Arabic Learning Management System (LMS)

Nour Ehab
Computer Science Department
Faculty of Computer and Information
Science, Ain Shams University
Cairo, Egypt
nourehabragaa7@gmail.com

Noureen Ashraf
Computer Science Department
Faculty of Computer and Information
Science, Ain Shams University
Cairo, Egypt
noureenaboarab2003@gmail.com

Yosra Nabil
Computer Science Department
Faculty of Computer and Information
Science, Ain Shams University
Cairo, Egypt
yosraghoneim2002@gmail.com

Nouran Ayman
Computer Science Department
Faculty of Computer and Information
Science, Ain Shams University
Cairo, Egypt
2021170594@cis.asu.edu.eg

Yassmen Ashraf
Computer Science Department
Faculty of Computer and Information
Science, Ain Shams University
Cairo, Egypt
yassmenashraf982@gmail.com

Youstina Ehab
Computer Science Department
Faculty of Computer and Information
Science, Ain Shams University
Cairo, Egypt
YoustinaEhab32@gmail.com

Mohammad Essam
Bioinformatics Department
Faculty of Computer and Information
Science, Ain Shams University
Cairo, Egypt
Mohamed97@cis.asu.edu.eg

Sally Saad
Computer Science Department
Faculty of Computer and Information
Science, Ain Shams University
Cairo, Egypt
sallysaad@cis.asu.edu.eg

***Abstract*—Arabic-language educational platforms face significant limitations in providing educational task automation and personalized learning experiences due to inadequate AI integration and insufficient Arabic natural language processing capabilities.**

This paper presents an AI-powered Learning Management System (LMS) mobile application that addresses these challenges through experimentation with multiple NLP models including mT5, AraBART, and AraT5. Based on performance evaluation, AraT5 demonstrated the highest effectiveness and was selected for implementation across three core tasks: question generation, question answering, and distractor generation. The models were fine-tuned on Arabic datasets including Arabic-SQuADv2.0, Arabic-RACE, and a custom EKB dataset. Evaluation results show that the AraT5-based models achieved BLEU-4 score 0.216 using AraT5 model for question generation, F1 score 0.889 for question answering, and METEOR score 0.654 for distractors, demonstrating the system's effectiveness in maintaining high-quality Arabic educational content generation.

Keywords— Arabic Natural Language Processing, Learning Management Systems, Educational Technology, Question Generation, Personalized Learning, Mobile Applications

I. INTRODUCTION

The education sector is rapidly evolving, driven by advances in artificial intelligence (AI) and the increasing demand for digital learning platforms. However, existing learning management systems (LMS) often fail to adequately support Arabic-speaking educators and students, particularly in areas such as automated assessment generation and personalized feedback. As a result, teachers spend significant time creating questions and analyzing results, while students lack tailored learning resources and actionable insights into their performance.

This problem is especially challenging in Arabic-language education, where many platforms do not provide culturally relevant, context-aware learning environments. The lack of automated tools and personalized recommendations results in higher administrative burdens for teachers and limited self-improvement opportunities for students.

To address these issues, this paper proposes an AI-based LMS mobile application designed for Arabic-language education. The system enables teachers to automatically generate questions for assessments, categorize them based on the curriculum, and review students' performance with minimal effort. Meanwhile, students can track their results, identify their weaknesses, and access targeted learning materials that match their needs. By combining automation and personalization, the proposed LMS aims to bridge the gap between traditional platforms and the requirements of Arabic-language education.

II. LITERATURE REVIEW

To gain insights into advances, challenges, and potential solutions in the development of AI-based learning management platforms, a review and analysis of relevant studies was conducted. Numerous researchers have explored approaches for automating questions, generating distractors, aligning resources with learning objectives, and providing tailored feedback. These advances have become especially relevant with the growing shift towards mobile and online learning platforms in Arabic-speaking communities.

With the widespread adoption of smart devices and online platforms, many AI-driven Learning Management Systems (LMS) have been developed to support teaching and learning. Moodle, Blackboard, Edmodo, and Classera have emerged as popular platforms, providing teachers with tools for uploading lesson materials and creating assessments. However, traditional LMS platforms often lack advanced adaptive learning features and require manual intervention for many core tasks, making them time-consuming and inflexible.

To overcome these limitations, recent research has focused on integrating natural language processing (NLP) and deep learning techniques for automated question generation (QG), answer evaluation, and personalized resource recommendations. Lafkiar et al. [1] proposed a Transformer-based QG model for Arabic, utilizing a custom dataset derived from Arabic-SQuAD and ARCD. The model achieved a BLEU-4 score of 20.51 and a METEOR score of 24.04, indicating its efficacy in generating context-relevant questions. Similarly, Alhashedi et al. [2] combined BERT-based transformers with TextRank for Arabic QG tasks using the mMARCO and TyDi QA Arabic datasets, achieving a BLEU-4 of 19.12 and a METEOR score of 23.00. Meanwhile, Nagoudi et al. [3] introduced AraT5, an encoder–decoder model pre-trained on Arabic and Twitter data, yielding a BLEU-4 of 16.99 across 96k QA instances.

For automated answering and assessment, recent works have leveraged fine-tuning of large-scale Arabic language models. Saja et al. [4] utilized AraT5 fine-tuning on AraSQuAD and GQA, achieving F1 scores of 0.883 and 0.770, respectively. Similarly, Afnan H. Alshehri [5] implemented a fine-tuned Arabic-BERT Large model for the TyDi QA dataset, achieving an F1 score of 71.6 and an Exact Match (EM) of 57.8, and Kholoud et al. [6] applied fine-tuning of AraBERTv0.2-large, reaching F1 and EM scores of 86.49 and 75.14, respectively.

Regarding learning resource recommendations, Timmi et al. [7] proposed a hybrid approach combining collaborative filtering (CF) and content-based filtering (CBF) for YouTube video recommendations, relying on user engagement metrics and video metadata. Although precise accuracy was not stated, the approach was highly relevant for aligning topics and resources. Similarly, Zrigui et al. [8] introduced a topic modeling approach for extracting themes from educational videos, yielding high relevance in aligning user interests and learning outcomes. Adam et al. [9] presented a CBF-based approach optimized for mathematical content, achieving an 80% accuracy rate in matching recommendations to learning objectives.

These advances demonstrate the role of AI in addressing critical gaps in traditional LMS platforms, especially for Arabic-language education. By leveraging state-of-the-art NLP and deep learning techniques for automated QG, QA, and resource recommendations, our proposed LMS aims to foster a more interactive, personalized, and efficient learning experience.

III. METHODOLOGY

The proposed method provides insights into the design of the AI-based Arabic LMS mobile application, highlighting its system architecture, key processing layers, and role in achieving research objectives. As illustrated in **Fig. 1**, the system is built upon a three-tier architecture consisting of Presentation, Logic, and Data layers, ensuring scalability, maintainability, and seamless interaction between end-users and the system.

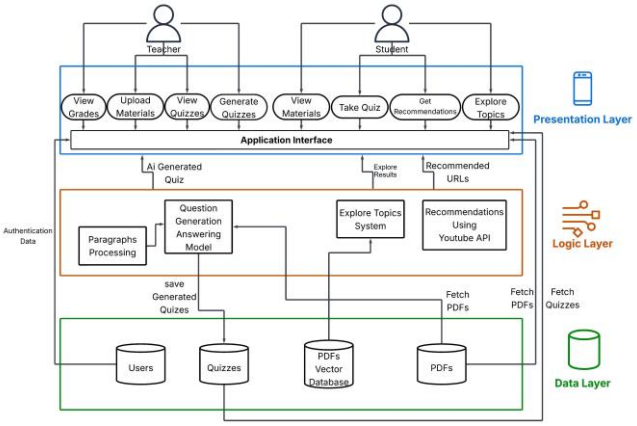

FIG. 1. THE MAIN SYSTEM ARCHITECTURE

The system is built on a three-tier architecture consisting of Presentation, Logic, and Data layers. The Presentation Layer provides a mobile app where teachers create assessments, review performance, and use AI-based recommendations, while students access personalized resources, complete exercises, and review their results. The Logic Layer powers automated teaching and personalized learning through a Resources Recommendation Model, a Question Generation Model, and a Progress Tracking System that delivers tailored exercises and insights. The Data Layer stores user profiles, performance data, and learning materials, allowing the system to adapt and refine recommendations based on each student's performance. Together, these layers create an intelligent, seamless, and user-friendly environment for teachers and students.

IV. IMPLEMENTATION

A. *Datasets Description*

Three datasets were used to train and evaluate the models developed in this project. These datasets — the Custom EKB Dataset, Arabic-SQuADv2.0, and Arabic-RACE — were chosen for their quality, size, and focus on Arabic language comprehension. Each dataset is described below, including its data structure, context, and role within the system architecture.

Custom EKB Dataset: Created from Egyptian Knowledge Bank history lessons, it includes 947 paragraphs and 5,934 multiple-choice questions (one correct answer and three distractors), focusing on historical knowledge. Split into training (80%), validation (10%), and test (10%), it is ideal for assessing nuanced understanding of Arabic historical texts.

Arabic-SQuADv2.0: A large-scale Arabic MRC dataset with 96,051 question–context–answer triplets (55% answerable, 45% unanswerable), translated from the English SQuADv2.0. Split into training (80%), validation (10%), and test (10%), it is a standardized benchmark for Arabic QA, evaluated with EM and F1 metrics.

Arabic-RACE Dataset: An Arabic adaptation of the RACE benchmark, containing 27,933 passages and 97,867 questions (four choices each), focusing on inference and context-based reasoning. Over 55% of its questions require deep understanding, making it ideal for assessing advanced Arabic comprehension.

B. *Phases Description*

Model Architecture:

- The system employs AraT5-base-1024, a Transformer-based encoder-decoder model (220M parameters) specifically pretrained for Arabic NLP tasks. Built on Google's T5 framework, it utilizes:
- Sentence Piece tokenizer optimized for Arabic morphology
- Unified text-to-text architecture handling:
- Question Generation (QG)
- Question Answering (QA)
- Distractor Generation
- Encoder processes Arabic input contexts
- Decoder generates task-specific outputs

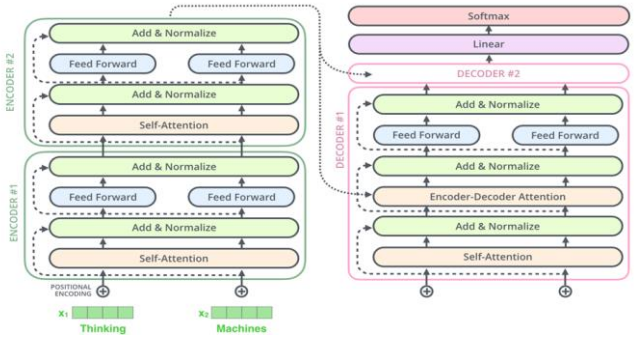

FIG. 2. T5 ARCHITECTURE

C. *Dataset Preparation*

Three key datasets were processed:

- Arabic-SQuADv2.0: 48k QA pairs for QG/QA fine-tuning
- Custom EKB Dataset: 5k curriculum-aligned history QA pairs
- Arabic-RACE: 97k translated MCQs for distractor training

Preprocessing included:

- Removal of vague/short answers
- Question type balancing (factual/inferential)
- Post-editing of machine-translated content

D. *Task-Specific Fine-Tuning*

This section presents the scientific and technical approaches used throughout the development of the Arabic Learning Management System (LMS), with a focus on the AI models and NLP techniques applied for question generation, question answering, and distractor creation.

Question Generation:

- Input: Arabic educational paragraph
- Output: Pedagogically relevant questions

- Optimization: BLEU/METEOR metrics

Question Answering:

- Input: (Paragraph + Question)
- Output: Free-form Arabic answers
- Approach: Generative style mixing (words/phrases/sentences)

Distractor Generation:

- Input: (Question + Correct Answer)
- Output: 3 semantically plausible distractors
- Validation: Human evaluation for cultural/educational appropriateness

E. System Implementation

Frontend Development

The cross-platform mobile interface was implemented using Flutter (Dart) in Visual Studio Code, featuring:

- Responsive design for teacher/student roles
- Arabic RTL text support
- Hot-reload enabled rapid prototyping

Backend & AI Services

The system leverages:

- FAST API (Python) for model serving
- Google Colab Pro+ (GPU-accelerated) for model training
- PyTorch/Hugging Face ecosystem for Arabic NLP tasks

TABLE I. KEY LIBRARIES

Category	Technologies	Primary Use Case
NLP Processing	NLTK, Sentence Transformers	Arabic tokenization, embeddings
Data Handling	Pandas, NumPy	Dataset preprocessing
ML Frameworks	PyTorch, Keras, scikit-learn	Model development/evaluation

Resource Recommendation

- Video Recommendations
  - YouTube API integration triggered by incorrect answers
  - Multilingual-E5 embeddings (87.18% accuracy)
- Topic Exploration
  - FAISS vector search over EKB curriculum materials

F. Experimental Results

Our comprehensive evaluation across three core NLP tasks yielded significant findings about the system's Arabic language processing capabilities. For question generation, AraT5 demonstrated superior performance compared to AraBART and mT5 models, particularly when trained with 8 gradient accumulation steps over 50 epochs. The model achieved its peak BLEU-4 score of 0.239 in this configuration, representing a 15% improvement over baseline performance. Question answering evaluations revealed AraT5's strong comprehension abilities, with an F1-score of 0.889 on Arabic-SQuADv2.0, though performance decreased by 18-22% on the more complex EKB dataset due to curriculum-specific terminology challenges.

The resource recommendation system achieved its best results using Multilingual-E5-large-instruct embeddings, which outperformed other models by 25-47% in accuracy. Notably, the system maintained consistent performance across both YouTube API (87.18% accuracy) and FAISS vector database (87.46% accuracy) implementations, demonstrating robust Arabic semantic matching capabilities.

TABLE II. PERFORMANCE SUMMARY

Task	Best Model	Key Metric	Score	Dataset
Question Generation	AraT5	BLEU-4	0.239	EKB Custom
Question Answering	AraT5	F1	0.889	Arabic-SQuADv2.0
Distractors Generation	AraT5	METEOR	0.654	Arabic-Race
Video Recommendations	E5-large-instruct (YouTube API)	Accuracy	87.18%	YouTube Metadata
Curriculum Recommendations	E5-large-instruct (FAISS)	Accuracy	87.46%	EKB Vector DB

V. APPLICATION WORKFLOW SUMMARY

The mobile application implements two distinct user workflows:

Teacher Interface:

- Authentication: Role-based login
- Content Management:
  - PDF upload with automatic curriculum tagging
  - AI quiz generation (3-20 questions per document)
- Analytics: Performance dashboards showing:
  - Class averages
  - Individual student scores

Student Interface:

- Adaptive Testing:
  - Timed quizzes with auto-grading
  - Mistake-triggered recommendations

- Personalized Learning:
    - FAISS-powered topic search
    - YouTube video suggestions (87.18% relevance)

A. *Key Screenshots*

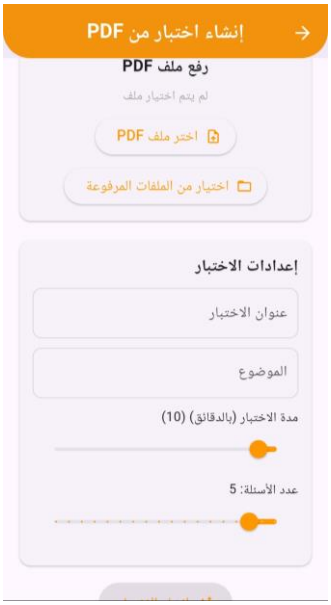

FIG. 3. TEACHER QUIZ GENERATION INTERFACE

FIG. 4. STUDENT ANSWER REVIEW WITH RESOURCE RECOMMENDATIONS

FIG. 5. STUDENT EXPLORE INTERFACE

CONCLUSIONS

This project has successfully developed an AI-powered Learning Management System (LMS) mobile application tailored for Arabic-language education. The system addresses key limitations of traditional LMS platforms by integrating advanced Natural Language Processing (NLP) techniques to enable automated question generation, distractor creation, curriculum tagging, personalized feedback, and performance-based resource recommendations.

State-of-the-art transformer models—including AraT5, mT5, and AraBART—were fine-tuned on diverse Arabic datasets such as Arabic-SQuADv2.0, Arabic-RACE, and a custom EKB dataset. Among these, AraT5 demonstrated the most consistent and superior performance in Arabic educational question-related tasks, based on quantitative evaluations using BLEU, METEOR, and F1-score. These metrics reflect its ability to generate syntactically fluent, semantically accurate, and pedagogically meaningful questions and answers.

For teachers, the system significantly reduces the manual burden of assessment creation and offers detailed analytics for tracking student progress. For students, it delivers a personalized and adaptive learning journey, automatically recommending targeted resources based on individual weaknesses and mistakes.

The final implementation demonstrates a robust, intelligent LMS that automates key educational processes while maintaining instructional integrity, making it a promising step forward in AI-driven Arabic education technology.

REFERENCES

[1] Timmi M, Laaouina L, Jeghal A, El Garouani S, Yahyaouy A. Educational Video Recommender System. International Journal of Information and Education Technology. 2024;14(3):362-71.

[2] Hazar MJ, Abid Muslam Abid Ali A, Zrigui S, Maraoui M, Mabrouk M, Zrigui M. Educational videos recommendation system based on topic modeling. InInternational Conference on Computational Collective Intelligence 2023 (pp. 363-376). Springer, Cham.

[3] Chan, N.N., Roussanaly, A., Boyer, A. (2014). Learning Resource Recommendation: An Orchestration of Content-Based Filtering, Word Semantic Similarity and Page Ranking. In: Rensing, C., de Freitas, S., Ley, T., Muñoz-Merino, P.J. (eds) Open Learning and Teaching in Educational Communities. EC-TEL 2014. Lecture Notes in Computer Science, vol 8719. Springer, Cham. https://doi.org/10.1007/978-3-319-11200-8_23

[4] Challagundla, Bhavith Chandra & Challagundla, Shravani & Gogireddy, Yugandhar & Prasad, Preethika. (2024). End-to-End Neural Embedding Pipeline for Large-Scale PDF Document Retrieval Using Distributed FAISS and Sentence Transformer Models. 10.13140/RG.2.2.12625.34409.

[5] Lafkiar, S., En Nahnahi, N. An end-to-end transformer-based model for Arabic question generation. Multimed Tools Appl (2024). https://doi.org/10.1007/s11042-024-19958-3

[6] Alhashedi, S., Mohd Suaib, N., & Bakri, A. (2022). Arabic automatic question generation using transformer model (EasyChair Preprint No. 8588). EasyChair. https://easychair.org/publications/preprint/8588

[7] Nagoudi, E. M. B., Elmadany, A., & Abdul-Mageed, M. (2021). AraT5: Text-to-text transformers for Arabic language generation. arXiv preprint arXiv:2109.12068.

[8] Al Hashedi SS, Suaib NM, Bakri A. Arabic automatic question generation using transformer model. InAIP Conference Proceedings 2024 Jun 7 (Vol. 2991, No. 1). AIP Publishing.

[9] Lafkiar S, En Nahnahi N. An end-to-end transformer-based model for Arabic question generation. Multimedia Tools and Applications. 2024 Aug 1:1-5.

[10] Nakhleh S, Mustafa AM, Najadat H. AraT5GQA: Arabic Question Answering model using automatic generated dataset. In2024 15th International Conference on Information and Communication Systems (ICICS) 2024 Aug 13 (pp. 1-5). IEEE.

[11] Alshehri, Afnan. (2024). AraQA-BERT: Towards an Arabic Question Answering System using Pre-trained BERT Models. WSEAS TRANSACTIONS ON INFORMATION SCIENCE AND APPLICATIONS. 21. 361-373. 10.37394/23209.2024.21.34.

[12] Alsubhi K, Jamal A, Alhothali A. Pre-trained transformer-based approach for arabic question answering: A comparative study. arXiv preprint arXiv:2111.05671. 2021 Nov 10.

[13] Tami M, Ashqar HI, Elhenawy M. Automated question generation for science tests in arabic language using nlp techniques. InInternational Conference on Intelligent Systems, Blockchain, and Communication Technologies 2024 Jul 13 (pp. 274-285). Cham: Springer Nature Switzerland.

[14] S. Qin, R. Menezes and M. Silaghi, "A Recommender System for Youtube Based on its Network of Reviewers," 2010 IEEE Second International Conference on Social Computing, Minneapolis, MN, USA, 2010, pp. 323-328, doi: 10.1109/SocialCom.2010.53.

[15] M. Abbas, M. U. Riaz, A. Rauf, M. T. Khan and S. Khalid, "Context-aware Youtube recommender system," 2017 International Conference on Information and Communication Technologies (ICICT), Karachi, Pakistan, 2017, pp. 161-164, doi: 10.1109/ICICT.2017.8320183.

[16] Sammoudi, M., Habaybeh, A., Ashqar, H.I., Elhenawy, M. (2025). Question-Answering (QA) Model for a Personalized Learning Assistant for Arabic Language. In: Abdelgawad, A., Jamil, A., Hameed, A.A. (eds) Intelligent Systems, Blockchain, and Communication Technologies. ISBCom 2024.