

LEARNIFY: An NLP-Driven Video Synthesizer for Interactive Learning Using Web Scraping

K.S.G.Vikas Reddy¹, D.Suvarna Lakshmi Manikanteswari Asst.prof², Mohammad Azmatulla³, Dindukurthi Chaturved⁴, Kasinikota Venkat⁵, R. Karthikeya Harshitha⁶

Department of Computer Science Engineering

¹⁻⁶ Sri Vasavi Engineering College, Tadepalligudem , Andhra Pradesh, India

ABSTRACT- The swift escalation in the number of digital learning videos has introduced difficulties such as excessive amounts of information, duplicated material, and learning paths that are not well-structured. Typical tools for creating summaries frequently shrink down individual videos or create quizzes that do not change, but they are not versatile and do not use different sources. To solve this issue, LEARNIFY presents a cutting-edge strategy that blends pertinent information from numerous online videos into a single, brief, and interactive educational experience. Its main strength is a flexible feedback system that measures how well students are learning through quizzes, pinpoints areas where they need more understanding, and remakes content that is easier to understand and customized, either as written material or as newly created videos. By bringing together web data extraction, Natural Language Processing (NLP), and video creation using artificial intelligence, LEARNIFY fosters learning that is tailored, flexible, and productive, all while reducing mental strain and turning simply watching videos into a stimulating journey of self-directed learning.

INDEX TERMS: Customized Education, Flexible Responses, Natural Language Processing (NLP), Video Creation, Web Data Extraction

1. INTRODUCTION

The Digital Age has revolutionized how learning resources are accessed, with the vast diversity of options, such as MOOCs, tutorial videos on YouTube, and other web repositories. The amount of information now available has created cognitive overload because evidence-based learning now requires learners to sift through many videos (often with little difference in content) about the same topics with potentially important differences in quality, depth, and relevance. During this time, learners are being forced to consider what is evidence-based and is relevant to their specific context - something that is not only time-consuming, but ineffective.

Early research was mostly interested in video summarization, in which algorithms produce textual synopses or pull together scenes from one source to produce a summary of key themes. While extremely useful, these efforts are constrained in extracting themes across sources and thinking differently across multiple resources. In a similar vein, recommendation systems proposed related videos, but again, this was just "one at a time," and the conceptual relationship or complementarity of the instructional media was often not assessed.

Personalized learning has emerged as an option because it adjusts the instruction to the learner's needs and progression. However, although many are adaptive, they are still usually anchored to datasets that are still somewhat static (or static but pre-curated), and then there is no pedagogical evolution to keep pace with the expanding learning landscape for online education.

Recent advances in Natural Language Processing (NLP) have made it possible to semantically connect ideas across a variety of sources. Efforts in semantic clustering, cross-document summarization, naming and semiotic recognition in the intent space allow machines to build a knowledge representation of the learning space from dissociated content. Together with the capability to create media driven by AI, these abilities open the door to the creation of educational videos in an automated fashion with new possibilities for dynamic and engaging learning opportunities. Nevertheless, challenges remain, How can systems avoid redundancy while presenting diverse perspectives? How can the generated content ensure factual accuracy and pedagogical soundness? In addition, how can adaptive feedback be built in so that learners not only learn new information but also build their understanding in real time?

To address these concerns, the study presents LEARNIFY, an integrated system that combines web-based data in real time, NLP-based content generation, and adaptive feedback. The objective is to provide personalized education and scalable learning experience while still preserving the richness and heft of educational resources that exist

2. LITERATURE SURVEY

Early work on summarizing educational videos was mainly based on the extraction of text data from the audio and video visuals. The researchers provided approaches to summarizing by detecting important audio and video elements, detecting important frames, and detecting scene transitions that allowed for summarizing videos into shorter videos. The learner could then utilize these shortened elements by seeking out important information they could learn or find without accessing the entire video [1].

Although the short forms using videos had an efficient and strategic approach, they missed the crucial oral explanation that occurred in the lecture or tutorial. To address this gap, audio analysis was used, especially with advancements in user-friendly automatic speech recognition (ASR) and speech-to-text technologies to retain key spoken information from the videos and present it in textual format which complemented the video- visual summaries [2]. Integrative approaches have developed multi layered processing frameworks that aligned the presentation of audio, video, and text for longer sequences of video [3]. These advances facilitated the understanding of difficult topics, as learners could simultaneously track the verbal explanations as well as the visual model of the demonstration.

As the creation of educational videos had exploded, researchers soon realized that summarizing individual videos alone was inadequate. Often, videos on similar topics had overlapping content with some also expanding on the subjects. This spurred the development of a multi-source aggregation process designed to synthesize content from multiple videos to mitigate redundancy and enhance learning [4]. The use of semantic clustering, topic modelling, and graph-based organization of the content around the concepts of the lesson allowed the system to surface content about similar information and develop organized memory representations around the knowledge [5]. In sum, by aggregating content from multiple sources, a summary was created that could organize and synthesize knowledge from multiple perspectives without redundancy.

In parallel, personalized learning powered by artificial intelligence" emerged as a new educational approach. Intelligent tutoring systems identified, on-the-fly, instructional material and assessments, based on the learner's performance. Reviewing each learner's progression through activities and quizzes, the system was able to suggest supplementary learning activities to address gaps in knowledge [6]. However, many of the personalized learning systems remained static in the content library the system used, all while attempts to integrate newer or evolving learning resources were difficult, and not dynamic to the learner's experience [7]. This limitation emphasized the necessity of a real-time adaptation capability in personalized learning platforms.

Subsequent developments were made with knowledge graphs and semantic analysis. Specific techniques, such as entity extraction, mapping relationship, and cross-document co-reference resolution allowed machines to model complex relationships between concepts and synthesize information from multiple sources [8][9][10][11]. Knowledge graphs aided this collaborative visualization by offering a visual and structural representation of related concepts, allowing even novice learners to view connected ideas in context rather than isolated, which assist with content navigation. These techniques also allowed intelligent systems to align differences in semantic conceptualization, various terminology, and terminology presentation across multiple video segments, resulting in a coherent learning experience.

In the last few years, AI-driven multimedia technologies have prompted the automatic creation of educational videos, interactive avatars [12][13]. These systems leverage the combination of assessments and skills-gap analyses, delivering interactive, engaging learning experiences that are dynamic to the progression of the learner in real time [14]. However, in the current environment, most platforms now focus solely on a singularly, evolving concept either in the summarization side or in the personalization side, while none have converged both concepts together into a singular system [15].

The literature shows an unambiguous gap: the need for integrated frameworks incorporating multi-source content aggregation, personalization, knowledge representation, and AI-enhanced multimedia content; all features that could support adaptive, context-aware, and engaging elements of the educational process in ways that address redundancy, learner needs, and utilize the richness of available online materials.

3. EXISTING SYSTEM

- Focusing on the Single Video:** Some texts also summarize a single video at a time (rather than take a broader, multi-video view) by retrieving highlights or key points, which focus is wholly inadequate for allowing a truly comprehensive understanding as learners will not have the benefit of views from additional video sources.
- Redundant Content:** Aggregate systems for multi-video utilization do a poor job of removing redundant or repeated content, such that it can overload learners with repetitive content and disruptively make it hard to get the learner to key concepts.
- Limited Personalization:** Adaptive learning platforms may use some method to evaluate the learner's performance but most do not leverage real-time synthesis of content from video utilization. Consequently, the materials presented to the learner do not address knowledge gaps for the particular learner and do not occur at a pace the learner is comfortable with.

- **Quality of the Data:** Captions, transcripts, and other forms of documentation are often inaccurate and/or incomplete. This inconsistency will not assist NLP summarization, knowledge extraction, nor trustworthiness of the content itself.
- **Overfitting of AI Models:** Most AI models are fit to a small data set, thus not generalizable for other topics, sources of video or for other groups of learners, which detracts from recommendations from such models.
- **Scalability:** Would it take significant computational power and luxurious portions of time to bring together a great deal of video content in real-time, analyze the information and synthesize into valid knowledge for the learner. Many systems in particular won't be able to loosen latency under great time or computational loads.
- **Limited Cross of features:** There are limitations in many systems to summarize multiple sources, produce a knowledge graph and interact the learner (provide a quiz or lesson based on videos), and AI-generated videos in one cohesive system and requires learners to switch from one item to another.
- **Lack of Dynamic Feedback:** In most systems, the content is not automatically modified based on the user's real-time performance or success, which relates to having the ability to target and adapt learning effectively.
- **Need for the Unified Platform:** These limitations reflect a need for a platform that offers an integrated, multi-source summary, adaptive personalization, representation of knowledge, and interactive learning features.

4. PROPOSED SYSTEM:

LEARNIFY aims to tackle these limitations through the integration of AI, NLP, and video synthesis to create an adaptive, personalized, and interactive learning platform.

Some core features of LEARNIFY:

1. Content Collection & Preprocessing:

Educational videos are captured using web scraping, audio is transcribed in text format, cleaned, and preprocessed for applicability.

2. Feature Extraction & Summarization:

Utilizing NLP approaches such as semantic analysis and key-term extraction will recognize key concepts and provide clean, concise, non-redundant summaries with appropriate context.

3. Adaptive Learning & Feedback:

Interactive quizzes are utilized to assess the learner's understanding of the material. Based on correct or incorrect, the LEARNIFY system will offer feedback catered to the learner, and regenerate their content in easier language, and further summarize what has already been summarized.

4. Knowledge Graph Creation:

Concepts and how they relate to each other are displayed to the user through a dynamic knowledge graph to allow exploration of the material in context.

5. AI Video Generation:

The summary generated will be converted into an AI-generated video incorporating text-to-speech, and visual avatars to enhance multimodal engagement with the instructional material.

6. Scalability:

The system can accept large datasets, and process them in a timely manner, across multiple subjects, and make improvements and adaptations in real-time.

7. Continuous Improvements:

The user's data and performance levels will feed back into the system over time for continual improvement of accuracy and adaptability.

8. User Interface:

A simple, intuitive interface allows users to search topics, access summaries, take quizzes, and watch AI-generated videos seamlessly.

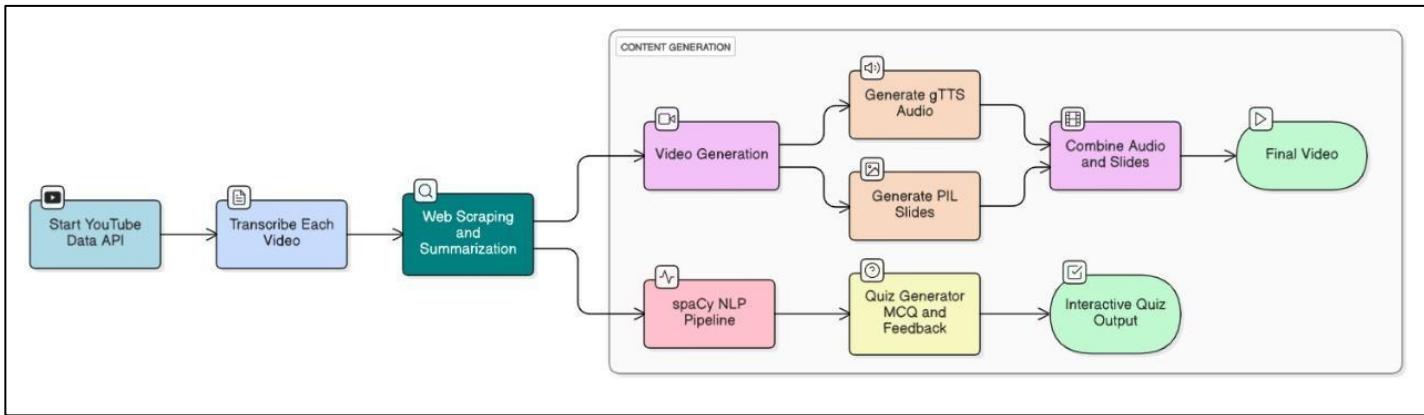


Fig1:System Design

5. IMPLEMENTATION PROCESS:

The completion of the proposed LEARNIFY system consists of multiple clearly identifiable stages that together promote efficient content generation, personalization, and adaptive learning.

- **Data Collection:** The entire process begins with methods of web scraping, where links to educational videos, titles, descriptions, and other relevant metadata (associated with the original educational video) are gathered from reputable sources online. The audio segments (i.e., video narrative) of the educational videos are extracted and subsequently processed by openly available speech-to-text (STT) software which converts the extracted audio segments into text. The process of transforming training video into text ensures both visual and verbal information are available for further processing and analysis.
- **preprocessing:** The text which was obtained from previously identified educational videos is subsequently preprocessed. This process will eliminate unnecessary data (e.g., filler words, noise, advertisement, and redundant sentences) from being part of a data sample. Preprocessing guarantees that the data is uniform, organized, and prepared for use in NLP tasks.
- **Summarization Based on NLP:** Semantic segmentation and text ranking algorithms individually or collectively can be utilized to pinpoint and extract the key observations. The entity resolution and topic analysis techniques will be utilized as a system to produce short yet coherent summaries of the initial training video material.
- **Video Generation:** After the above process, the condensed text can be converted to educational videos. AI-powered avatars make this process efficient, and Text-to-Speech (TTS) synthesis is used. Both animation as well as visual imagery shown on the screen along with the voice of the narrator for the educational video are synchronized during this step.
- **Tradition of Quizzes and Feedback:** The quizzes are automatically formed from the summary. The learner will be given immediate feedback that will provide details about incorrect options to support understanding and retention of the information.
- **Content Adaption to Learner Performance:** Gleaned user data will be analyzed for the personalization of the next content based on their performance metrics, like quiz scores and time spent on the topic. Our system will dynamically present the user with a difficulty level and recommendations for optimizing their learning performance.
- **Interface:** The user interface will present a format that will be simple for the learner to search for topics, watch dynamically created videos, take quizzes and read summaries. It will incorporate backend support for database storage and processing, and AI and NLP module integration.
- **Testing and Evaluation:** The platform will be tested to ensure accuracy, latency, scalability, and learner satisfaction. The platform will be continuously improved by analyzing performance metrics and the learner experience so we can create the optimal and sustainable learning platform.

6. EXPERIMENTAL RESULTS:

6.1 Register & login

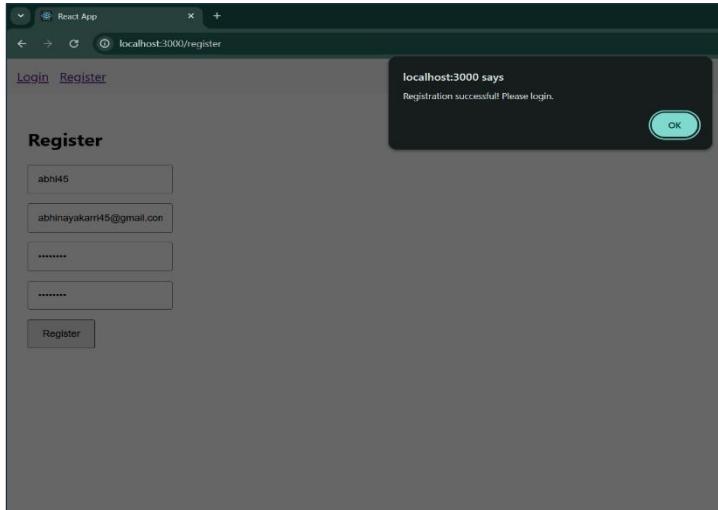


Fig1: Register page

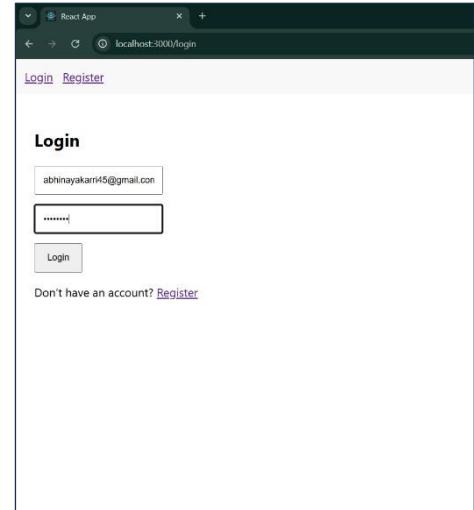


Fig2: Login

6.2 Search:

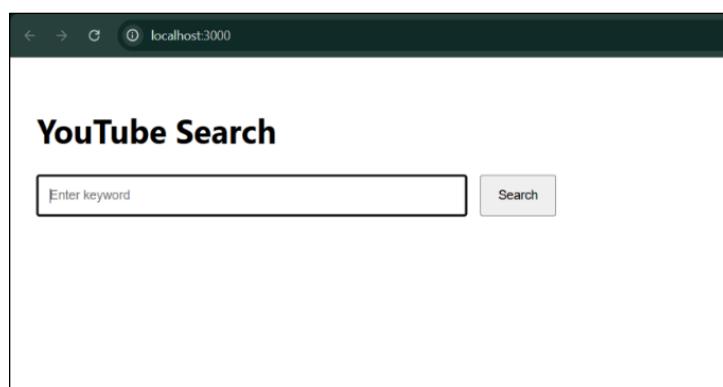


Fig3:Search

6.3 Filtered Videos

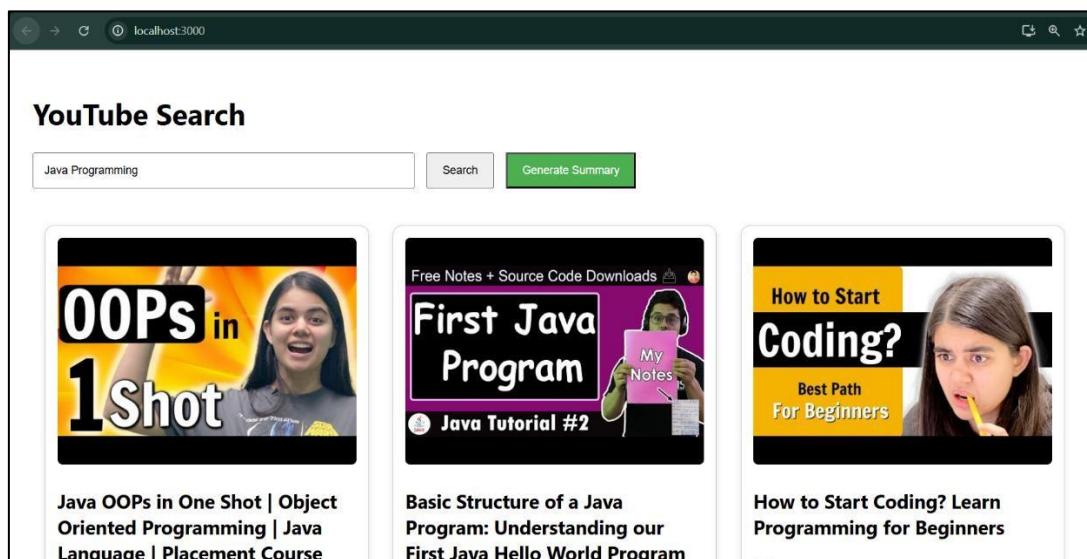


Fig4:Filtered videos

6.4 Transcript for each video

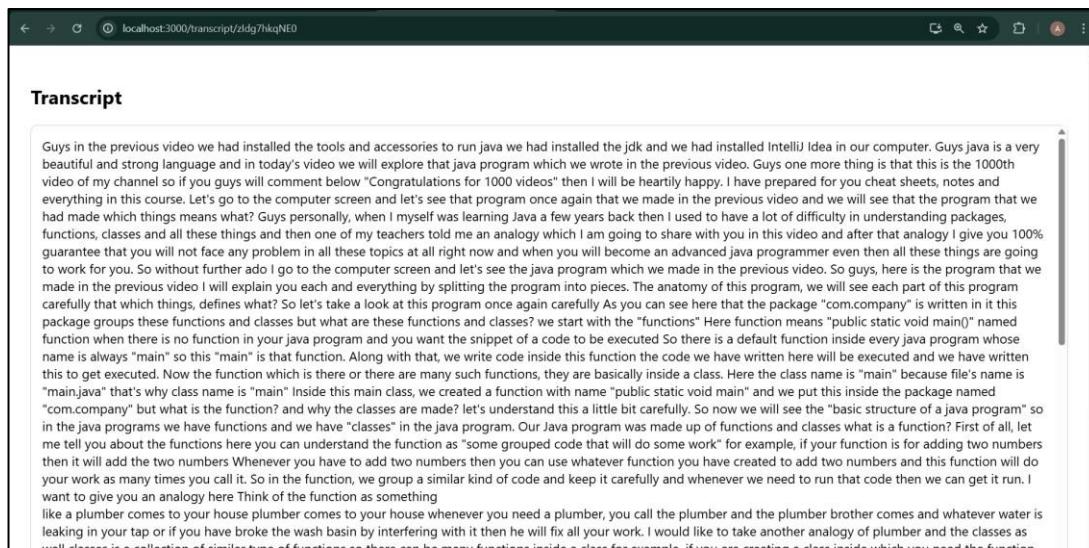


Fig 5: Transcript for each video

6.5 Summary of selected videos

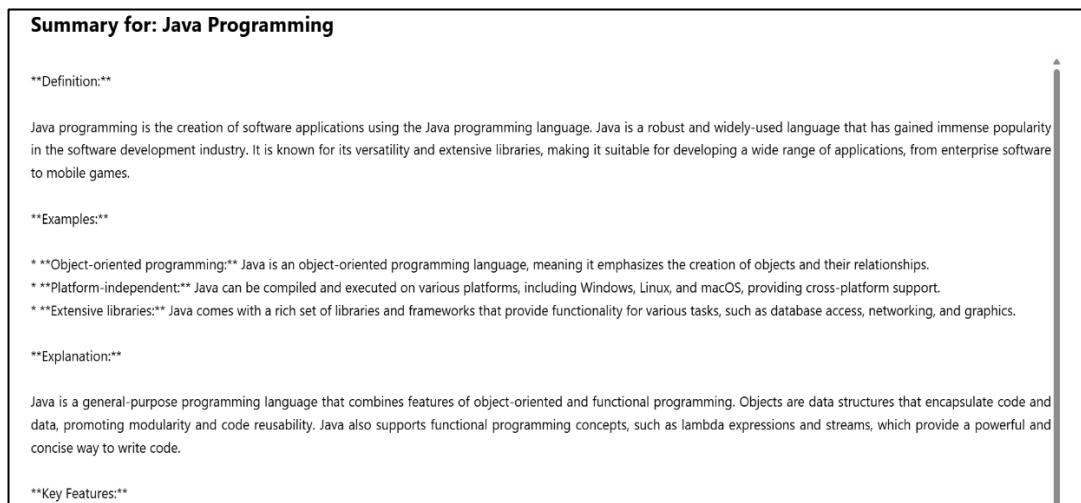


Fig 6: Summary

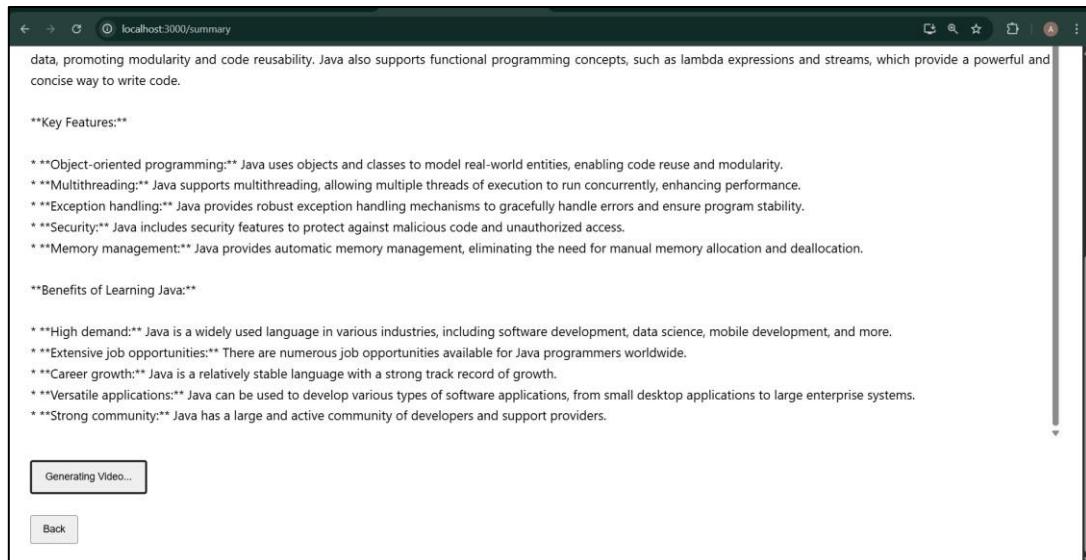


Fig 7: Summary

6.6 Video Generation

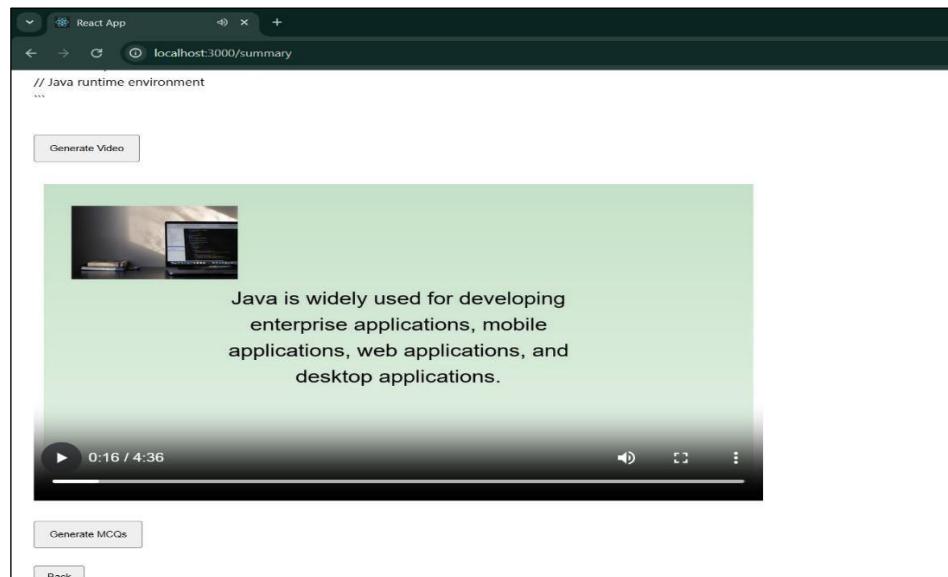


Fig 8: Video

6.7 Mcq's for Interactive Learning

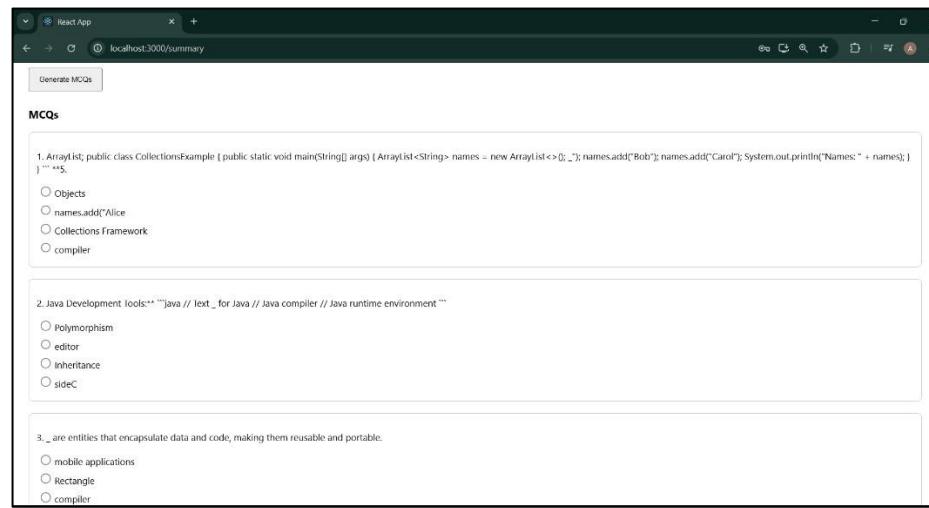


Fig 9: mcq's

6.8 Feedback

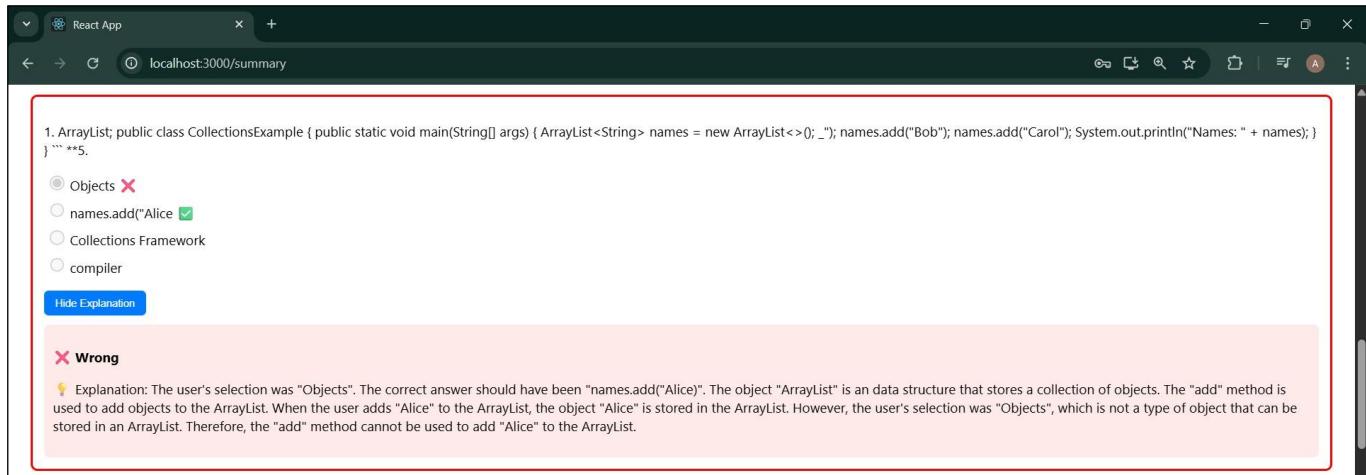


Fig 10: Feedback

7. CONCLUSION:

LEARNIFY offers a detailed structure for AI-based, NLP-founded learning video synthesis. In contrast to conventional static learning websites, it scrapes dynamically, analyzes, summarizes, and creates multimedia learning material from various online sources. Using smart summarization, adaptive testing, and feedback, it provides a dynamic and interactive learning environment.

Through redundancy reduction, improved understanding, and active learning facilitation, LEARNIFY shows the world how AI can transform online education. The future might hold emotion-sensitive feedback, voice command, and multi-language support, making LEARNIFY universally accessible and pedagogically sound.

REFERENCES:

- [1] A Review on Video to Text Summarization Techniques, Journal of Electrical Systems. Accessed on September 24, 2025. [Online]. Available: <https://journal.esrgroups.org/jes/article/download/4977/3628/9161>
- [2] Summarization of Video using Audio, IJRASET. Accessed on September 24, 2025. [Online]. Available: <https://www.ijraset.com/research-paper/summarization-of-video-using-audio>
- [3] Automated Video Summarization Using Speech Transcripts. ResearchGate. Accessed on September 24, 2025. [Online]. Available: https://www.researchgate.net/publication/220979489_Automated_video_summarization_using_speech_transcripts

- [4] A Cascaded Architecture for Extractive Summarization of Multimedia Content via Audio-to-Text Alignment, arXiv. Accessed on September 24, 2025. [Online]. Available: <https://arxiv.org/html/2504.06275v1>
- [5] Multi Aggregator for Content Aggregation and Contextual Learning, ResearchGate. Accessed on September 24, 2025. [Online]. Available: https://www.researchgate.net/publication/358040863_Multi_Aggregator_for_Content_Aggregation_and_Contextual_Learning
- [6] The Best Content Aggregation Strategies for Effective Online Marketing, Juicer.io. Accessed on September 24, 2025. [Online]. Available: <https://www.juicer.io/blog/content-aggregation-strategies>
- [7] Systematic Literature Review on Artificial Intelligence-Driven Personalized Learning, The Science and Information (SAI) Organization. Accessed on September 24, 2025. [Online]. Available: https://thesai.org/Downloads/Volume16No6/Paper_36- Systematic_Literature_Review_on_Artificial_Intelligence.pdf
- [8] Rapid Evidence Review: Technology-Supported Personalised Learning, EdTech Hub. Accessed on September 24, 2025. [Online]. Available: https://edtechhub.org/wp-content/uploads/2020/09/Rapid- Evidence-Review_-Technology-supported-personalised-learning.pdf
- [9] KG-Hub—Building and Exchanging Biological Knowledge Graphs, Oxford Academic. Accessed on September 24, 2025. [Online]. Available: <https://academic.oup.com/bioinformatics/article/39/7/btad418/7211646>
- [10] MERGE: A Modal Equilibrium Relational Graph Framework for Multi-Modal Knowledge Graph Completion, MDPI. Accessed on September 24, 2025. [Online]. Available: <https://www.mdpi.com/1424- 8220/24/23/7605>
- [11] A Beginner's Guide to Building Knowledge Graphs from Videos, Towards Data Science. Accessed on September 24, 2025. [Online]. Available: <https://towardsdatascience.com/a-beginners-guide-to-building- knowledge-graphs-from-videos-6cafcb5f3e5>
- [12] Free Text to Video AI – Create AI Videos from Text, VEED.IO. Accessed on September 24, 2025. [Online]. Available: <https://www.veed.io/tools/ai-video/text-to-video>
- [13] 10 Best Educational Video Making Software in 2025, Synthesia. Accessed on September 24, 2025. [Online]. Available: <https://www.synthesia.io/learn-training-videos/educational-video-making-software>
- [14] How AI Identifies Learning Gaps, Quizcat AI. Accessed on September 24, 2025. [Online]. Available: <https://www.quizcat.ai/blog/how-ai-identifies-learning-gaps>
- [15] Deep Learning Based Knowledge Tracing in Intelligent Tutoring Systems, ResearchGate. Accessed on September 24, 2025. [Online]. Available: https://www.researchgate.net/publication/393264654_Deep_learning_based_knowledge_tracing_in_intelligent_tutoring_systems