

AI-Based Presenter Coach

Shrawan M. Mokale¹, Titiksh K. Zaware², Utkarsha V. Dunde³, Swayam K. Pajgade⁴, Nilima Ramteke⁵

^{1,2,3,4,5}School of Computing, MIT ADT University, Pune, Maharashtra, India

Peer Review Information	Abstract
Type: Article Received: 27 March 2026 Revised: 12 April 2026 Accepted: 26 May 2026 Published: 16 June 2026	Speaking in Public and Presentation Delivery is Great Challenge for Professionals and Students. While there are some tools which provides a context or summary for the presentation or visual understanding but the user is not confident on his presentation and not get feedback except manual which is traditional way and not accessible when needed. This Paper Represents the Presentify an AI Based Full stack application which provides the script using RAG framework which creates chunking mechanism and store in vector database chroma, chat with presentation using API or Local LLM, and practice-based evaluation which will prepare you end to end to present your ppt confidently and accurately. Keywords: Retrieval-Augmented Generation (RAG); Large Language Models (LLMs); Vector Database; ChromaDB; Artificial Intelligence; Intelligent Presenter; Analytics.

How to Cite This Article

Mokale, S. M., Zaware, T. K., Dunde, U. V., Pajgade, S. K., & Ramteke, N. (2026). AI-based presenter coach. *Multidisciplinary Journal of Research in Engineering and Technology*, 13(2s), 255–260.

Introduction

Traditional Presentation involve a physical or manual feedback which evaluates the presenter but it has its own limitations like, the presenter will be not evaluated when he is needed, not understanding of the core concept or idea behind the presentation. This is where other tools came up like ChatGpt and Google, they provide the core idea or summary of the presentation which makes them clear about the concept but then also presenter is not confident because they are not practicing it and not getting feedback if they do. Here they rely on the traditional way so presenter just read and understand and repeat without understanding his growth.

Presentify solves this bottleneck using AI implementation and make presenter confident about his presentation. It is User Friendly application which uses RAG, Local LLM and performs the script generation, chat with power point and practice session where user will practice and get his performance score and analytical dashboard.

From real-world learning patterns, we observed that consistent learning is important for achieving good skill. Just like human skills improve through practice and feedback, our AI Presenter Coach follows a similar iterative learning approach. The system is iterated as per user usage and is trained, so it provides more precise and accurate context and response over time.

A major role in this system is played by Retrieval Augmented Generation (RAG) It enables fast retrieval of the context from the slides and store it to the vector DB also as the user needs the information retrieval is fast using cosine similarity and is used to generate impactful script and response. Also, we have used React.js, Typescript, and Analytical Library to generate the user-friendly dashboard for analyse the growth and report of the presentation you did.

Contributions:

- We developed Presentify that transforms PPT slides into structured, natural-speaking scripts using advanced language and RAG.
- We used a voice analysis module that evaluates user performance based on confidence, clarity, speaking pace, and filler word usage, and generates the report.
- We designed a RAG-based retrieval system using vector databases Chroma to ensure accurate script generation and intelligent chat coach to answer your doubt.
- A complete stack prototype was created in Python (Flask). backend) and recent frontend frameworks (React/Next.js.)
- The system exhibits better quality of presentation, decreased hesitation, and quantifiable skill development over in comparison to traditional practice methods

Related Work

AI driven coaching, teaching and learning have swiftly adjusted to educational market. As AI is continuous to enhance numerous learning stakeholders are transforming it and causing work to be quicker and easier. Traditional learning involves the human interaction and feedback which is not when you need always available when you need In contrast, digital interaction powered by AI provides instant support and individualized instructions, rendering it more effective to user.growth and development.

New improvement in Large Language Model. Natural Language Processing (NLP) and (LLM) which began to do well the context generation, conversation like human, summarization, and. But it was failing in response with Ground level knowledge-based questions. Model begins to hallucinate.

In order to overcome this shortcoming, a new methodology named Retrieval-Augmented Generation (RAG) is being used. It essentially integrates information search and AI generated responses, meaning that the answers are not arbitrary but on factual and current information. Such a method is already. application in fields such as document analysis, question answering systems and even in learning devices.

Meanwhile, speech analysis systems have also. improved a lot. They employ methods such as speech-to-text. conversion, identification of filler words (uh, um etc.), and analyzing speaking speed. It is due to this that they are able to give. feedback on communication skills like clarity, etc., is to be provided properly. fluency, and confidence. To process information, there are the use of vector databases such as ChromaDB are growing exceedingly popular. They store data as embeddings and help in fast and accurate searching based on meaning, not just exact words. Compared to traditional this enables them to be a lot more helpful as systems.

In addition, such modern web technologies as React, Next.js, and TypeScript help in building interactive dashboards. Such dashboards are capable of providing realtime feedback and insights, increasing the engagement and usability of the entire system.

The Proposed Approach

- The proposed system, AI Presenter Coach, is created to assist the users by enhancing their presentations. with doubt using script gen and performance evaluation solving coach.

- First, the user places a presentation file (e.g., .pptx). The system analyzes the slides, and reads the textual content and render pictures with images. This data is then converted into vector database so that it can be semantically retrieved.

The slide text is processed with the RAG engine followed by. The default prompt is shared with the LLM similar retrieved. chunks to the Model and Script is produced. For chat response user question based similar chunks are chosen and then it is given to the LLM. After creating the scripts, the user is able to practice sessions. Audio is recorded in the system during practice writes down the speech, and examines several parameters, including talking rate, filler words, pauses, confidence score and overall clarity.

The resulting product is:

- AI-generated scripts
- Voice analysis scores
- Feedback and suggestions

System Architecture

The system has three layers:

1. Input Layer

- Upload presentation(.pptx)
- Slide rendering and text extraction
- Create Structure json document

2. Processing Layer

- RAG creation
- Script gen using groq or llama
- Voice analysis (speech-to-text, filler detection, WPM calculation)
- Evaluation engine (scoring and feedback generation)

3. Output Layer

- Storage of scripts and session history (JSON/DB)
- Dashboard visualization (charts, scores, insights)
- Real-time interaction via web interface

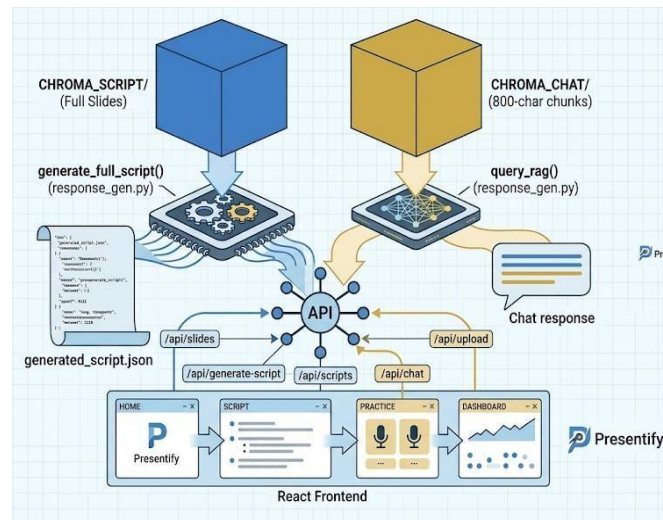


Fig. 1. System architecture

Data Representation

Each slide is converted into json:

```
{
  "slide_number": 1,
  "content": "Introduction to English",
}
```

```
"keywords": ["English", "Introduction"],
"embedding": [vector representation]
}
```

Practice session is stored in :

```
{
  "session_id": "001",
  "transcript": "...",
  "wpm": 135,
  "filler_words": 3,
  "confidence_score": 82,
  "clarity_score": 78
}
```

Retrieval-Augmented Generation (RAG)

The system creates scripts and responds to queries with the help of RAG:

- Find similar slide material using vectors.
- LLM (Llama or Groq) retrieved context.
- Produce correct, situational output.
- This ensures:
 - Reduced hallucination
 - Improved relevance
 - Improved response and training.

Voice Analysis and Evaluation

The system allows assessment of presentation quality based on:

- Words Per Minute (WPM) - Speaking pace - Measures the speed at which a person speaks.
- Filler Words Detection → Finds the words that are not needed.
- Pause Detection → Fluency Measuring.
- Clarity Score → Vocabulary based.

The overall score is computed as a weighted combination:

$$Score = w_1 \cdot Confidence + w_2 \cdot Clarity + w_3 \cdot Pace + w_4 \cdot FillerPenalty$$

Frontend Integration

The frontend is built using React/Next.js, providing:

- Slide preview and script display
- Real-time practice interface
- Interactive dashboards
- Session history tracking

APIs enable communication between frontend and backend for seamless data flow using flask framework.

Experimental Results

Parameter	Value
Total Practice Sessions	4
Slides in Presentation	6
Unique Slides Practiced	4
Total Words Spoken	120
Sessions with Transcript	3
Date Range	2026-03-19 to 2026-03-20
Presentation Topic	Emotional Intelligence in the Workplace

Fig. 2. Dataset Summary

AI-Based Presenter Coach

Dataset Configuration

- Presentations: Multiple .pptx files
- Slides per presentation: 5–20
- Practice sessions: Multiple per user

Evaluation Parameters

- Speech speed (WPM)
- Filler word count
- Pause duration
- Clarity with confidence score

Results and Analysis

- Improved speaking confidence over repeated sessions
- Reduction in filler words by ~30–50%
- More consistent speaking pace (optimal 120– 150 WPM)
- Faster feedback compared to manual evaluation

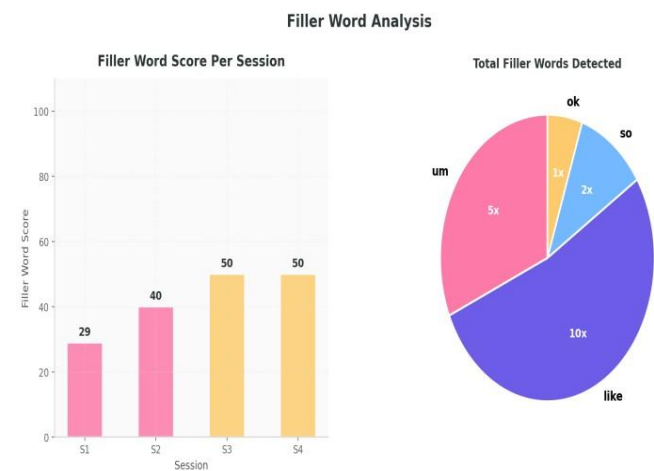


Fig. 3. Filler words analysis

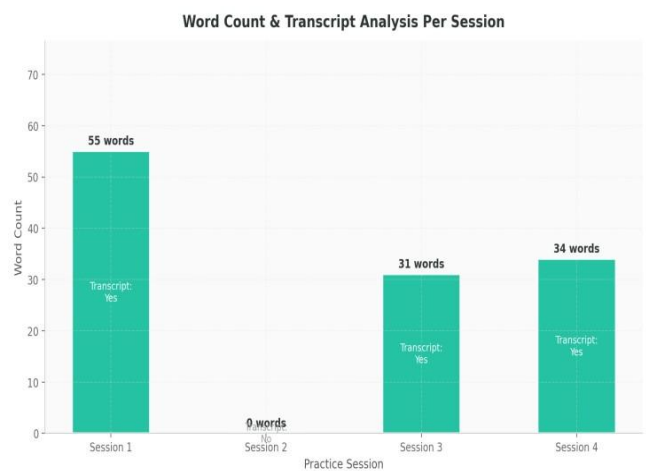


Fig. 4. Word count and Transcript analysis

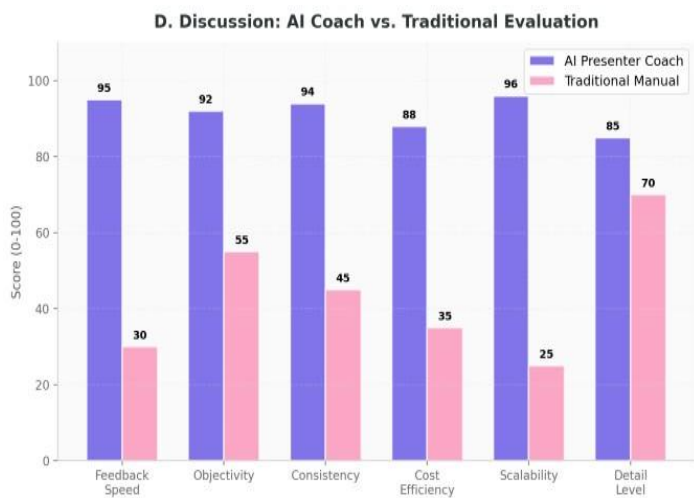


Fig. 5. AI Coach vs Traditional Evaluation



Fig. 6. Timeline Analysis of Presentation Practice Sessions and Overall Scores

Future Work

The envisioned system offers real-time feedback during presentations, which allows users to get feedback as they present. Rather than offering feedback at the end of the presentation, this system provides real-time analysis of voice tone, speed, clarity, and confidence, with

discreet cues or suggestions. This allows users to correct errors like speaking too quickly, filler words, or disengagement, and improves the overall quality of the presentation. Another important system is the learning system, which provides feedback tailored to the user's previous performance.

It monitors the users' progress, recognises patterns in their performance, and tailors advice. If a presenter frequently lacks confidence or struggles with vocal tone, the system will focus on these aspects and give specific feedback. This results in a personalised learning experience, akin to having a coach that knows the presenter's progress. The system is also able to accommodate multiple languages, allowing for international usage. Participants will be able to present in various languages and the system will still be able to provide feedback on speech patterns, tone and clarity. This makes the system accessible to a wider audience, and helps people from different language backgrounds to enhance their speaking skills without any restrictions. Moreover, the system also uses emotion recognition using speech tone analysis where it measures the pitch, volume and intonation of speech to detect emotions like confidence, anxiety and excitement. This helps the presenter gain feedback on their presentation. This is complemented by gesture and posture analysis via computer vision, where the system, using a camera, assesses body posture, hand gestures, and other aspects of body language. It helps them adopt confident posture, appropriate eye contact and effective gestures, which are essential for effective presentations. Overall, this makes a powerful AI-based presenter coach that provides in-depth feedback to improve communication skills, both spoken and non-verbal, in an interactive and personalised manner.

Conclusion

This AI based system for presentation provides smart and scalable way to improve the presentation skills. By combining the RAG and Analysis with feedback generation, it helps user to practice more effectively and flexibly which gradually improves the skills to present it. This system reduces the manual time and gives consistent results based on the previous session as well. It has the modular design which well suited for the students, teachers even professional to have good user experience. Overall, this AI system keep a great step toward the AI based learning where it improves their skills in personalized and impactful way.

References

1. Lewis, P., et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," 2020.
2. Fan, W., et al. "RAG Meets LLMs: A Survey," 2024.
3. Gao, L., et al. "Retrieval-Augmented Generation for Large Language Models: A Survey," 2023.
4. Karpukhin, V., et al. "Dense Passage Retrieval for Open-Domain Question Answering," 2020.
5. Izacard, G., & Grave, E. "Leveraging Passage Retrieval with Generative Models," 2021.
6. Chen, S., et al. "PresentCoach: Dual-Agent Presentation Coaching System," 2025.
7. Wang, Y., et al. "AI-Assisted Presentation Training with Real-Time Feedback," 2024.
8. Kumar, R., et al. "Intelligent Systems for Public Speaking Skill Enhancement," 2024.
9. Yang, G., et al. "MaLa-ASR: Large Language Model-Based Speech Recognition," 2024.
10. Google. "Speech Recognition API Documentation," 2024.
11. Jurafsky, D., & Martin, J. H. "Speech and Language Processing," 3rd ed., 2023.
12. Shriberg, E. "Spontaneous Speech: How People Really Talk and Why Engineers Should Care," 2005.
13. Brown, T., et al. "Language Models are Few-Shot Learners," 2020.
14. Touvron, H., et al. "LLaMA: Open and Efficient Foundation Language Models," 2023.
15. OpenAI. "GPT Models and Applications," 2023.
16. Johnson, J., et al. "Billion-Scale Similarity Search with GPUs (FAISS)," 2019.