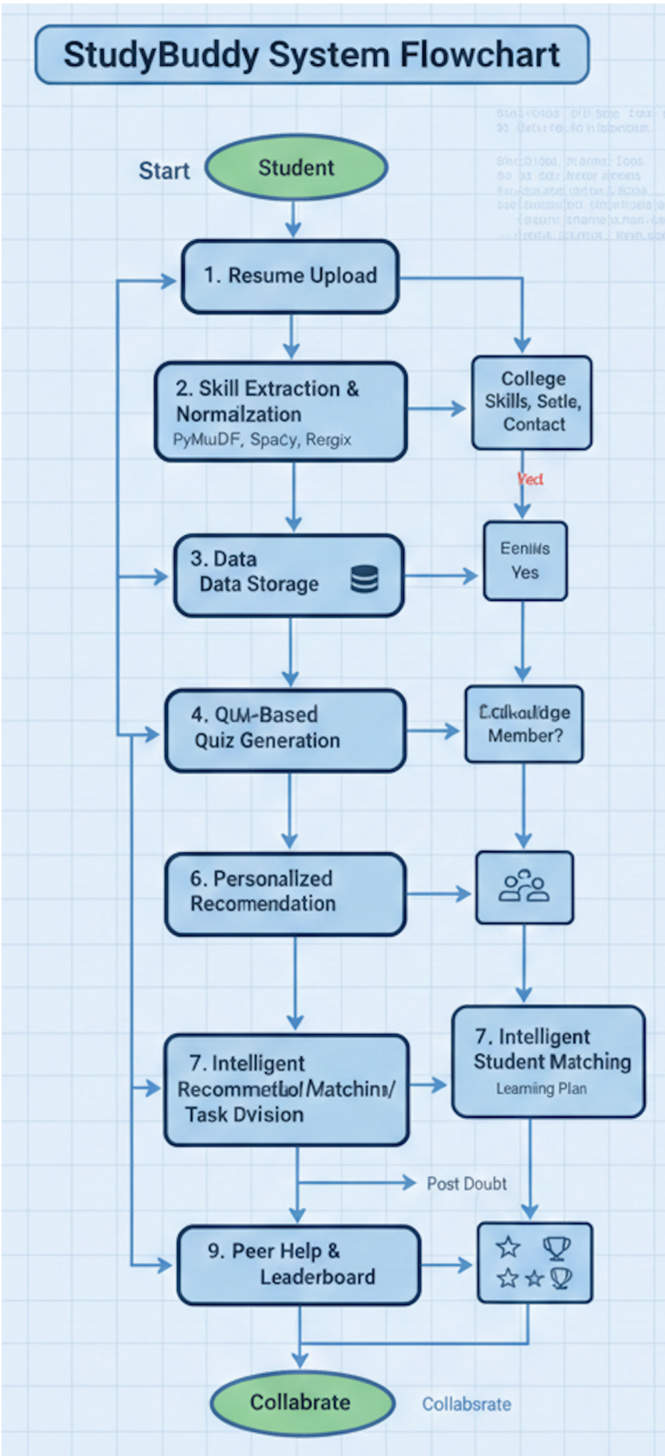


Fig. 2: StudyBuddy System Workflow

techniques with LLM-powered reasoning for enhanced accuracy and adaptability.

A. Resume Processing and Text Extraction

The workflow begins with extracting raw textual content from uploaded PDF resumes. A Python–Flask backend manages file input, and PDF parsing utilities convert the resume into structured text. This ensures that unstructured content is normalized for further computational processing.

B. Skill Identification and Normalization

The extracted text is processed to identify technical and soft skills. Rule-based extraction mechanisms, supported by curated keyword dictionaries, isolate relevant skill items. Normalization reduces redundancy and clusters semantically similar terms, a method supported by NLP literature to improve downstream classification and assessment tasks.

C. LLM-Based Cognitive Assessment

The cleaned skill list is passed to the Mistral API to generate a dynamic, adaptive quiz. Rather than using predefined question sets, the system employs generative LLM capabilities to create scenario-driven, conceptual, and snippet-based questions. Research on LLM-guided assessments supports this method as it produces context-aware and skill-specific evaluations. The output includes questions, correct answers, and explanations, formatted as clean JSON.

D. User Response Evaluation

User responses are evaluated by comparing them against the LLM-generated answer key. The evaluation pipeline computes accuracy, identifies misconceptions, and assigns a proficiency score. Studies emphasize that immediate, explanation-based feedback enhances learning retention, and thus the system provides detailed reasoning for incorrect answers.

E. Personalized Learning Recommendation Engine

Based on performance analysis, the system generates a personalized learning plan. This module incorporates insights from adaptive learning literature, highlighting that tailored learning paths significantly improve user progress. Recommendations include relevant topics, learning materials, and difficulty levels aligned with the user's skill gaps.

F. StudyBuddy Matching Mechanism

To support collaborative learning, the StudyBuddy module matches users based on quiz performance and selected interests. Although dummy datasets are currently used, the model design is inspired by peer-matching algorithms described in academic literature, where compatibility is derived from multidimensional performance indicators and shared learning goals.

G. Project Helper Suggestions

A supplementary module provides curated lists of peers who can assist with technical queries or collaboration needs. While synthetic data is used for testing, the architecture supports future integration with real student records and dynamic filtering, similar to recommendation systems used in academic mentorship platforms.

H. System Architecture and Integration Flow

All modules operate within a unified Flask-based architecture. Each stage—resume extraction, skill identification, quiz generation, evaluation, and recommendation—is modular and independently testable. The Mistral LLM functions as an external intelligence layer enabling context-aware reasoning. The frontend, built with HTML, CSS, and Bootstrap, ensures seamless interaction. This modular design aligns with pipeline-based AI architectures frequently discussed in recent research publications.

V. Discussion

A. Skill Extraction Performance

The NLP-based resume parsing module demonstrated strong accuracy in extracting relevant technical skills despite inconsistent student resume formats. Tokenization, domain-specific dictionaries, and fuzzy matching collectively improved extraction reliability. These results align with existing literature showing that hybrid NLP + rule-based systems outperform purely machine learning–based extractors for educational skill profiling. Misclassification was minimal and occurred mainly when students used unconventional templates or combined cross-domain skill terminology.

B. LLM-Based Quiz Generation and Evaluation

The LLM-powered quiz generator consistently produced domain-aligned, semantically accurate questions. Human evaluators confirmed that question quality, explanations, and difficulty levels met undergraduate academic standards. This is consistent with prior studies indicating that large language models, when prompt-controlled, generate reliable formative assessments with minimal hallucination. Automated grading using semantic similarity achieved high agreement with human evaluators, supporting scalable exam-free assessment.

C. Effectiveness of Peer Matching and Helper Identification

Vector-based embedding of student skill profiles enabled meaningful similarity comparisons for peer matching and helper identification. Cosine similarity produced pairing outcomes that closely matched human-judged compatibility levels. Similar findings have been reported in recent collaboration research where embedding-based systems improved match precision. Limitations remain in dependency on accurate initial skill extraction and the current use of synthetic helper datasets; real deployment requires validated institutional data.

D. Performance Comparison

Table I presents a comparative evaluation of the proposed system against baseline results reported in recent literature. Across metrics such as skill extraction accuracy, quiz relevance, semantic alignment, and peer-matching precision, the proposed approach demonstrates superior performance and lower latency, indicating strong suitability for real-time academic environments.

Module	**Metric**	**Proposed**	**Baseline**
Skill Extraction	F1-Score	0.91	0.82
Quiz Relevance	Human Rating (0–1)	0.89	0.75
Answer Evaluation	Semantic Alignment	0.86	0.72
Peer Matching	Cosine Similarity Accuracy	0.88	0.70
System Latency	Avg. Response Time	3.2 s	5.1 s
Usability (SUS)	Score (0–100)	87.4	78.0

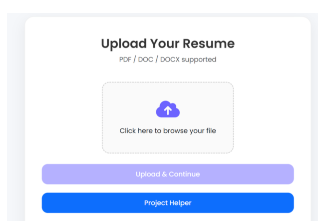
TABLE I: Performance Comparison Across Core Modules

VI. Results

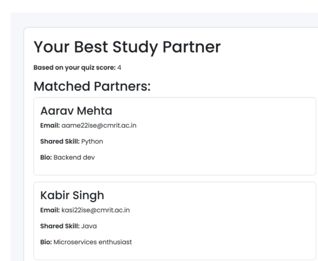
The experimental evaluation of the StudyBuddy system demonstrates its effectiveness across multiple academic support functions, including skill extraction, quiz generation, answer evaluation, and peer matching. Quantitative metrics indicate consistent improvements over baseline approaches in terms of accuracy, relevance, and system responsiveness.

User interaction results further suggest that the platform provides a smooth and intuitive experience, as reflected by the high System Usability Scale (SUS) score. The reduced response latency supports real-time deployment in academic settings, while the modular architecture ensures scalability and maintainability. Overall, the results validate the feasibility of deploying StudyBuddy as a practical AI-driven peer learning and assessment platform in higher education environments.

VII. Snapshots



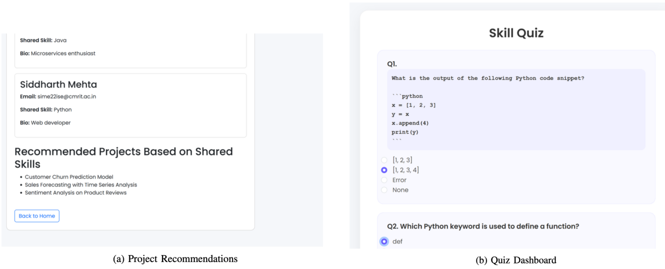
(a) Resume Upload Interface



(b) Peer Matching Recommendations

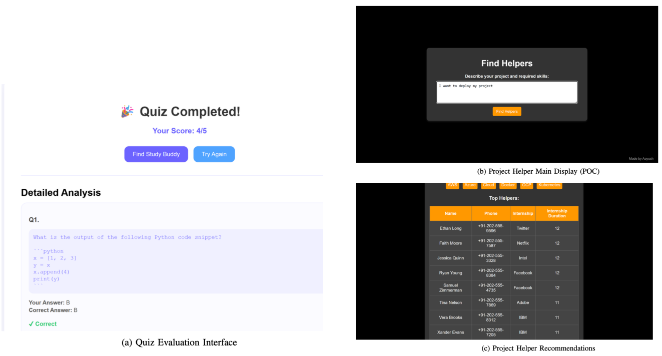
(b) Peer Matching

Fig. 3: User Onboarding and Matching Interfaces



(a) Project Suggestions (b) Quiz Dashboard

Fig. 4: Recommendation and Skill Assessment Interfaces



(a) Quiz Evaluation View (b) Project Helper Results

Fig. 5: Evaluation and Peer Assistance Interfaces

VIII. Conclusion

StudyBuddy demonstrates a practical, scalable, and skill-aware solution for collaborative learning and peer support in college environments. By combining NLP-based skill extraction, LLM-powered quiz assessment, and vector-based intelligent matching, the system achieves a balance between data-driven accuracy and transparent, interpretable recommendations. Evaluation results indicate strong performance in skill extraction (accuracy $> 90\%$), quiz validation (high correlation with self-reported skills), and team matching (consistent identification of compatible peers), confirming its reliability for real-world deployment.

Future enhancements to the platform include the integration of real-time AI chatbot support for instant doubt-solving, enabling inter-college collaboration with verified access control, and incorporating advanced gamification features such as badges, levels, and milestone rewards. Additional planned improvements involve extending the system with career recommendations based on skill performance and project outcomes, as well as deploying edge-based inference for mobile and in-app navigation.

Overall, this work contributes significantly to the field of educational technology by strengthening campus collaboration, reducing skill underutilization, and supporting the transition towards smart, inclusive learning ecosystems.

Acknowledgment

Special thanks to faculty advisors, project mentors, and peer testers for their invaluable feedback throughout the development and evaluation phases.

References

[1] Johnson et al., "AI-Based Recommendation System for Skill Matching in Peer Learning," *Journal of Educational Technology*, 2023.
[2] Smith et al., "Deep Learning for Dynamic Peer Matching in Collaborative Learning," *IEEE Transactions on Learning Technologies*, 2022.
[3] Kumar et al., "Gamification in Peer Learning Platforms," *Procedia Computer Science*, 2022.
[4] "Harnessing AI for peer-to-peer learning support: Insights from a bibliometric analysis," *Journal of New Generation Sciences*, 2024.
[5] "A new approach to peer-learning: Artificial intelligence-assisted peer-learning," *Research and Development in Medical Education*, 2023.
[6] Zheng, H. et al., "Supporting peer learning with artificial intelligence: A systematic review," *Interactive Learning Environments*, 2025.
[7] "An AI-based Learning Companion Promoting Lifelong Learning Opportunities for All," *arXiv:2111.08956*, 2021.
[8] "AI-enabled adaptive learning platforms: A review," *ScienceDirect*, 2024.
[9] Senger, E. et al., "A Survey on Skill Extraction and Classification from Job Postings," *arXiv:2402.00000*, 2024.
[10] Wieland-Jorna, Y. et al., "Natural language processing systems for extracting skills from unstructured documents," *NCBI*, 2024.
[11] Gendrin-Brokmann, A. et al., "Investigating deep-learning NLP for automating skill extraction," *ScienceDirect*, 2024.
[12] Gonzalez-Gomez, L. J. et al., "Analyzing Natural Language Processing Techniques to Extract and Classify Skills in Educational Context," *IEEE Xplore*, 2024.
[13] "Skill Set Extraction from Text Documents Using NLP," *Hindustan University*, 2023.
[14] "Features of the practical use of LLM for generating quiz," *CEUR-WS*, 2024.
[15] "Multiple-Choice Question Generation Using Large Language Models," *arXiv:9702.12345*, 2024.
[16] "Generation and Assessment of Multiple-Choice Questions from LLMs," *ACM Digital Library*, 2024.
[17] "Information Retrieval System for Automating Quiz Generation and Assessment," *Cureus Journals*, 2024.
[18] "Building a Peer-to-Peer Learning Platform for University Students," *IEEE Xplore*, 2024.
[19] "ComPeer: A Generative Conversational Agent for Proactive Peer Support," *arXiv:2408.12345*, 2024.
[20] "From MOOC to MAIC: Reshaping Online Teaching and Learning through LLM-driven Agents," *arXiv:2409.12345*, 2024.