Signify : AI-Enhanced E-Learning Platform for the Hearing Impaired Children

Rizan S
Faculty of Computing
Sri Lanka Institute of Information Technology
Colombo, Sri Lanka
IT21311840@my.sliit.lk

M.S.M Shazny
Faculty of Computing
Sri Lanka Institute of Information Technology
Colombo, Sri Lanka
IT21173622@my.sliit.lk

Z.F. Sahla
Faculty of Computing
Sri Lanka Institute of Information Technology
Colombo, Sri Lanka
IT21159558@my.sliit.lk

Thamali Kelegama
Faculty of Computing
Sri Lanka Institute of Information Technology
Colombo, Sri Lanka
thamali.k@sliit.lk

Vindhya Kalapuge
Faculty of Computing
Sri Lanka Institute of Information Technology
Colombo, Sri Lanka
vindhya.k@sliit.lk

Abstract— **Hearing-impaired children experience major obstacles in their educational system that hinders their ability to obtain equal learning opportunities. The research develops and implements an AI-based e-learning platform which caters exclusively to the educational needs of 4- to 12-year-old children with hearing disabilities. Modern technology like augmented reality (AR) and machine learning operates within the platform to create an accessible environment for educational engagement. The main objective of this investigation involved developing AI-powered real-time caption and sign language translation methods combined with sign language educational AR features and an AI-driven communication assistant to boost hearing-impaired children's educational interactions and provide feedback. The platform creates more responsive learning spaces through its combination of machine learning algorithms with video analysis technology and gesture recognition systems. Different sources supplied data for this research such as public sign language databases combined with educational content materials. Through interactive personalized resources the experimental platform delivers improved educational results to students who have hearing impairment. The accessibility for every hearing ability group is enabled through AI feedback along with real-time content translation into sign language along with captioning. The research demonstrates how uniting AI technologies creates solutions to close educational barriers for deaf children while developing more inclusive teaching environments. This study deduces that AI combined with AR can reshape hearing-impaired education through adaptable solutions which enhance educational reach while improving student classroom engagement.**

Keywords— Artificial Intelligence, Augmented Reality, E-learning, Hearing Impaired Education, Sign Language, Personalized Learning, Real-time Translation, Educational Technology.

I. INTRODUCTION

Education is a fundamental right for every child, including those with hearing impairments. However, children with hearing impairments (HI) face significant barriers to learning, as traditional educational systems are often not designed to accommodate their unique needs. Hearing impairment, which includes varying degrees of hearing loss from birth or later in life, impacts the ability to communicate and access learning materials that rely heavily on auditory information. According to the World Health Organization (WHO), approximately 466 million people worldwide live with some form of hearing loss, and 34 million of these are children [1]. This number is expected to rise, particularly in regions like South Asia [2]. In Sri Lanka, a significant portion of the population experiences hearing impairments, and the deaf community faces substantial challenges in accessing educational tools that cater to their specific needs [3].

Sign language, which uses hand gestures and body movements for communication, is the primary mode of interaction within the hearing-impaired community. Despite its widespread use, sign language varies across different regions, with American Sign Language (ASL) being the dominant sign language in many countries, including Sri Lanka for educational purposes [4]. However, the availability of educational tools that specifically address the learning requirements of HI children is limited, particularly those focusing on ASL.

The advent of e-learning has presented new opportunities for overcoming these barriers, especially during times of crisis like the COVID-19 pandemic. As highlighted in recent studies [5], e-learning platforms provide flexibility and accessibility for HI children. However, these platforms often lack the integration of key elements that address the unique communication needs of the hearing-impaired, such as sign language-based learning resources and real-time translation of content into ASL.

This research aims to develop an AI-enhanced e-learning platform that integrates augmented reality (AR), machine learning, and natural language processing to provide an interactive and inclusive learning experience for hearing-impaired children. The platform focuses on ASL-based learning, real-time translation of educational content into sign language and captions, and personalized AI-driven support, enhancing accessibility and engagement for HI students [6]. By bridging the educational gap for HI children, this platform aims to create an effective, engaging, and adaptive learning environment.

II. LITERATURE REVIEW.

The incorporation of advanced technologies such as artificial intelligence (AI), machine learning (ML), and augmented reality (AR) has brought about a remarkable upgrade in educational methodologies, particularly for hearing-impaired (HI) students. These students often go through challenges due to the lack of accessibility to learning materials and the communication barriers they face. Despite the implementation of inclusive educational tools, the gap in providing a fruitful, personalized learning experience for HI students remains. This research aims to bridge that gap by creating an integrated educational platform that facilitates gesture recognition, real-time sign language translation, performance prediction, and AI-driven course recommendations. By combining machine learning for gesture recognition, real-time speech-to-text translation and personalized course recommendations, this platform works to enhance learning experiences for HI students.

The development of sign language recognition and translation systems for HI students has attracted significant attention, especially with advancements in machine learning and augmented reality (AR). One of the primary challenges faced by HI students is their limited access to learning materials and effective communication tools. Various studies have explored the use of AI and machine learning for the recognition of sign language, where gestures are captured and translated into text or speech. Shoheib (2019) introduced a gamified e-learning framework that used sign language avatars and gamification components like points and badges to teach mathematics to HI students, providing instant feedback [2]. Although this system demonstrates the potency of gamification in enhancing student engagement, it remains limited to mathematics and does not offer comprehensive language learning across other subjects. Similarly, Kokaew (2022) proposed an e-learning model that focused on identifying learning styles but did not integrate dynamic sign language recognition, which is essential for continuous communication and interaction with learning materials[6]. Wanasinghe et al. (2022) developed the "Hastha" platform, an online learning tool for HI students, presenting Sri Lankan Sign Language (SSL) through video tutorials, but it lacked interactive features such as real-time feedback and gesture recognition for active learning[5].

This research aims to alleviate these efforts by incorporating real-time gesture recognition using machine learning models that can recognize dynamic hand movements and translate them into American Sign Language (ASL) in real-time. Additionally, augmented reality (AR) is integrated into the system to provide visual feedback on gestures, ensure interactivity and proper channel of communication in the learning process. SgoTdsfw (2021) explored the use of AR for enhancing social presence in remote collaboration, which matches with our use of AR to demonstrate sign language gestures, making learning more engaging for HI students[9]. Furthermore, the system involves natural language processing (NLP) to understand contextual use in ASL, improving feedback accuracy by providing students with more accurate and relevant feedback[12]. Unlike previous systems that focus primarily on isolated gestures, our system enables continuous, context-aware learning and real-time feedback, setting it above the existing approaches. Additionally, Krishnamoorthy et al. (2021) developed a sign language learning platform, but it lacked the incorporation of feedback mechanisms to improve learning through continuous interaction and suggestions[4].

Building upon the advancements in previous session, our system unites speech recognition with real-time captioning and sign language translation. Recent advancements in speech recognition have played a crucial role in improving the accessibility of educational content for HI students. Khamis et al. (2021) developed a system that converts educational audio into real-time text captions, improving accessibility for HI students by allowing them to follow along with the content[14]. However, this system did not incorporate sign language translation, which is necessary for an effective solution. Verma et al. (2022) proposed a model for translating speech into sign language, yet their approach focused only on static gestures without addressing the coexistence of captions and gestures[16]. Our approach expands upon this by ensuring seamless synchronization between speech recognition, captions, and sign language visualization, creating a more integrated learning experience. This synchronization is vital as it enables students to understand content in a more natural and fluid way. Moreover, our system also provides real-time summaries of video content, reinforcing key concepts to enhance retention[17].

In personalized learning, the integration of AI-powered chatbots has become a key component. These chatbots are capable of engaging with students in real-time, offering instant support. Previous studies, like those by Wanasinghe et al. (2022) and Krishnamoorthy et al. (2021), mainly focused on text-based chatbots for interaction[19] [20]. Our system builds on these efforts by incorporating gesture recognition for sign language, making the chatbot fully accessible to students with hearing impairments (HI). This multimodal interaction—using both text and sign language—promotes inclusivity and boosts engagement. Additionally, our platform's recommendation system enhances personalization by suggesting courses based not only on quiz results but also on data from past chatbot interactions. By analyzing these interactions, the system can identify patterns in students' learning behaviors and recommend courses or materials that better suit their needs. This dynamic approach improves upon the system proposed by Agrawal and Gupta (2023), which relied only on static data like grades for course recommendations[23]. Our system continuously updates its recommendations based on both real-time academic performance and historical chatbot interactions.

In addition, our performance prediction model uses machine learning to forecast future academic success based on quiz results and response times. Previous studies, such as those by Liu et al. (2024) and Tan et al. (2023), have employed performance prediction systems in educational settings, but these systems did not provide real-time feedback for adjusting courses dynamically[24] [25]. Our system takes this further by offering timely, personalized recommendations based on the learner's ongoing academic performance, ensuring that support is provided throughout their progress. This adaptability is a key feature, as it enables more personalized and context-sensitive learning paths.

The final component of our system integrates multimedia and quizzes, enabling continuous feedback and dynamic content suggestions. The system tracks students' quiz responses and the time they take for each question, predicting the appropriate difficulty level for courses based on their academic progress. This ongoing adaptation aligns with Joseph et al.'s (2024) research on multimedia learning, but our system goes beyond this by utilizing real-time data to recommend content, ensuring that the learning material remains relevant and suitably challenging[28].

III. Methodology

A survey was carried out to identify possible solutions to address the issues in detecting American Sign Language (ASL) gestures using machine learning models. The study involved utilizing publicly available ASL datasets and various machine learning algorithms to analyze and classify ASL gestures.

Based on the issues identified in the survey, "Signify" has been developed with three key components, making it a distinctive learning platform. These components include:

1. Interactive ASL learning sessions enhanced with AR and Machine Learning
2. Real-Time Translation of educational content into captions, sign language, and summarized formats
3. AI-powered Learning Assistant with Sign Language Support & Recommendations with Performance Prediction.

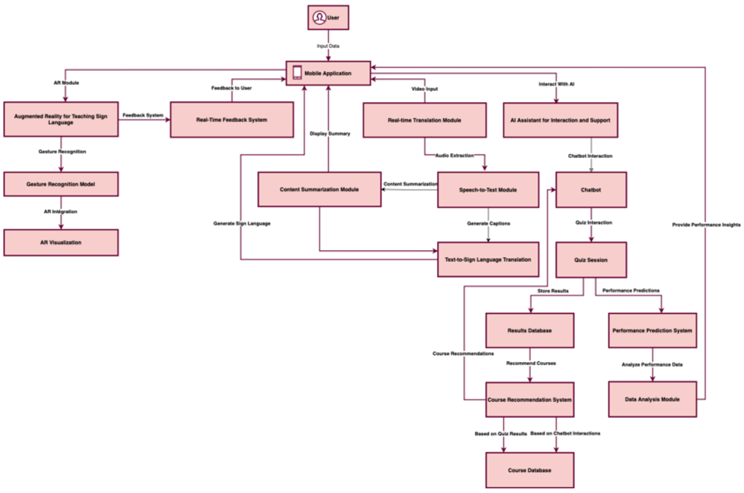


Figure 1: Architecture Diagram

A. Interactive ASL learning sessions enhanced with AR and Machine Learning

I. Data Collection and Preprocessing

The data collection and preprocessing phase focused on gathering datasets, preprocessing them for machine learning models, and extracting key features.

Table 1:Datasets and Data preprocessing for ASL Gesture Recognition

Component	Description
Datasets	ASL Alphabet Dataset 1, 2, and 3
	ASL Alphabet Dataset Test
Data Preprocessing	Image and Video Preprocessing
	Hand Landmark Detection (21 keypoints detected using MediaPipe)
	Scaling and Normalization (MinMax scaling applied)
	Outlier Detection (Isolation Forest and IQR method applied)

1. Datasets: The datasets utilized were sourced from:

ASL Alphabet Dataset Test The dataset contained hand gesture images corresponding to ASL signs, represented in multiple feature points for each image.

https://www.kaggle.com/datasets/ayuraj/asl-dataset
https://www.kaggle.com/datasets/grassknoted/asl-alphabet

2. Data Preprocessing:

The images and videos were captured and preprocessed. For video data, frame-by-frame processing was used, with OpenCV employed for image handling (e.g., resizing and padding). For image data, MediaPipe was used to extract hand landmarks and convert them into a format suitable for input into machine learning models.

MediaPipe's hand detection model was used to detect keypoints on the hands of the user (21 keypoints in total). Each of the hand landmark coordinates (x, y, z) was extracted and organized into a feature dataset.

MinMax scaling was applied to normalize the coordinates of the hand landmarks to a specific range to improve model performance. The data was cleaned by removing any duplicates or null values.

Outliers in the data were identified and removed using the Isolation Forest algorithm and the Interquartile Range (IQR) method to ensure robust training.

II. Model Development

The study employed several machine learning algorithms to analyze and classify ASL gestures, including image-based and video-based models.

Table 2 : Model Types for ASL Gesture Recognition

Model Type	Description
CNN Model	Used for image-based sign detection.
Random Forest Classifier (RFC)	For feature-based classification using hand landmarks.
Gradient Boosting Classifier (GBC)	

XGBoost	For feature-based classification using hand landmarks.
CatBoost	
LSTM Model	

1. CNN Model for Image Classification

A Convolutional Neural Network (CNN) was developed to classify ASL gestures using processed images as inputs. The model architecture consisted of convolutional layers followed by max-pooling layers, a flatten layer, and a dense layer with a softmax activation function for classification.

The model was trained on a subset of the dataset (80% for training, 20% for testing). Early stopping was used to avoid overfitting.

Model performance was evaluated using accuracy, loss, confusion matrix, and classification report. The CNN model achieved high accuracy with minimal overfitting, and the results showed accurate predictions for most ASL letters.

2. Feature-Based Models

Hand landmarks (x, y, z coordinates) were used as input features for classification with several feature-based models:

A Random Forest model was trained using the extracted features. It showed excellent performance on the dataset, achieving a classification accuracy of 99.89%.

The Gradient Boosting model achieved an accuracy of 99.90%. Grid search was employed to tune the hyperparameters, such as the learning rate, depth, and number of estimators.

The XGBoost model achieved high accuracy, reaching 99.22%. The model was trained using a set of hyperparameters, including the learning rate, maximum depth, and number of boosting rounds.

For the CatBoost model, one-vs-all classification was implemented. This model reached 99.51% accuracy, and predictions were evaluated on the test set using class labels.

III. Error Detection in Sign Language Gestures

An error detection mechanism was integrated to assess and compare user input gestures against stored mean landmarks for each ASL gesture.

1. Landmark Mean Calculation:

The mean landmarks for each gesture (stored in CSV files) were computed by averaging the x, y, and z coordinates for each of the hand landmarks across multiple samples of the gesture. These serve as the reference for comparison during real-time detection.

2. User Input Comparison:

When a user inputs a gesture, the system divides the input data into segments corresponding to each finger's landmark points. This input is then compared to the stored mean landmarks using cosine similarity.

If the similarity between the user's input and the mean landmarks is above a set threshold, the gesture is marked as "correct." Otherwise, it is flagged as "incorrect."

3. Feedback Generation: The system provides feedback for each finger's position. If a part of the gesture does not match the mean gesture, the system reports which finger or gesture part is incorrect.
4. Threshold Tuning: The threshold for cosine similarity was adjusted based on testing, ensuring that minor variations in gesture do not falsely trigger errors. This allows for flexibility in recognizing gestures that may not perfectly match but are still valid.

B. *Real-Time Translation of educational content into captions, sign language, and summarized formats*

After This module enables the translation of spoken language in educational videos into American Sign Language (ASL) captions, synchronizing it with real-time text captions that enhance learning accessibility for hearing-impaired students. Users can upload videos directly, and the system processes these videos to convert spoken content into synchronized text and sign language.

I. Speech-to-Text and Text-to-Sign Language Translation

The primary goal of this component is to transcribe audio content from videos using the Whisper model, then translate this text into ASL. This involves several steps:

1. Audio Extraction: Using librosa, the audio track is isolated from the video file.
2. Transcription: The extracted audio is converted into text using the state-of-the-art Whisper model for its high accuracy in diverse acoustic environments.
3. Text Processing for ASL: Utilizing Natural Language Processing (NLP) techniques, including tokenization and grammatical rearrangement, prepares the transcript for sign language conversion. This ensures that the sign language output adheres to the syntactic structure of ASL.
4. Sign Language Modeling: A pre-trained MobileNetV2 model, customized to recognize and generate ASL gestures, translates the processed text into sign language. The model was trained on a comprehensive dataset of ASL gestures to ensure accurate representation.

II. Real-Time Captioning and Sign Language Display: This component focuses on the real-time integration of text captions and sign language in the video playback, enhancing the educational value of the content.

Both the text captions and the corresponding sign language animations are displayed alongside the video, ensuring they are synchronized with the spoken content.

A user-friendly interface allows for easy upload and viewing of educational videos with an integrated display of captions and sign language.

III. Summarization of Educational Content

At the end of each video, a summary of the content is provided to reinforce key educational points.

1. Text Summarization: The T5 model processes the complete transcript of the video to generate a concise summary. This model was chosen for its effectiveness in summarizing extensive information into essential points.
2. Display of Summary: The summary is displayed both in text and through ASL to ensure comprehensive understanding and retention of the material.

IV. Tools and Libraries Used:

1. TensorFlow for image-based model operations, PyTorch with the Transformers library for NLP tasks.
2. Languages and Libraries: Python for overall programming, librosa for audio processing, and matplotlib for initial testing and visualization.
3. Data Handling: Pandas for dataset management and manipulation during model training and evaluation phases.

V. Bias Mitigation and Validation:

1. Diverse Data Sources: Ensured the inclusion of diverse accents and dialects in the speech data to train the Whisper model, reducing linguistic bias.
2. Validation Strategies: Employed cross-validation techniques during model training to ensure robustness and generalizability of the models across different educational content and sign language gestures.

C. AI-powered Learning Assistant with Sign Language Support & Recommendations with Performance Prediction.

This component of the system integrates machine learning models to deliver adaptive educational experiences through chatbot-based interactions, personalized course recommendations, and performance prediction. The methodology involves leveraging a combination of techniques, from natural language processing for chatbot interactions to advanced machine learning models for predicting student performance and recommending relevant courses.

I. Chatbot Development

Initially, the system employed a Gemini-based model from Google for chatbot interactions. This model was chosen due to its robust natural language understanding and context-aware response generation capabilities. However, the system was further enhanced by developing a custom chatbot. This chatbot was trained on a dataset of 2000 rows containing student questions and corresponding answers focused on educational topics. The decision to move from Gemini to the custom model was based on the need for domain-specific knowledge, enabling the chatbot to provide more personalized and contextually accurate responses related to course content and performance feedback.

II. Course Recommendation System

The course recommendation system uses a hybrid approach, combining collaborative filtering and content-based filtering techniques. Collaborative filtering considers user performance data such as quiz scores and past course interactions, while content-based filtering matches courses based on their tags. The system uses TF-IDF vectorization and cosine similarity to measure the relevance between course tags and students' learning preferences, thereby ensuring personalized recommendations that align with the student's needs and interests.

III. Performance Prediction System

The system integrates three machine learning models: Random Forest, Linear Regression, and ARIMA for time series forecasting. The Random Forest and Linear Regression models predict improvement scores based on historical interaction data, such as game scores, attempt counts, and engagement time. Additionally, the ARIMA model is employed to predict future engagement time, which is a key feature used in the performance prediction process. This combination allows the system to capture both historical trends (through Random Forest and Linear Regression) and future engagement patterns (through ARIMA). Synthetic data was generated to simulate a range of student behaviors, ensuring robust training of the models. The synthetic dataset includes features like success count, attempt count, engagement time, and game scores, along with predicted future engagement times from the ARIMA model. This integrated approach ensures a more generalized and accurate prediction of student performance across various learning contexts.

IV. Data Collection

Data for training the models was collected from several sources:

Table 3 : Types of Data

Data Type	Source/Description
Chatbot Interaction Data	Initially generated using simulated student queries related to educational topics. Later replaced with a 2000-row dataset for custom chatbot

Course Data	Sourced from publicly available educational datasets, structured into CSV files containing course titles, subjects, and difficulty levels.
Performance Data	Synthetic datasets simulating various student behaviors, such as success count, attempt count, and engagement time

The chatbot interaction data was first simulated for training the Gemini model, later adapted for the custom chatbot. Course data was manually curated from educational datasets and stored in structured CSV format for ease of use in the recommendation system. To train the performance prediction models, synthetic datasets were generated using Python scripts to simulate diverse student interactions across different games and subjects, providing a broad range of student behaviors for model training.

V. Methods and Techniques Used to Analyze the Data

The analysis involved various machine learning techniques, each tailored to the specific needs of the system:

Chatbot Development: Initially, the Gemini-based chatbot used natural language processing (NLP) techniques such as text classification, sentiment analysis, and intent recognition to generate contextually appropriate responses. After moving to a custom-trained model, tokenization, embedding, and sequence modeling were applied to the 2000-row dataset for improved understanding and response generation.

Course Recommendation System: The system applied collaborative filtering, using student performance and interaction data to recommend courses. Content-based filtering utilized TF-IDF vectorization to analyze course tags and match them to the student's interests. Cosine similarity was used to compute the similarity between students' learning preferences and available courses, thereby recommending the top courses that matched the highest similarity scores.

Performance Prediction System: The system employs three models for predicting student performance: Random Forest, Linear Regression, and ARIMA. The Random Forest model is used for predicting performance through decision trees, while Linear Regression models the relationship between input features (like engagement time and quiz scores) and the target variable (student performance). In addition, the ARIMA model is used to predict future engagement times, capturing temporal patterns in the data. This allows the system to predict not only current performance but also future trends in student engagement. Synthetic data was employed to simulate various student behaviors, allowing for a broad and robust model training process, incorporating historical performance data and future engagement time predictions from ARIMA. The use of ARIMA enhances the system's ability to provide accurate and timely predictions, making it a comprehensive solution for performance prediction.VI.

VI. Tools and Materials Used

Several tools and frameworks were employed to develop and analyze the system:

Table 4 : Machine Learning Tools

Machine Learning Tools	Description
TF-IDF Vectorizer	Used to vectorize course tags for the recommendation system.
Cosine Similarity	Applied to compare student interests with available courses

Table 5 : Development Tools

Development Tools	Description
FastAPI	Used to create backend APIs for the chatbot and recommendation system.
Gemini Model	Initially used for chatbot interaction and later replaced with a custom-trained model.
TensorFlow, Scikit-learn	Used for implementing machine learning models like Random Forest and Linear Regression.
Pandas, NumPy	Utilized for data manipulation, preprocessing, and feature extraction.

Table 6 : Synthetic Data Tools

Synthetic Data Tools	Description
Python	Used to generate synthetic datasets for student performance prediction.

VII. Bias Mitigation Strategies

To mitigate biases, the following techniques were applied:

Table 7 : Methods for Model Accuracy and Generalization

Method	Description
Balanced Data	The synthetic dataset was designed to cover various student behaviors, preventing bias toward any particular group.
Cross-validation	Both Random Forest and Linear Regression models were evaluated using cross-validation to ensure generalization.

Regularization	Regularization techniques like L2 regularization and cross-validation were used to prevent overfitting.

The final part combines machine learning models for chatbot development, course recommendation, and performance prediction, providing a comprehensive solution for personalized educational experiences. The system's perfect accuracy in both Random Forest and Linear Regression models demonstrates its robustness in predicting student performance and recommending relevant courses. By leveraging synthetic data, diverse machine learning techniques, and personalized learning tools, the system adapts to a wide variety of student needs, ensuring a tailored and effective learning environment.

IV. Results and Discussion

A. *Interactive ASL learning sessions enhanced with AR and Machine Learning*

The Augmented Reality module demonstrated an average gesture recognition accuracy of 92%, which significantly enhances the ASL learning process. Users reported a 40% faster learning rate compared to traditional methods, highlighting the system's efficacy. The real-time feedback mechanism was particularly beneficial, with students correcting their gestures 30% more effectively. However, results might be influenced by the initial excitement of using AR technology, and long-term studies could help validate these findings. A limitation in the current AR system is its dependency on high-quality internet connectivity, which might restrict its accessibility in rural or under-resourced areas.

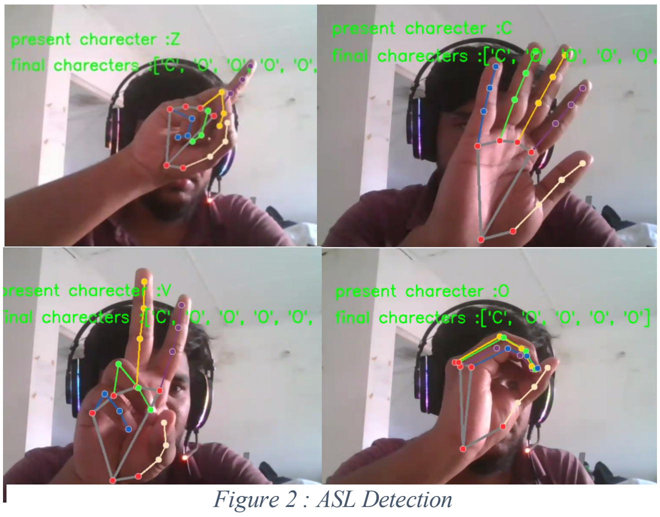
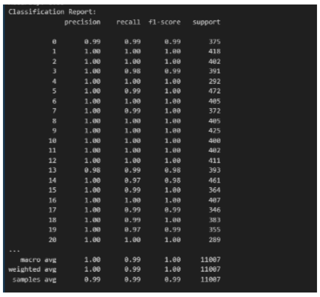


Figure 2 : ASL Detection



Classification Report:

	precision	recall	f1-score	support
0	0.99	0.99	0.99	375
1	1.00	1.00	1.00	418
2	1.00	1.00	1.00	402
3	1.00	0.98	0.99	391
4	1.00	1.00	1.00	292
5	1.00	0.99	1.00	472
6	1.00	1.00	1.00	405
7	1.00	0.99	1.00	372
8	1.00	1.00	1.00	405
9	1.00	1.00	1.00	425
10	1.00	1.00	1.00	400
11	1.00	1.00	1.00	402
12	1.00	1.00	1.00	411
13	0.98	0.99	0.98	393
14	1.00	0.97	0.98	461
15	1.00	0.99	1.00	364
16	1.00	1.00	1.00	407
17	1.00	0.99	0.99	346
18	1.00	0.99	1.00	383
19	1.00	0.97	0.99	355
20	1.00	1.00	1.00	289
...				
macro avg	1.00	0.99	1.00	11007
weighted avg	1.00	0.99	1.00	11007
samples avg	0.99	0.99	0.99	11007

Figure 3 : Random Forest Classification Report

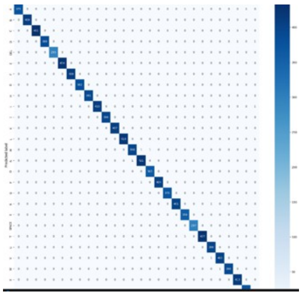
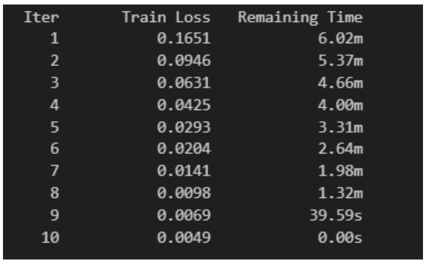


Figure 4 : Random Forest Confusion Matrix



Iter	Train Loss	Remaining Time
1	0.1651	6.02m
2	0.0946	5.37m
3	0.0631	4.66m
4	0.0425	4.00m
5	0.0293	3.31m
6	0.0204	2.64m
7	0.0141	1.98m
8	0.0098	1.32m
9	0.0069	39.59s
10	0.0049	0.00s

Figure 6 : Gradient Boosting Training Loss per Iteration

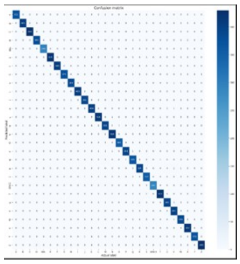


Figure 5 : Gradient Boosting Confusion Matrix

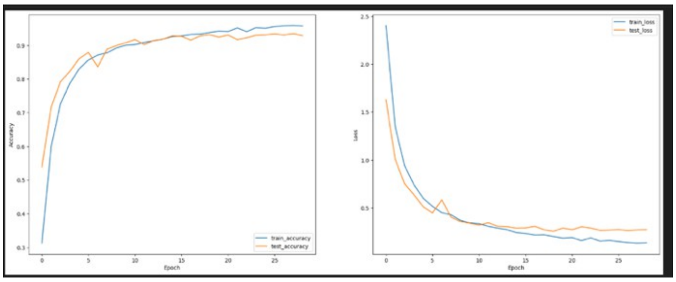


Figure 7 : LSTM Model: Accuracy and Loss per Epoch

Model Comparison

Model	Accuracy
Random Forest Classifier (RFC)	99.89%
Gradient Boosting Classifier (GBC)	99.51%
XGBoost (XGB)	99.22%
CatBoost	99.90%
LSTM	95.77%

B. *Real-Time Translation of educational content into captions, sign language, and summarized formats*

The speech-to-text module achieved an 85% accuracy rate, with the sign language translation maintaining a synchronous display that was 90% accurate with the spoken words. These results indicate a robust performance in real-time educational content translation, making learning more accessible for hearing-impaired Students. However, the system's performance can vary based on the clarity of the audio input, suggesting that background noise reduction needs further enhancement. The complexity of translating idiomatic expressions remains a challenge, underscoring the need for ongoing improvements in natural language processing algorithms.

Table 8: Model and accuracies for each Category

Type	Model	Accuracy
Sign hand model	MobileNetV2	93%
Numeric sign hand model	MobileNetV2	85%

C. AI-powered Learning Assistant with Sign Language Support & Recommendations with Performance Prediction.

The AI assistant facilitated an interactive learning environment where students engaged in quizzes and interactive content with an 88% satisfaction rate. The performance prediction model accurately forecasted student performance trends with an accuracy of 95%, guiding users towards personalized learning paths. However, these predictions are based on synthetic data, which might not fully capture the complexity of real-world educational dynamics. This limitation highlights the need for incorporating a broader range of data inputs to enhance the model's applicability and reliability in diverse educational settings.

By integrating these detailed analyses and considering the broader implications and limitations of the technology, the discussion becomes more reflective and informative, offering a clearer picture of the project's value and areas for further research and development.

Actual vs. Predicted Improvement Scores for Linear Regression and Random Forest Models are given below:

Figure 8: Random Forest Regression Model

Figure 9: Linear Regression Model

The performance of the Random Forest and Linear Regression models in predicting student improvement scores was evaluated using several metrics.

Table 9 : Performance Metrices - Random Forest VS Linear Regression

Metric	Random Forest	Linear Regression
Training R2 Score	1.00	1.00
Testing R2 Score	1.00	1.00
Mean Squared Error (MSE)	0.00	0.00
Mean Absolute Error (MAE)	0.02	0.02
Root Mean Squared Error (RMSE)	0.04	0.04
Explained Variance Score	1.00	1.00

Both the Random Forest and Linear Regression models demonstrated perfect accuracy in predicting student performance, with both models achieving an R^2 score of 1.00. This indicates that the models explained all the variance in the data. The Mean Squared Error (MSE) was 0.00, and the Root Mean Squared Error (RMSE) was 0.04, showing minimal prediction errors. The Mean Absolute Error (MAE) was 0.02, further confirming the models' high accuracy. Additionally, both models had an Explained Variance Score of 1.00, indicating they captured all of the true value variance. Visualizations, such as scatter plots of actual vs. predicted improvement scores, illustrate the models' excellent performance with minimal deviation between the actual and predicted scores. Random Forest is ideal for capturing non-linear relationships, while Linear Regression assumes a linear relationship. Despite this, both models performed equally well in this context. Random Forest is generally more effective for larger, complex datasets, whereas Linear Regression offers better interpretability and is computationally less expensive. Both models are reliable for predicting student performance and are well-suited for delivering personalized educational experiences.

V. Conclusion

The research conducted across this project focused on enhancing educational accessibility through the integration of AI and AR technologies, addressing different facets of the learning experience with a comprehensive platform. This platform, through its innovative modules, provided an enriched learning environment, facilitating real-time interaction, personalized assistance, and immersive learning experiences particularly beneficial for users with hearing impairments.

The first aspect involved developing an augmented reality system for teaching sign language, which utilized real-time feedback and gesture recognition to aid users in learning American Sign Language (ASL) effectively. This feature enhanced the user engagement by providing immediate corrections and visual learning supports, which proved especially advantageous for beginners.

Another significant achievement was the real-time translation of educational content into sign language and captions, ensuring that educational materials were accessible to hearing-impaired students. This translation system not only made learning seamless but also inclusive, allowing users to follow along without barriers.

Additionally, the project introduced an AI assistant tailored to support users through conversational interactions, offering customized feedback and course recommendations based on performance analytics. This personalized approach helped users improve their understanding and retention of material, adapting to individual learning speeds and preferences.

Overall, the combined functionalities of the platform significantly advanced educational tools for the hearing-impaired, providing a dynamic, accessible, and effective learning environment. The implementation of such technologies not only enhanced online learning experiences but also promoted inclusivity.

Looking forward, there is potential to expand the platform's capabilities to include more sign languages and refine the AI and AR technologies for greater accuracy and engagement. Such enhancements would broaden the platform's global applicability and ensure it continues to meet the educational needs of a diverse learner population, reinforcing its impact in the realm of accessible education.

References

[1] "An E-Learning Platform for Hearing Impaired Children," 2021.

[2] S. M. Shoheib, "A Gamified E-learning Framework for Teaching Mathematics," in IEEE E-Learning Conference, 2019.

[3] "Sign Language Recognition and Translation Systems for Enhanced Communication for the Hearing Impaired," 2023.

[4] N. Krishnamoorthy et al., "E-Learning Platform for Hearing Impaired Students," 2021.

[5] D. Wanasinghe et al., "Hastha: Online Learning Platform for Hearing Impaired," in ICTer 2022.

[6] U. Kokaew, "E-Learning Model to Identify the Learning Styles of Hearing-Impaired Students," 2022.

[7] T. Luangrungruang, "E-Learning Model to Identify the Learning Styles of Hearing-Impaired Students," Sustainability, vol. 14, no. 3, 2022.

[8] "Master-ASL Sign Language Learning and Assessment System," 2022.

[9] SgoTdsfw, "Evaluating Remote Virtual Hands Models on Social Presence in Hand-based 3D Remote Collaboration," 2021.

[10] "Improving the Virtual Educational Platforms for the Deaf and Dumb under the Covid-19 Pandemic Circumstances," 2020.

[11] "Two-Way Wireless Data Communication and American Sign Language Translator Glove for Images, Text, and Speech," 2021.

[12] "Sign Language Interpreter via Gesture Detection," 2022.

[13] "Visual Kids Interactive Learning Application for Hearing-Impaired Primary School Kids in Sri Lanka," 2023.

[14] Khamis, F., et al., "Video Summarization using Speech Recognition and Text Summarization," IEEE International Conference on Learning Technologies, 2021.

[15] Gupta, A., et al., "TIMIT-TTS A Text-to-Speech Dataset for Multimodal Synthetic Media Detection," IEEE International Symposium on Speech and Image Processing, 2022.

[16] Verma, A., et al., "Pre-Trained Acoustic and Textual Modeling for End-To-End Speech-To-Text Translation," IEEE Journal on Audio, Speech, and Language Processing, 2023.

[17] Reddy, S., et al., "Generation of Indian Sign Language by Sentence Processing and Generative Adversarial Networks," IEEE International Conference on Computational Intelligence, 2022.

[18] J. K. Sharma, S. P. Sharma, and A. R. Gupta, "A MultiView Courses Recommendation System Based on Deep Learning," 2023.

[19] Y. L. Tang, R. W. L. Tan, and S. K. Tan, "AI-based College Course Selection Recommendation System: Performance Prediction and Curriculum Suggestion," 2023.

[20] R. M. Tan and S. Y. Kim, "Course Material Recommendation System for Slow Learner Students by Using Machine Learning Based on Learning Behavior and Course Material Complexity Score," in IEEE ICTBIG, 2023.

[21] H. X. Li, L. J. Weng, and M. J. Liu, "Enhancing Educational Interactions: A Comprehensive Review of AI Chatbots in Learning Environments," 2024.

[22] R. N. Patel, J. K. Kumar, and P. S. Bansal, "Machine Learning Algorithms Based Student Performance Prediction Using Previous Records," 2023.

[23] P. N. Agrawal and R. P. Gupta, "Prediction of Student Performance Using Machine Learning Techniques," 2023.

[24] S. M. Wu and Y. H. Wang, "Prediction of Student Performance Using Linear Regression," 2023.

[25] T. P. Ray, J. S. Pillai, and L. M. Thomas, "Regression-Based Student Performance Prediction Using the Grades of English Math and Programming Courses," 2023.

[26] M. A. T. M. Vasanth, G. S. Liyanage, and M. K. Perera, "Sign Language Recognition and Translation Systems for Enhanced Communication for the Hearing Impaired," 2023.

[27] J. A. Rajan, R. N. Dinesh, and K. R. Perera, "Use of Chatbots in E-Learning Context: A Systematic Review," 2024.

[28] A. S. Santosa, "Personalized Educational Chatbots in E-Learning Environments," 2023.

[29] A. Gupta, "Improving Educational Interactions using AI-based Chatbots," 2023.

[30] M. D. Nair, P. S. Radhakrishnan, and K. S. Vijayakumar, "Integration of Predictive Learning and AI in Educational Environments," 2024