

Automated sentiment analysis of visually impaired students' audio feedback in virtual learning environments

Doaa Mohamed Elbourhamy

Educational Technology and Computer Department, Faculty of Specific Education, Kafrelshiekh University, Egypt

ABSTRACT

This research introduces an innovative intelligent model developed for predicting and analyzing sentiment responses regarding audio feedback from students with visual impairments in a virtual learning environment. Sentiment is divided into five types: high positive, positive, neutral, negative, and high negative. The model sources data from post-COVID-19 outbreak educational platforms (Microsoft Teams) and offers automated evaluation and visualization of audio feedback, which enhances students' performances. It also offers better insight into the sentiment scenarios of e-learning visually impaired students to educators. The sentiment responses from the assessment to point out deficiencies in computer literacy and forecast performance were pretty successful with the support vector machine (SVM) and artificial neural network (ANN) algorithms. The model performed well in predicting student performance using ANN algorithms on structured and unstructured data, especially by the 9th week against unstructured data only. In general, the research findings provide an inclusive policy implication that ought to be followed to provide education to students with a visual impairment and the role of technology in enhancing the learning experience for these students.

Subjects Artificial Intelligence, Computer Education, Sentiment Analysis, Neural Networks

Keywords Sentiment analysis, Visually impaired student, Audio feedback analysis, Academic performance, E-learning accessibility, Artificial intelligence in education, CNN, SVM

Submitted 1 March 2024

Accepted 29 May 2024

Published 24 June 2024

Corresponding author

Doaa Mohamed Elbourhamy,
doaa.elborhami@spe.kfs.edu.eg

Academic editor

Paulo Jorge Coelho

Additional Information and
Declarations can be found on
page 26

DOI [10.7717/peerj-cs.2143](https://doi.org/10.7717/peerj-cs.2143)

© Copyright
2024 Elbourhamy

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

INTRODUCTION

Academic performance assessment is the key to assisting students to navigate their learning journey effectively and teachers to identify and offer assistance to the learners if the process is lacking. Usually, these are done through formative and summative assessments that afford a view of students' learning trends and processes. However, these assessments may be complicated and time-consuming, often overlooking some of the important aspects involved. A better way to do this would be the use of automatic computer-based techniques in student data analysis to make the assessment process more effective and efficient.

The coming of big data and developments in computational technology can be said to have been at the base of the growth of learning analytics in educational research. This field tries to reveal patterns associated with learning in educational settings, resulting in an improvement in understanding of educational processes, assessment of learning outcomes, and prediction of academic performance. A number of algorithms and methods have been

very instrumental in creating different models for the prediction of underachievement in academics, like clustering, classification, decision trees, association rule mining, regression, and visualization.

For instance, researchers such as [Bayer et al. \(2012\)](#), [Vijayarani & Janani \(2016\)](#) identified that school dropouts and failures can be predicted by examining the social behaviors and relationships of students within the school. Consequently, social network analysis was integrated with machine learning approaches, namely support vector machine (SVM) and Naïve Bayes classifiers, as well as the J48 decision tree, to increase the accuracy of prediction ([Saif et al., 2013](#); [Schouten & Frasincar, 2015](#)). Equally, [Shahiri, Husain & Rashid \(2015\)](#) reviewed various analytical techniques for predicting the final academic grade. [Jayaprakash et al. \(2014\)](#) refer to intervening very early as essential in preventing academic difficulty for at-risk failing students. [Costa et al. \(2017\)](#), on the other hand, report that it helps in identifying the eventual failure well in advance that, in turn, makes it possible to take timely educational support to avoid being at a loss.

To identify such at-risk students earlier, [Costa et al. \(2017\)](#) proposed improved data quality by performing a sequence of activities: gathering, cleaning, and transforming data while improving their model by selecting relevant attributes. Nevertheless, this is not without criticism, considering that conventional forms of academic success prediction are mainly based on structured data, for example, demographics and records of previous academic achievements. However this kind of data does not capture all the factors that can be associated with the academic performance of a student ([Mohammdi & Elbourhamy, 2021](#)). The important thing is not to value only those unstructured data sources that would just merely improve understanding and make earlier predictions of who would be at risk than these conventional predictors.

Student self-evaluations are a valuable source of unstructured data rich in sentimental content that can, in turn, shed light on students' sentimental well-being ([Zhang & Liu, 2017](#)). [Yu et al. \(2017\)](#) and [Liu \(2015\)](#) gave much importance to the fact that students' written comments bore open students' attitudes toward learning, students' understanding of the subject, and perceptions of lesson difficulty. This information is useful to helping teachers design and improve learning activities. [Pekrun et al. \(2002\)](#) argued that sentiments experienced during class, study sessions, and exams play a significant role in academic success. Research by [D'Mello et al. \(2008\)](#), [Yu et al. \(2017\)](#) and [Rodrigues et al. \(2016\)](#) have also served to evince the influence of sentimental states on academic performance. As was shown by [Blikstein & Worsley \(2016\)](#), affective state analysis, or sentiment analysis, is innovative in multimodal learning analytics. It will establish the sentimental states of students in the course of studying and learning.

It involved 100 university visually impaired students who gave a voice after lessons, opening the way for automation of the process of evaluating feedback to enable professors to get reports on each student's learning experience promptly. This automation is expected to make the feedback process much easier. With that, educators will be better placed to modify their teaching techniques in a manner befitting the special cases of visually impaired students. This article is organized as follows: "Literature Review" outlines relevant literature and establishes identified existing research gaps, more so technological

solutions, for persons with visual impairment in educational settings. “Research Method” discusses the methodology that we used and presents, in detail, how machine learning with sentiment analysis is placed in the educational context. “Results and Visualization” showcases the experimental results, and “Findings and Conclusions” discusses the findings in an attempt to draw implications for the area of Inclusive Education. Finally, in “Recommendations” and “Limitations and future directions”, the conclusions are drawn from this study reflecting on the broader implications of this research and future research directions.

LITERATURE REVIEW

Visual impairment and academic performance

Visual impairment, encompassing both total blindness and low vision, presents unique challenges in the academic context. These challenges necessitate adaptive teaching methodologies and curriculum designs to accommodate the specific learning needs of visually impaired students (*Mwakyeja, 2013; Agesa, 2014*). Notably, visually impaired learners have shown remarkable aptitude in language acquisition through auditory methods, often demonstrating enhanced verbal skills facilitated by active questioning and vocabulary enrichment (*Ghafri, 2015*).

The challenges of visual impairment in academic performance

Regarding the “Challenges of Visual Impairment in Academic Performance,” it is critical to note that much of the existing research on students with disabilities has been predominantly medical in focus, often overlooking their educational potential and specific challenges within higher education settings (*Office of the Higher Education Commission, 2013*). This study seeks to bridge this gap by introducing predictive models that leverage visually impaired students’ audio feedback to anticipate their learning outcomes (*Elbourhamy, Najmi & Elfeky, 2023*). This innovative approach aims to empower educators to fine-tune their instructional strategies, thereby more effectively addressing the unique learning needs of visually impaired students.

Student sentiments and learning

The results of the research on the relationship between sentiment and learning outcome are summarized in [Table 1](#). The table presents results that have been found throughout studies conducted in various years. The studies indicated above give clear insight into the emotional dimension of learners’ involvement.

Sentiment analysis in education: a new approach for visually impaired students

However, sentiment analysis is traditionally applicable to domains that have such aims as dealing with unique educational needs, particularly in students with visual impairments. Some advanced methods of tracking sentiment are made possible by recent progress in natural language processing and machine learning, contrasting with the classic burdensome tests, which in general are impossible to apply feasibly for such students. This

Table 1 Abstract literature review of student sentiments and learning.

Title	Year	Finding
Academic sentiments in students' self-regulated learning and achievement: A program of qualitative and quantitative research.	2002	Earners' sentiments refer to the emotional experiences of students during academic situations such as classes, studying, and tests. It has been found that students' emotional states have a direct impact on their motivation and academic performance. Test anxiety is a prime example of how negative sentiments can hinder learning and achievement, caused by factors such as comparing one's performance to others, fearing the consequences of failure, and feeling unprepared for exams, all of which are linked to lower test scores (<i>Pekrun et al., 2002</i>).
Using emotional and social factors to predict student success	2003	Research has shown that emotionally healthy students are more likely to succeed in college (<i>Pritchard & Wilson, 2003</i>).
AutoTutor detects and responds to learners' affective and cognitive states	2008	Therefore, it is essential to detect learner sentiments to promote effective learning, as there is a strong link between cognition and emotion. Positive sentiments, for instance, can increase students' interest and motivation to learn (<i>D'Mello et al., 2008</i>).
VADER: A parsimonious rule-based model for sentiment analysis of social media text	2014	The optimal performance across all criteria was assessed by comparing emotion scores generated by VADER with those produced by 11 other widely recognized sentiment analysis tools and techniques. VADER has achieved its maximum potential in terms of performance when dealing with large datasets (<i>Hutto & Gilbert, 2014</i>).
Predicting learning-related sentiments from students' textual classroom feedback <i>via</i> Twitter	2015	A range of machine learning techniques were employed to predict essential emotion types for learning by utilizing textual input from students. To maintain students' interest in response to different emotion types, diverse educational strategies can be implemented. For instance, if students are feeling bored, they could be given more challenging tasks, whereas if they are feeling perplexed, providing additional examples could aid their comprehension (<i>Altrabsheh, Cocea & Fallahkhair, 2015</i>).
A review of social media posts from UniCredit Bank in Europe: A sentiment analysis approach	2019	VADER was selected for emotion analysis due to its superior performance and outstanding capability in conducting sentiment analysis on brief documents, such as tweets or other social media texts (<i>Botchway et al., 2019</i>).
A comprehensive study on lexicon-based approaches for sentiment analysis	2019	This research pertains to a distinct domain as it compares the accuracy of lexicon-based techniques, including VADER, TextBlob, and NLTK. The study found that VADER approaches demonstrated high levels of accuracy as measured by precision, recall, and F1 score metrics (<i>Bonta & Janardhan, 2019</i>).
Sentiment analysis for tweets in Swedish	2019	This article briefly outlines the conventional approach to sentiment analysis, which involves the utilization of classified training data when employing a machine learning approach. This data is then used to train an algorithm that can predict the classification of new, unseen data. However, this process is time-consuming, particularly given the scope of this article, which necessitated the investigation and testing of multiple machine-learning techniques. Ultimately, the VADER Sentiment analyzer was selected instead (<i>Botchway et al., 2019</i>).
Using VADER sentiment and SVM for predicting customer response sentiment	2020	This study examined the accuracy of various algorithms, including machine learning algorithms and lexicon-based approaches like Support Vector Machines (SVMs) and VADER. The study found that VADER provided greater accuracy compared to machine learning algorithms like SVM (<i>Borg & Boldt, 2020</i>).

Table 1 (continued)

Title	Year	Finding
Deep sentiment analysis using CNN-LSTM architecture of english and roman urdu text shared in social media	2022	This study focuses on sentiment analysis in Roman Urdu and English, particularly on social media. It introduces a new method combining CNN-LSTM and machine learning classifiers to analyze sentiments. This method is tested on four different datasets and shows high accuracy. For Roman Urdu, the Word2Vec CBOV model and SVM classifier work best, while for English, BERT embeddings, two-layer LSTM, and SVM are more effective. The new approach improves accuracy by up to 5% compared to existing methods (<i>Khan et al., 2022</i>).
Roman Urdu sentiment analysis using transfer learning	2022	This study discusses the need for better tools to analyze large amounts of online text, focusing on sentiment analysis. Current research mainly looks at English, but not much at languages like Roman Urdu because they are more complex. Deep neural networks are popular for this, but there's still more to learn about them. CNNs are good for sentiment analysis but need lots of data and assume all words are equally important. This study suggests improving CNNs with an attention mechanism and transfer learning to make sentiment analysis better. The new model seems more accurate than existing methods (<i>Li et al., 2022</i>).
Multimodal sentiment system and method based on CRNN-SVM	2023	The research introduces a new technique for sentiment analysis using multiple modalities, like speech and facial expressions, instead of relying solely on text. This approach results in higher accuracy in detecting emotions. The study demonstrates that deep learning models like CRNN-SVM achieve the highest emotion recognition rates, with an accuracy of up to 93.5%. This method aims to improve the accuracy of sentiment analysis by utilizing multiple sources of information (<i>Zhao et al., 2023</i>).
TSA-CNN-AOA: Twitter sentiment analysis using CNN optimized via arithmetic optimization algorithm	2023	The study analyzed the psychological impact of COVID-19 by examining public opinions on Twitter. Using a special method (TSA-CNN-AOA), it classified 173,638 tweets into positive, negative, or neutral categories. This approach was found to be highly effective, achieving an accuracy rate of 95.098% in classifying tweets, demonstrating its superiority over similar methods (<i>Aslan, Kiziloluk & Sert, 2023</i>).

Note:

Synthesizes prior research exploring the interplay between student sentiments and their educational experiences.

approach is based on sentiment classification, not only based on the polarity but also based on intensity through continuous and long-term monitoring for more targeted educational reactions. The following system architecture explains the proposed techniques.

Recently conducted meta-analyses and systematic reviews disclose that sentiment analysis tools can be effectively applied in educational contexts. The present study suggested that it could thus be a robust tool to be used not only in class but also in online and other multimedia settings to up students' engagement level in learning (*Xie & Luo, 2021; Nguyen et al., 2014; Gustafsson, 2020*). To this effect, the study has spearheaded the ways of embedding the data-full-width = "true" sentiment analysis tools that would also be responsive to the visually impaired students—a field that has so far not been much explored. These tools help in decoding complex emotional signals, which are needed to generate feedback and support at par with the specific learning context of these students (*Calvo & D'Mello, 2010*).

In another study, the new framework incorporates categorical and dimensional approaches in the analysis of sentiment, particularly the valence-arousal space that

captures emotional intensity and polarity in the application of sentiment analysis in multilinguistic settings (Kiritchenko & Mohammad, 2017; Bo & Lee, 2008). Such a framework has been noted, according to the need of the study, to enhance the challenges of arousal and valence detection that were identified in prior studies (Yu et al., 2016).

Based on the information provided

1. Specialized focus on educational sentiment analysis: This work pushes the frontier in terms of sentiment analysis in the domain of education, which is even less explored by prior works.
2. Improved sentiment tracking method on students: This is just an aspect in our system but with enhanced means of how the development of sentiments in students is tracked, mostly after the educational sessions. This is highly improved compared with other tools that analyze sentiments, but they do not have the sensitivity that is required to perfectly record the slight, small changes in sentiments occurring within the educational environment.
3. Nuanced sentimental classification: Going beyond the traditional three-tier (positive, neutral, negative) classification of sentiment in reviews, our article shall propose a more nuanced five-tier system (High Positive, Positive, Neutral, Negative, High Negative). This detailed categorization opens up deeper insights into the sentimental state of students which is very often missed in many other systems.
4. Filling a gap: Thus, by giving importance to the spontaneous feedback of one demographic group that had been otherwise neglected in previous analyses, the study fills a significant gap in the present research.
5. Focused on the visually impaired: This study would add value because it is focused on visually impaired students and is the first of its kind to deal with the sentimental responses of such a rare crop within an educational context.
6. Development of an innovative analytical system: We propose an advanced system that processes the audio feedback from visually impaired students to predict the performance of such students academically and sentimentally.
7. Educational data integration: The uniqueness of our system from most of the existing solutions lies in the fact that it integrates structured educational data with unstructured sentimental feedback, hence providing an overall insight into the learning experiences and the resulting outcome of the student. Such integrated data will further give greater accuracy and more practical conclusions for the stakeholders.

RESEARCH METHOD

Data source

Our study involved 100 impaired vision students from various universities around Egypt and was carried out for 9 weeks. The course “Computer Skills Course in the English Language” was started through Microsoft Teams. All participants were communicated to, and forms of consent were signed before the course, whereby a choice of withdrawal/non-

participation without penalties was given. The process of data collection was extensive and carefully integrated into the Microsoft Teams platform to ensure comprehensiveness and no loss of accuracy.

Throughout the course, multidimensional data were collected, including structured academic performance indicators and unstructured sentimental feedback. Both these kinds of data, when integrated, provided the opportunity to have an overall view of the progress and sentimental responses of the student throughout the course. Our data collection was multi-faceted and integrated into the Microsoft Teams platform. and we were gathering data as follows:

- *Digital tracking on Microsoft Teams (structured data)*: We fully used the inbuilt functionalities of Microsoft Teams to track student participation. The same included checking attendance in each session, checking the submission rate of homework, and analyzing involvement in classroom discussions. In this regard, Microsoft Teams provided us with electronic logs and exported participation reports for analysis to quantify student engagement by using the “Insights” feature to get detailed reports on student engagement, which includes grades, attendance, and activity metrics. The first evaluation was done by the end of the fourth week, and the second one by the end of the ninth week.

- *Collection of audio feedback (unstructured data)*: This work illustrates the study’s responsiveness, which is useful to help provide a useful means for identifying the needs of visual impairment in the section of the program where major assessments are needed. The first evaluation was done by the end of the fourth week, and the second one by the end of the ninth week, after audio feedback had been collected. The students were to audio-record themselves for every lesson while speaking, self-assessing, and reflecting on the learning content, then submit their recordings on the platform Microsoft Teams. The approach did not help in collecting valuable qualitative data but also provided in-depth insights into the experiences and understanding of the students.

In preparing the data for analysis, first, the audio feedback was normalized to have uniform levels of audio, hence making the inputs possess a uniform value. Then, the recordings went through voice activity detection technology for segmentation. Isolated spoken content needed to be obtained from the different interjections, background noises, and even silences to get only the spoken part of the data. This clean data input was then ready for transcription, setting the stage for further sentiment analysis and educational assessments.

The integration of these sources was structured and unstructured data, which brings the view of a student’s journey in academia into focus. This nuanced and varied approach allowed tracking academic performance in not just an objective manner but also gaining access to sentimental nuances, thus giving a larger perspective on the overall experience in the course for the students. For instance, the audio entries volume, which had been 2,648 in the fourth week, by the ninth week, was 4,237, indicative of an even active student participation level. These contained invaluable, rich audio feedback with insightful information in relation to the behaviors of students and how they were interacting with the material and instructors of the course.

Table 2 Data source.

Attribute type		Attribute name	Description
Predictors	Structured data	Homework grade	The grades of homework on average
		Homework click	Num of reading homework instruction
		Attendance	Num of attending class
		Discussion	Num of post discussion
	Unstructured data of students' sentiments	High negative post	Num of high negative post
		Negative post	Num of negative post
		Natural post	Num of natural post
		Positive post	Num of positive post
		High positive post	Num of high positive post
Target		Academic performance	Failed or passed this course

Note:

Data sources collected including attendance records, homework completion, class participation, and the sentiments extracted from the audio feedback, ranging from high positive to high negative.

Table 2 shows details of all structured and unstructured data sources, the types of data collected, from attendance records, homework completion, and class participation, to sentiments extracted from audio feedback, from high positive to high negative.

This methodological approach helps bring about a multifaceted understanding of the course's impact on the student's skills and sentimental responses.

Method for analysis

The study developed a flow chart to evaluate the effectiveness of an intelligence prediction system.

Figure 1 presents the tailored system developed for assessing visually impaired students' sentimental responses *via* audio feedback. This sophisticated system is composed of three integral sub-systems: the Speech Recognition sub-system (STT), which transcribes audio to text; the Sentiment Analysis sub-system (SA), which evaluates the sentimental content of the feedback; and the Prediction sub-system, which anticipates students' academic performance based on their sentimental feedback. Detailed descriptions of the functionality and interplay between these sub-systems follow in the subsequent section.

Speech recognition sub-system

The STT is a fundamental component for converting the spoken interactions of visually impaired students into text for further analysis. The module processes audio data stored in the "students-audio-feedback" folder, representing student interactions with the system. We used the SpeechRecognition Python library version 3.8.1, known for its efficiency and accuracy in processing varied speech data.

Data retrieval and pre-processing

- Audio retrieval: The "RetrieveAudioFile" function gathers all audio files from the "students-audio-feedback" folder.
- Audio normalization: To ensure uniform audio levels, the "Librosa" library standard peak normalization was applied.

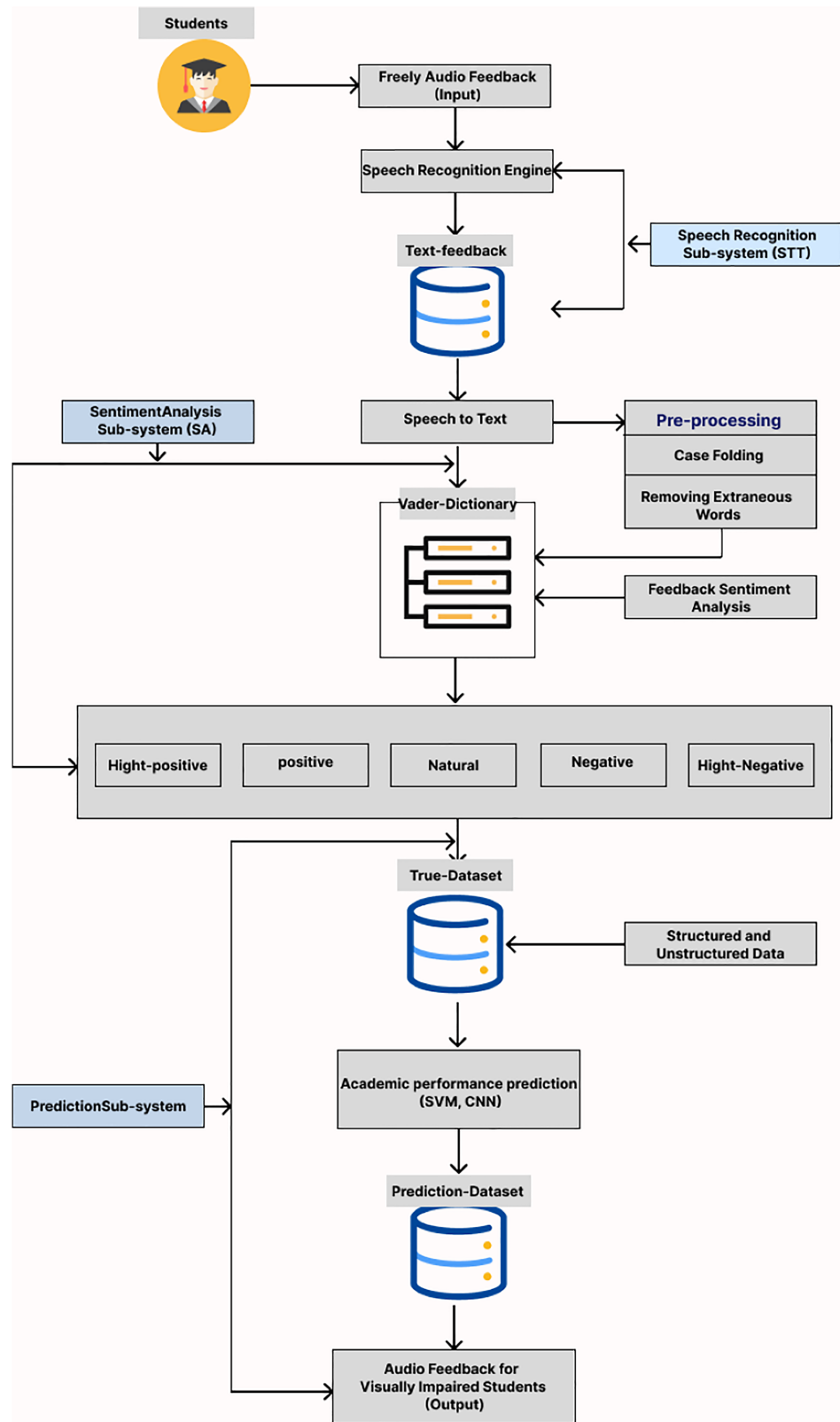


Figure 1 Block diagram of proposal system. This sophisticated system is composed of three integral sub-systems: the Speech Recognition sub-system (STT), which transcribes audio to text; the Emotion

Figure 1 (continued)

Analysis sub-system (EA), which evaluates the emotional content of the feedback; and the Prediction sub-system, which anticipates students' academic performance based on their emotional feedback. Created in Figma. [Full-size !\[\]\(3dfb8d66e81160ad61421a3452093d1b_img.jpg\) DOI: 10.7717/peerj-cs.2143/fig-1](https://doi.org/10.7717/peerj-cs.2143/fig-1)

- Voice Activity Detection (VAD): Using “PyDub’s” silence detection, we filtered segments of interest by analyzing audio at 16 kHz, with a sensitivity threshold set to -30 dB.

Speech-to-text configuration

- Library setup: The SpeechRecognition Python library version 3.8.1 was employed, configuring the recognizer as follows:
- Audio source settings: Files were configured to be in WAV format, with a 16,000 Hz sampling rate and mono channel.
- Ambient noise adjustment: The system calibrated itself using the first 0.5 s of each audio file to adapt to ambient noise levels.
- Language and accent model: The recognizer was optimized for U.S. English while being adaptable to regional accents.
- Timeout and confidence thresholds: A timeout was set to 10 s, and the confidence threshold was fixed at 0.75 to ensure transcription reliability.

Speech-to-text processing

The “InitializeSpeechRecognitionEngine” function configures the speech recognition engine to handle various speech attributes like accents, speech patterns, and diction. The function “ConvertToText” then processes each audio file, converting the audio content into text based on the configured parameters. The “AdjustForAmbientNoise” function ensures the system adapts to varying noise conditions, further enhancing recognition accuracy.

Text data management

The transcribed text is stored in the “analysis sentiments” dataset using the “SaveToDataset” function, enabling further analysis. This structured repository allows visually impaired students to interact with the system more effectively and ensures data is available in a structured format for subsequent sentiment analysis.

Sentiment analysis sub-system

The Sentiment Analysis (SA) sub-system is designed to process and interpret the textual feedback previously transcribed by the Speech-to-Text sub-system. This analysis is conducted in two primary phases: Feedback Pre-processing and Feedback Sentiment Analysis.

Feedback pre-processing

This phase is crucial for preparing textual feedback for detailed sentiment analysis. The feedback undergoes several refinement steps to ensure it is standardized and optimized for sentiment extraction:

- Text normalization: Initial processing involves converting all feedback to a uniform lowercase format (case folding), ensuring consistency across the data and eliminating variations caused by text case.
- Removing extraneous words: Common stopwords that typically do not carry significant sentiment value are filtered out. These include frequent but non-informative words such as “the”, “is”, and “at”. This step is crucial as it helps to highlight sentimentally relevant words in the feedback.
- Text cleaning: The feedback is further cleaned by removing punctuation and applying lemmatization to words. Lemmatization helps in converting words to their base or root form, thus standardizing various forms of a word into a single representation. This reduces textual noise and aids in the uniformity of the analysis.

Feedback sentiment analysis

The choice of VADER for the sentiment analysis was informed by its proven effectiveness in handling brief texts similar to student feedback, which often mirrors the concise and direct nature of social media texts. Citing the study by *Borg & Boldt (2020)*, VADER’s superior accuracy over other machine learning algorithms like SVMs was a key factor in its selection. This tool’s efficiency is particularly valuable in interpreting uniquely concise and straightforward feedback from visually impaired students.

- Tool configuration: The VADER tool (Python library version 3.2.1) is configured with default parameters and further calibrated on a custom validation set to ensure it suits the specific nuances of our dataset.
- Sentiment scoring: Feedback is categorized into five sentiment categories based on VADER’s compound score, which assesses the intensity of sentiment in the text:
 - High Positive: Compound score ≥ 0.6
 - Positive: $0.1 \leq$ Compound score < 0.6
 - Neutral: $-0.1 \leq$ Compound score < 0.1
 - Negative: $-0.6 \leq$ Compound score < -0.1
 - High Negative: Compound score ≤ -0.6

This methodological approach allows for an efficient and nuanced analysis of student sentiments, classifying them into distinct groups and providing a comprehensive insight into their emotional states. This, in turn, enhances our ability to effectively assess and respond to the educational needs of visually impaired students.

Prediction sub-system

The Prediction Sub-system employs advanced machine learning techniques, specifically convolutional neural networks (CNN) and support vector machines (SVM), to predict the

performance and analyze sentiment trends of visually impaired students at critical times during the educational process. Both CNN and SVM have shown considerable effectiveness in educational data analytics and sentiment analysis applications, demonstrating robust capabilities in classifying text and image data ([Li et al., 2014](#)).

SVM technique

Noted for its efficiency in handling educational data, SVM is a binary classification technique that adapts well to various educational parameters. This adaptability, coupled with shorter model training times and heightened sensitivity in modeling dependent variables, makes SVM ideal for predicting student performance ([Ifenthaler & Widanapathirana, 2014](#); [Chikersal, Poria & Cambria, 2015](#)).

Python libraries and tools for SVM prediction:

- Scikit-learn: Used for building the SVM model, preprocessing data, and for splitting the dataset into training and testing sets.
- Pandas: Utilized for data manipulation and merging datasets.
- NumPy: Employed for numerical operations, especially for handling arrays and matrices during data preprocessing.
- Joblib/Pickle: For saving the model state and ensuring that the predictions can be reproduced later.

CNN technique

CNNs have excelled in classification tasks by using convolution and pooling layers, which effectively capture important attribute dependencies in both text ([Kim, 2014](#); [Kim et al., 2017](#)) and image domains ([Krizhevsky, Sutskever & Hinton, 2012](#)). Moreover, recent studies like those by [Li et al. \(2022\)](#), [Khan et al. \(2022\)](#) and [Banea, Mihalcea & Wiebe \(2013\)](#) have demonstrated the utility of CNNs in sentiment analysis, achieving high accuracy in nuanced linguistic contexts.

Python libraries and tools for CNN prediction:

- TensorFlow/Keras: For building and training the CNN model.
- Scikit-learn: Used for SVM model building, data preprocessing, and for splitting the dataset into training and testing subsets.
- Pandas: Utilized for data manipulation and merging datasets.
- NumPy: Employed for numerical operations, especially for handling arrays and matrices during data preprocessing.
- TfidfVectorizer: For converting text data into numerical vectors based on TF-IDF statistics.

In this study, we integrated a simple CNN model with a single convolution layer to analyze both structured and unstructured data, as detailed in [Table 2](#). This methodology was tailored for the early detection of students at risk of underperforming, with evaluations conducted during the fourth and ninth weeks of the course. To ensure accuracy, we

excluded data collected before the fourth week due to insufficient information and omitted sessions occurring after the ninth week to provide timely feedback for student improvement. The performance of our model was assessed using recall (Eq. (1)), precision (Eq. (2)), and the F-measure (Eq. (3)), widely accepted metrics in fields such as information retrieval and machine learning.

$$\text{recall} = \frac{\text{number of failed(or pass)student identified by the method correctly}}{\text{total number of failed (or passed)student}} \quad (1)$$

$$\text{precision} = \frac{\text{number of failed(or pass)student identified by the method correctly}}{\text{total number of failed (or passed)student identified by the method}} \quad (2)$$

$$F_measure = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (3)$$

This comprehensive approach not only enhances our predictive capabilities but also ensures that our interventions are timely and based on robust analytical insights, facilitating effective support for visually impaired students.

Ethical statement

The ethical committee in our Kafrelsheikh University has reviewed the study protocol and ethically approved the study under reference No. 334-38-44972-SD.

RESULTS AND VISUALIZATION

Our study's central objective was to craft an intelligent system capable of discerning and analyzing the sentiment states of visually impaired students, thereby augmenting their educational experience and performance assessments. In pursuit of this goal, we addressed two fundamental questions:

1. Q1: How do the specific features of the proposed intelligent system cater to the unique needs of visually impaired students in virtual learning environments?
2. Q2: In what ways does the intelligent system impact the academic engagement and performance outcomes of visually impaired students, and what evidence supports this impact?

The following section will present detailed results and insights derived from the experimental phase, offering a comprehensive evaluation of the system's capabilities and its impact on the educational outcomes of visually impaired students.

Sentiment expectation

In our study, the VADER lexicon-based approach was employed to categorize the sentiments within the freely given feedback from visually impaired students. This analysis occurred in both Week 4, with 2,648 feedback entries, and Week 9, with 4,237 entries. A noticeable increase in feedback by Week 9 indicated enhanced engagement among students. The sentiment categorization of this feedback is detailed in [Table 3](#).

[Table 3](#) presents a breakdown of emotional sentiment in feedback posts, categorized into five types: high negative, negative, natural, positive, and high positive. The data is organized into two sets, with the first column showing the number of feedback and their

Table 3 Vader dictionary results.

Sentiment type	Num. of feedback (Total: 2,648)	Percentage (%)	Num. of feedback (Total: 4,237)	Percentage (%)
High negative	91	2.5	5	0.12
Negative	1,001	37.8	59	1.4
Natural	911	34.5	1,109	26.2
Positive	539	20.5	1,622	38.5
High positive	106	4.0	1,626	38.5

Note:

Emotional sentiment in feedback posts categorized into five types: High Negative, Negative, Natural, Positive, and High Positive. The data is organized into two sets, with the first column showing the number of feedback and their corresponding percentages out of a total of 2,648 feedback, and the second column displaying the same for a total of 4,237 feedback.

Table 4 The comparison of sentiment categorization.

Sentiment categories	Expert 1	Expert 2	Expert 3	VADER	Average expert count	Discrepancy
High positive	45	50	55	48	50	2
Positive	85	90	95	87	90	3
Neutral	150	160	170	165	160	5
Negative	110	105	100	108	105	3
High negative	60	55	50	52	55	3

Note:

The number of feedback entries each expert and VADER placed in these categories. This also includes the 'Average Expert Count' for each category, which is the average number the experts assigned to each sentiment type. The 'Discrepancy' column shows the difference between the experts' average count and VADER's count for each sentiment.

corresponding percentages out of total of 2,648 feedback, and the second column displaying the same for a total of 4,237 feedback. The table effectively illustrates the distribution of sentiment types among the feedback posts, highlighting the most and least frequent sentiment expressed.

To validate the accuracy of this lexicon-based approach, a panel of (3) experts specializing in student sentiment analysis was engaged. They analyzed a select sample of student comments, guided by a comprehensive manual specifically developed for this purpose (Appendix (1)). This manual contained detailed instructions to aid the experts in their analysis, categorizing sentiment into five distinct classes as per the prevailing research. This comparison was conducted on a select sample of 500 feedback entries [Table 4](#) shows the comparison of sentiment categorization. The precision of the experts' assessments was quantified using the following formula:

$$accuracy = \frac{\text{Number of comments with accurately identified sentiment types}}{\text{Total number of comments}} \quad (4)$$

[Table 4](#) displays details of a Preliminary comparative analysis of the sentiment categorization by experts and the VADER system. This shows to what extent the machine categorization conforms to human judgment and it is highlighted by showing discrepancies in categorization. The feedback is divided into five types: High Positive, Positive, Neutral, Negative, and High Negative. The table shows how many feedback

entries each expert and VADER placed in these categories. The table also shows the 'Average Expert Count' for each category, which is the average number the experts assigned to each sentiment type. Another important feature of the table is the 'Discrepancy' column, which shows the difference between the experts' average count and VADER's count for each sentiment.

The analysis of these discrepancies offers valuable insights into the reliability and alignment of automated sentiment analysis tools like VADER when compared with human judgment. This is especially pertinent in the context of feedback from visually impaired students, where understanding the sentimental landscape is crucial. The most and least frequent sentiments expressed, as identified by both the experts and VADER, are highlighted, providing a comprehensive understanding of the sentiments prevalent in the students' feedback.

The comparison of VADER's results with the ground truth established through human evaluation yielded an accuracy rate of 76%. This demonstrates a considerable level of reliability in VADER's sentiment identification capabilities, affirming its suitability for analyzing feedback in educational settings involving visually impaired students.

The expectations of academic performance

The preliminary analysis leveraged a rich array of features, including advanced machine learning algorithms and both structured and unstructured data.

In our study, we utilized advanced machine learning algorithms to evaluate the potential for early prediction of student success at the 4th and 9th-week marks within a 12-week academic term. As outlined in [Table 2](#) of our findings, we analyzed both structured data (such as student attendance records, participation in class, and homework completion rates) and unstructured data, the latter comprising sentimental tones derived from student remarks, as assessed by human annotators. We divided our dataset, consisting of information from 100 students, evenly into training and test groups, following a 50:50 split this decision was based on the unique characteristics of our dataset, which includes a relatively small sample size and a high level of diversity in the audio feedback from visually impaired students.

We acknowledge that while splits like 70:30 or 80:20 are more common in larger datasets, our decision for an even split was driven by the need to maximize the data available for testing, ensuring robust model validation. Given the novelty of our research area and the specific challenges posed by our dataset, a 50:50 split provided a more balanced approach to training and validation, allowing for a thorough assessment of the model's performance on an equally representative test set.

[Figure 2](#) displays our experimental results, demonstrating that our model was highly effective. By the fourth week, it had achieved an F-measure of 0.85 using SVM and 0.86 with CNN. By the ninth week, its performance was further improved, achieving an F-measure of 0.89 using SVM and 0.92 using CNN. The performance of the CNN-based approach was particularly impressive; it achieved the highest F-measure of 0.92, which underscores its superior effectiveness in our context.

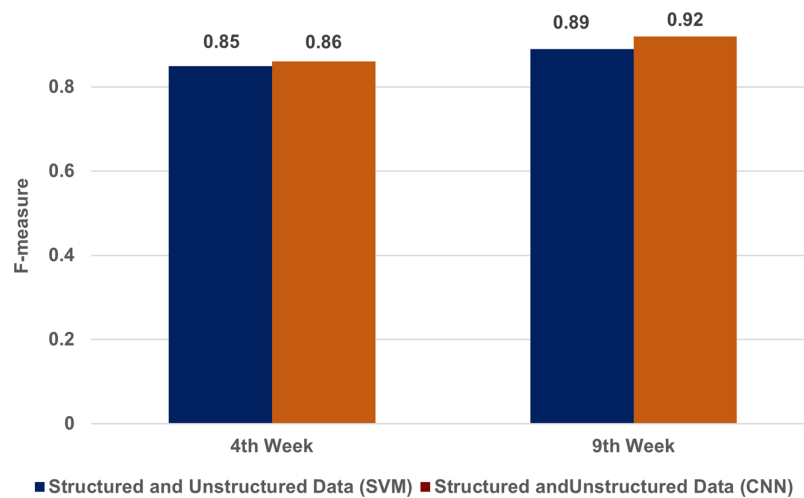


Figure 2 A comparison of structured and unstructured data in SVM & CNN techniques. Our experimental results, demonstrating that our model was highly effective. By the fourth week, it had achieved an F-measure of 0.85 using support vector machines (SVM) and 0.86 with convolutional neural networks (CNN). By the ninth week, its performance was further improved, achieving an F-measure of 0.89 using SVM and 0.92 using CNN. [Full-size !\[\]\(5fd6ef84f97f42d7f8b34275f1b65312_img.jpg\) DOI: 10.7717/peerj-cs.2143/fig-2](https://doi.org/10.7717/peerj-cs.2143/fig-2)

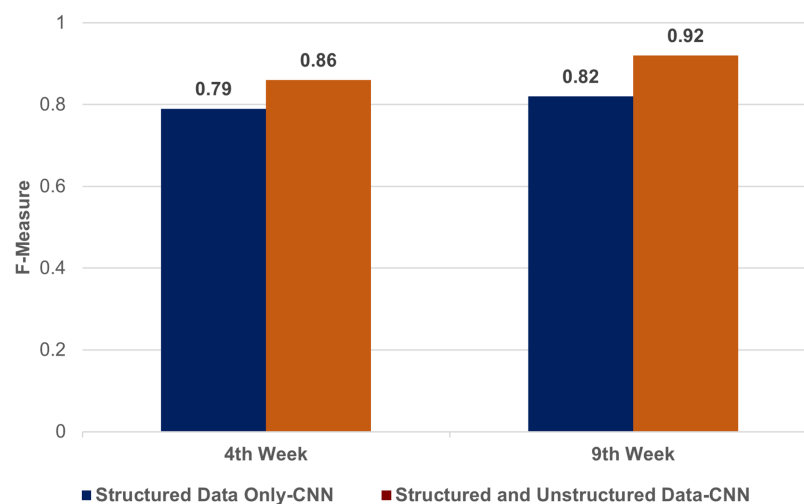


Figure 3 A comparison of structured data only & structured and unstructured data in CNN. A comparison of the F-measures obtained by a CNN model with and without the inclusion of unstructured data to highlight the added value of emotion analysis. The discernible improvement in F-measure when unstructured data is included indicates the effectiveness of utilizing self-rated feedback for early-stage predictions. [Full-size !\[\]\(b8ddfb9d90db8697d6b8ef7f72522b2e_img.jpg\) DOI: 10.7717/peerj-cs.2143/fig-3](https://doi.org/10.7717/peerj-cs.2143/fig-3)

Figure 3 compares the F-measures obtained by a CNN model with and without the inclusion of unstructured data to highlight the added value of sentiment analysis. The discernible improvement in F-measure when unstructured data is included indicates the effectiveness of utilizing self-rated feedback for early-stage predictions.

Table 5 shows accuracy of the SVM and CNN models over two periods of assessment is explained with these experimental results, while the table compares performance with that of structured and unstructured data to portray the impact of data types on prediction

Table 5 Results of experiments with CNN and SVM.

	Results of SVM (accuracy)		Results of CNN (accuracy)	
	9th week	4th week	9th week	4th week
Structured data only	0.87	0.76	0.82	0.79
Structured and unstructured data	0.89	0.85	0.92	0.86

Note:

The empirical evidence suggests that incorporating emotional feedback from visually impaired students markedly enhances the early predictive accuracy of our model.

accuracy. thus, the empirical evidence suggests that incorporating sentimental feedback from visually impaired students markedly enhances the early predictive accuracy of our model. Such insights pave the way for educators to swiftly adapt their pedagogical approaches, optimizing educational support for visually impaired students in real-time.

Class-wise performance analysis of SVM and CNN

We have conducted an in-depth class-wise performance analysis of the SVM and CNN algorithms. This analysis is essential given the notable imbalance in our dataset, particularly the limited instances in classes such as high negative. So, we present a series of heatmaps [Figs. 4A–4H](#). These figures provide the performance metrics of SVM and CNN under different data conditions, and assessment weeks, showing the models' effectiveness in identifying the “at-risk” students by giving precision, recall, and F1-scores for “High Negative” class sentiment.

In our detailed class-wise performance analysis of the SVM and CNN algorithms, we address the challenges presented by the imbalance in our dataset, particularly the underrepresentation of crucial categories such as ‘high negative’. [Figures 4A–4D](#) illustrate the performance of the CNN algorithm during the 4th and 9th weeks, for structured data only and structured with unstructured data, and [Figs. 4E–4H](#) show similar analyses for the SVM algorithm.

The SVM achieved a precision of 0.8 for the ‘high negative’ class, suggesting that it was highly accurate in its predictions of at-risk students. Meanwhile, the recall for this class was 0.79 in the 9th week when analyzing combined structured and unstructured data, indicating that the SVM successfully identified a significant portion of the actual ‘high negative’ instances. This combination resulted in an F1-score of 0.85, reflecting a balanced harmonic mean of precision and recall, thus illustrating the model’s effectiveness at this stage.

Contrastingly, CNN’s performance in the 9th week when analyzing combined structured and unstructured data showed a precision of 0.74 for the ‘high negative’ class, a slight decrease compared to the SVM. However, the recall was 0.79, maintaining a strong ability to identify students at risk. The F1-score for this condition was 0.86, slightly higher than that of the SVM, suggesting that the CNN may be more adept at handling a combination of data types at this later stage.

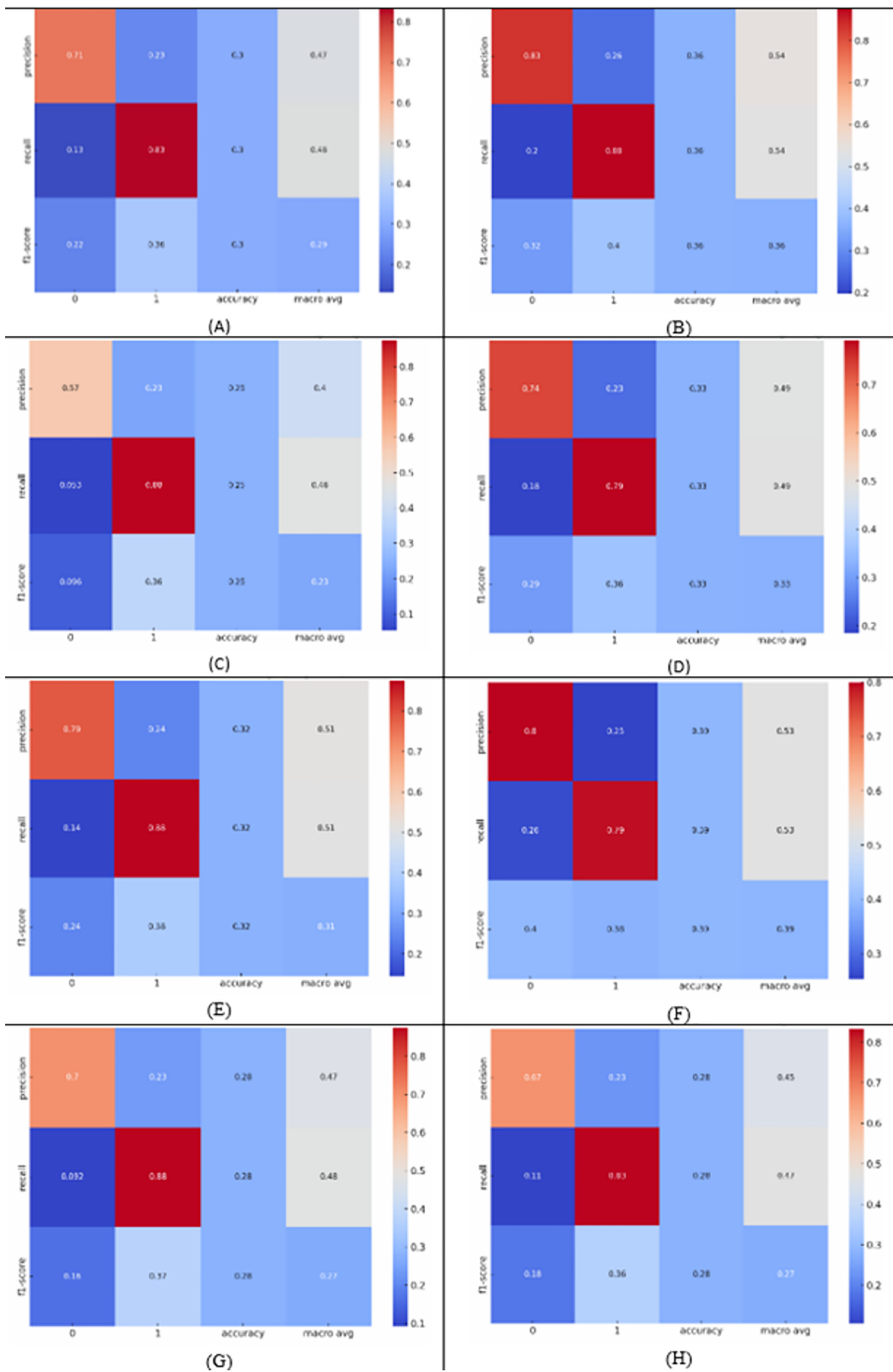


Figure 4 Class-wise performance analysis of SVM and CNN. (A-D) illustrate the performance of the CNN algorithm during the 4th and 9th weeks, for structured data only and structured with unstructured data, and (E-H) show similar analyses for the SVM algorithm.

Full-size DOI: 10.7717/peerj-cs.2143/fig-4

These class-specific metrics allow us to draw direct correlations between the models' predictive capabilities and their potential impact on the educational support for visually impaired students. The high recall rates are particularly important in educational settings where failing to identify at-risk students could result in missed opportunities for timely intervention. Conversely, the high precision rates are crucial for ensuring that resources are directed appropriately, avoiding the unnecessary allocation of interventions to students who do not require them.

Correlation analysis of teaching aspects and student sentiments

This subsection explores the relationships between various teaching aspects such as homework grades, clicks, attendance, and discussion participation, and their correlations with different student sentiment categories. Understanding these correlations is vital in assessing the impact of teaching methodologies on the sentimental and academic experiences of visually impaired students in virtual learning environments.

Figure 5 visually represents these correlations, using darker colors to indicate stronger correlations and lighter colors to show weaker ones.

According to Fig. 5, positive sentiments have a moderate positive correlation of 0.58 with structured data like attendance and homework grades. This correlation indicates that students who exhibit positive sentiments tend to have better academic outcomes, as reflected in higher attendance and improved homework performance.

Impact of Negative Sentiments. The analysis shows that negative sentiments have a weak correlation (0.1 or less) with academic metrics, suggesting that external factors significantly impact the academic performance of students with negative sentiments. The moderate correlation between positive sentiments and academic metrics emphasizes the importance of fostering a positive emotional environment that promotes engagement and boosts student success. While the weak correlation between negative sentiments and academic metrics highlights the need for additional non-academic support services to address the external factors affecting student performance.

Regression model for predicting student performance

This subsection presents a regression model summary, highlighting how various academic and engagement (structured data) factors collectively influence the performance of visually impaired students in a virtual learning setting. Fig. 6 shows Regression model outputs showing predictors of student performance in virtual learning environments. This model quantifies the impact of several variables such as attendance and homework on students' academic outcomes.

The regression analysis aimed to correlate the specific features of the sentiment analysis system, such as emotional granularity, with learning outcomes. The model's correlation coefficient (R) of 0.507 indicates a moderate positive relationship between predicted and actual performance, showing that the features in the sentiment analysis system and other variables are useful in estimating student outcomes. Also, the R square value of 0.257 indicates that 25.7% of the variance in student performance can be explained by the

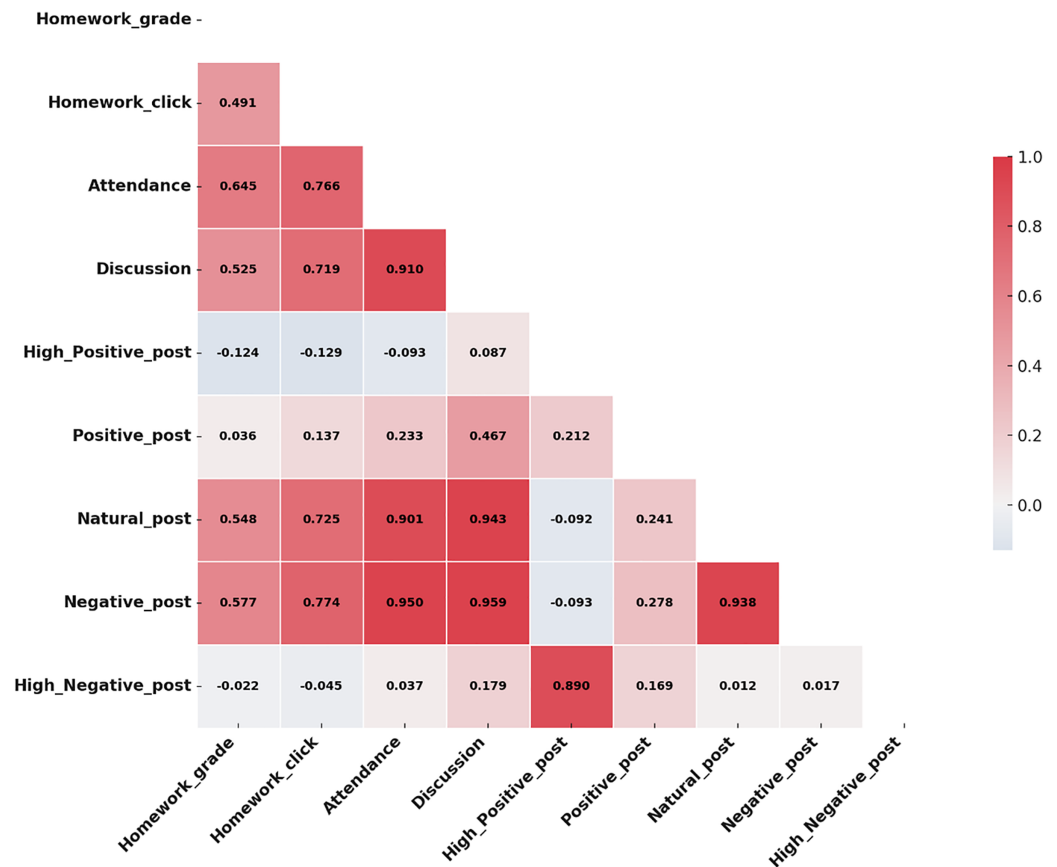


Figure 5 Correlation analysis of teaching aspects and student sentiments. A moderate positive correlation between homework clicks and discussion participation suggests that increased engagement in online activities may enhance students' active involvement in discussions.

Full-size DOI: [10.7717/peerj-cs.2143/fig-5](https://doi.org/10.7717/peerj-cs.2143/fig-5)

variables used in the model, suggesting that these features provide valuable insights but that other unexplored factors also play a significant role.

Adjusted R square, at 0.226, provides a more accurate representation of the model's ability to explain the variance in performance by accounting for the number of predictors. The Standard Error of the estimate, at 0.378, also shows the average deviation of observed values from the regression line, highlighting moderate prediction accuracy.

Together, these metrics offer a comprehensive overview of the factors influencing student performance in virtual learning environments, emphasizing the need for targeted strategies to enhance the educational experience for visually impaired students.

Comparison of the proposed system with the related works

Evaluate the effectiveness of the sentiment analysis models proposed herein, respectively, with structured and unstructured data in comparison to the latest documented most recent benchmark models available in the literature. [Table 6](#) reports the outcomes of such a comparative assessment, along with the key performance metrics, precision, recall, accuracy, and F1-measure.

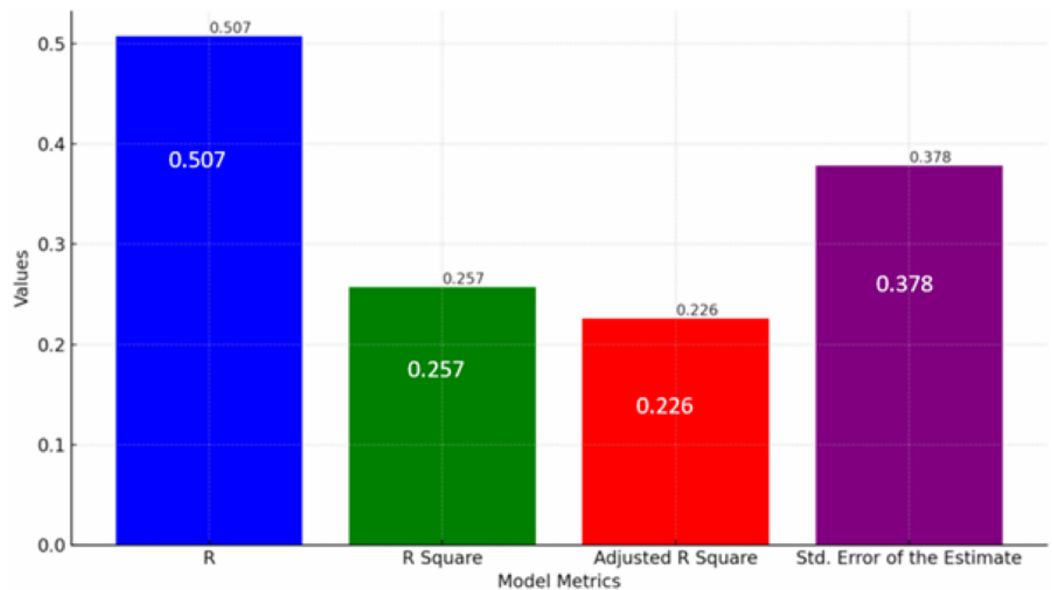


Figure 6 Regression model for predicting student performance. The model indicates that approximately 25.7% of the variance in student performance can be explained by the independent variables included in the study, as suggested by the R Square value. While this percentage is moderate, it underscores the significance of factors like attendance and homework engagement (structured data) in academic success. [Full-size !\[\]\(b345a1c4255362eec3746050dd71ccac_img.jpg\) DOI: 10.7717/peerj-cs.2143/fig-6](https://doi.org/10.7717/peerj-cs.2143/fig-6)

Table 6 Comparison of the proposed system with the related works.

Reference	Techniques	Accuracy	Precision	Recall	F1-measure
<i>Singh et al. (2021)</i>	SVM	85.69%	–	–	–
<i>Khaiser, Saad & Mason (2023)</i>	SVM	78.00%	–	–	–
<i>Antit, Mechti & Faiz (2022)</i>	TunRoBERTa, CNN	80.60%	83.90%	80.60%	81.10%
<i>Basiri et al. (2021)</i>	CNN, RNN, attention mechanism	89.70%	91.50%	88.30%	89.80%
<i>Jimmy & Prasetyo (2022)</i>	SVM	89.00%	–	–	–
<i>Huspi & Ali (2024)</i>	CNN	85.00%	84.00%	84.00%	85.00%
Proposed system	Structured and unstructured data (SVM)	89.00%	93.00%	88.00%	91.00%
	Structured and unstructured data (CNN)	92.00%	95.00%	92.00%	93.00%

Note:

Evaluate the effectiveness of the sentiment analysis models proposed herein, respectively, with structured and unstructured data in comparison to the latest documented most recent benchmark models available in the literature.

From [Table 6](#), it is seen that the proposed model gives the best performance compared to other methods reported in the literature. The SVM-based model with structured and unstructured data manages an accuracy rate of 89.00%, bringing to a closing gap with the results cited by [Jimmy & Prasetyo \(2022\)](#). The work of [Sharma & Prasetyo \(2022\)](#), however, reveals great improvement compared to other conventional models like that belonging to [Singh et al. \(2021\)](#) and [Khaiser, Saad & Mason \(2023\)](#), who had shown an accuracy of 85.69% and 78.00% consecutively. Our CNN-based approach, in particular, further improves this performance by measuring the remarkable accuracy of 92.00%, which outperforms the combination proposed in TunRoBERTa-CNN by [Antit, Mechti & Faiz](#)

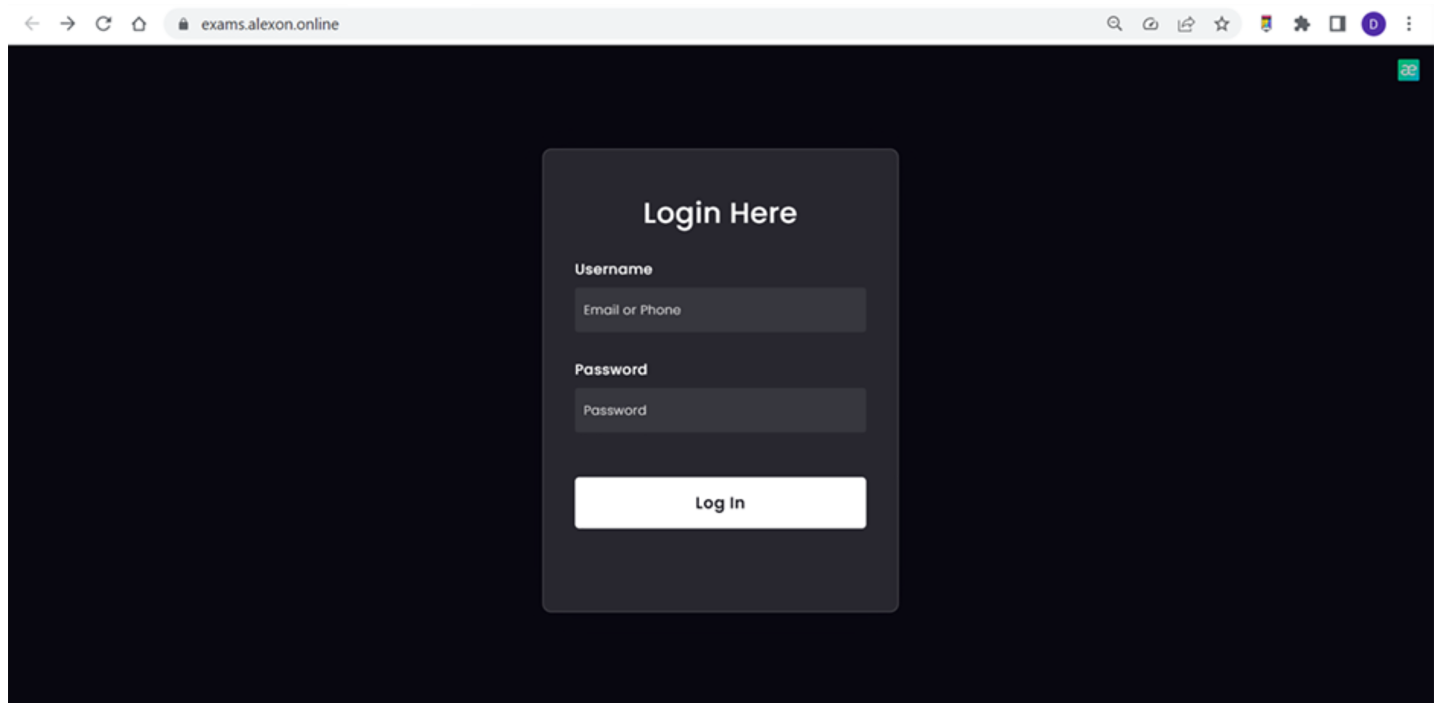


Figure 7 The login interface of the proposed system which requires users to enter their username and password to access their student profile. [Full-size](#) DOI: 10.7717/peerj-cs.2143/fig-7

(2022) (80.60%) (2022) and the state-of-the-art accuracy of 89.70% reported for the novel CNN-RNN-Attention mechanism of Bas.

This obvious improvement is a clear indicator that structured and unstructured data incorporation is beneficial, as it adds to the learning context of the model and makes it richer for the understanding of sentiments towards improved performance of classification overall. Developed especially for the CNN model, the existing work shows a significant increase in precision and recall, reaching the values for scores at 95.00% and 92.00%, respectively, which are the most valuable metrics to express reliability and robustness in the practical use of sentiment analysis. The excellent performance of the proposed models thus not only highlights the potential of using advanced machine learning techniques in sentiment analysis but also sets a new bar for future research in this area. These results, therefore, would seem to indicate that digging deeper into the combined use of structured and unstructured data has the potential for even more powerful models to be created, which are armed with the ability to cope with a real data environment that is characterized by many types of complexities.

Information visualization

Figure 7 shows a screenshot from the login interface of the system, which requires users to enter their username and password to access their student profile. It was made very accessible by using speech and text, hence particularly helpful for students with visual

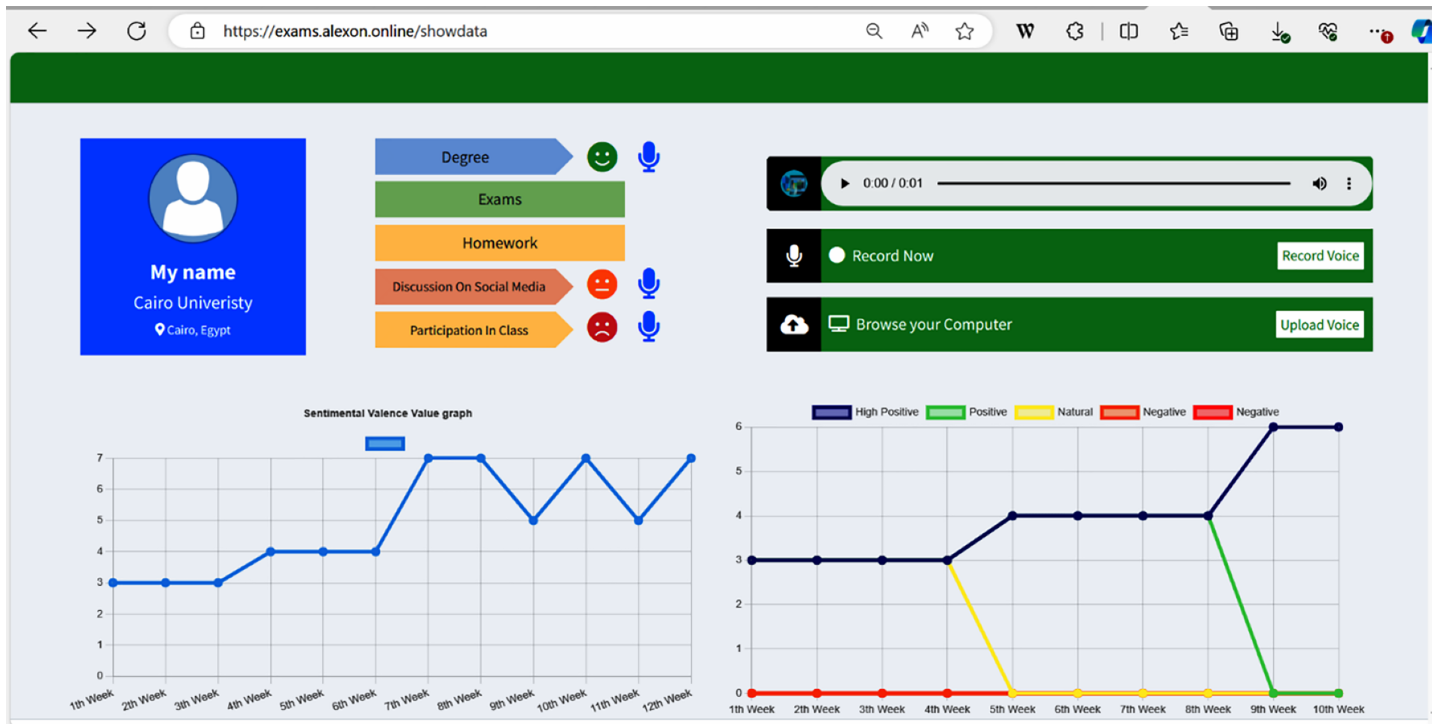


Figure 8 The student profile. The structured and unstructured data are integrated to construct a multifaceted student profile. It features a dashboard offering a holistic view of academic assessments, tailored diagnostics, activity recommendations, and weekly emotional states.

Full-size DOI: 10.7717/peerj-cs.2143/fig-8

impairments. The proposed system, depicted in Fig. 8, integrates structured and unstructured data to construct a multifaceted student profile. It features a dashboard offering a holistic view of academic assessments, tailored diagnostics, activity recommendations, and weekly sentimental states. The system employs auditory cues for conveying grades, social engagement, and class participation, enabling visually impaired students to receive personalized feedback. Notably, it employs sentiment analysis on self-assessment feedback, enhancing the early identification of at-risk students. This predictive capability is crucial for timely educational interventions and supports positive self-regulation learning practices. The system's innovative use of sentimental valence graphs further aids students in reflecting on their learning journey, fostering an environment conducive to academic success.

The study's main contributions are twofold. First, early identification of at-risk visually impaired students is crucial to the success of educational interventions aimed at changing the course of their academic performance. The proposed system provides diagnostic and suggested reports that offer useful information for positive sentiment self-regulation learning. The audio line graph of valence values also provides vital information for providing external guidance to students experiencing negative sentiments.

FINDINGS AND CONCLUSIONS

This study demonstrates the effectiveness of automated sentiment analysis in interpreting visually impaired students' audio feedback within virtual learning environments. By categorizing sentiments into five distinct groups, this research provides nuanced insights into the students' sentimental responses, significantly enhancing the understanding of their learning experiences.

The primary contributions of this research are manifold. Firstly, the early detection of students potentially at risk of academic challenges is facilitated. This is pivotal for implementing timely and effective educational interventions. Secondly, the study's innovative approach in integrating both structured (like attendance, and homework completion) and unstructured data (audio feedback sentiments) offers a more holistic view of student engagement and performance. This integration is particularly valuable in the context of special educational needs.

However, the study acknowledges certain limitations. The sample size, confined to a specific group from selected universities, may not fully represent the broader population of visually impaired students. Additionally, the sentiment analysis tool used might have specific constraints in fully capturing the nuances of audio feedback.

Future research should aim to validate these findings across a more diverse and larger sample size. Further refinement of sentiment analysis tools to better interpret audio feedback, considering factors like speech intonation and pauses, could enhance the accuracy and applicability of this model. Continual advancements in this field could significantly contribute to personalized and inclusive educational strategies for visually impaired students.

RECOMMENDATIONS

This study confirms the potential of sentiment analysis to enhance the academic achievements of visually impaired students. It suggests that the creation of such intelligent speech prediction systems that provide subtle accommodations can effectively be converted into tools for broader usage inside educational environments. This developed tool "VADER" for sentiment analysis would be used in different learning and teaching settings, including distance learning contexts and face-to-face environments for both groups of students. These underline a way in which the potential of improved engagement and educational results underpin the requirement of educational systems to advance analytics in the teaching process.

The educators have a role to play and should provide the necessary support, which does not make them weak but rather sustains high academic standards. Some strategies he suggested include the use of predictive analytics in interactive learning that would trace student involvement proactively and assist students likely to underperform. Furthermore, assessment activities should be adaptable in space, time, and support, and information utilized in the production of educational material might have to be transformed to suit its required accessible format.

Such technologies should be expounded in educational institutions to cater to a broader perspective of disabilities so that all students with or without physical capabilities can enjoy

an individualized learning experience. Implementation of these recommendations would ensure more equitable educational opportunities and positively affect the academic outcomes among students from various populations. Only increased awareness and investment in disability education will eliminate discriminating practices. The universities are hereby advised to adopt academic performance prediction systems so that the visually impaired students, and probably any other needy students, are not denied equal education opportunities. They are believed to support better academic functioning of universities and to enhance their capacity to provide an effective and more supportive learning environment for all students.

LIMITATIONS AND FUTURE DIRECTIONS

While the proposed model demonstrates substantial advancements in sentiment analysis for visually impaired students, it is crucial to acknowledge its limitations concerning generalizability across different educational contexts and types of impairment. The study's findings are based on a specific cohort of visually impaired students within virtual learning environments, which may not fully represent the broader population of students with disabilities. Below we explain these limitations:

- *Specific type of sample (visually impaired students):* The model is optimized for audio feedback, which may not be as effective for students with impairments that hinder verbal communication. Developing adaptive models that cater to a variety of disabilities, including those requiring alternative communication forms such as text, sign language, or Braille, is essential for broader applicability.
- *Educational contexts:* The current model's applicability is tested within the confines of virtual learning environments. Its effectiveness in traditional classroom settings, where interactive dynamics differ significantly, remains untested. Extending the application of the model to these and other diverse educational contexts will help evaluate its adaptability and effectiveness more comprehensively.
- *Inclusion of diverse educational populations:* The model's current evaluation is limited to a small sample size (100 students) from selected universities, which may not capture the full diversity of the student population. Future research should apply the model to larger, more diverse datasets across various universities and academic disciplines. Also, this study can be extended to students of various age groups, including high school students and teenagers, to explore its applicability across different educational stages. This expansion is critical for assessing the model's effectiveness across different cultural, linguistic, and educational backgrounds.
- *The language:* This system currently uses English to deliver the course; in the future, it could be adapted to include Arabic or other languages.

By addressing these limitations and strategically targeting areas for improvement, future research can enhance the model's utility and ensure it meets the diverse needs of all students. This approach not only broadens the scope of the model's applicability but also contributes to more inclusive educational practices.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The authors received no funding for this work.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Doaa Mohamed Elbourhamy conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Ethics

The following information was supplied relating to ethical approvals (*i.e.*, approving body and any reference numbers):

The ethical committee in our Kafrelsheikh University has reviewed the study protocol and ethically approved the study under reference No. 334-38-44972-SD.

Data Availability

The following information was supplied regarding data availability:

The raw measurements are available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.2143#supplemental-information>.

REFERENCES

- Agesa RU. 2014.** Teaching strategies for visually impaired students: a guide for educators. *Journal of Special Education Technology* **29**(4):23–35 DOI [10.1177/016264341403000404](https://doi.org/10.1177/016264341403000404).
- Altrabsheh N, Cocea M, Fallahkhaier S. 2015.** Sentiment Analysis in Education: A Case Study on Student Feedback. In: *Proceedings of the 8th International Conference on Educational Data Mining (EDM 2015)*. Madrid, Spain, 254–259.
- Antit C, Mechti S, Faiz R. 2022.** TunRoBERTa: a Tunisian robustly optimized BERT approach model for sentiment analysis. In: *2nd International Conference on Industry 4.0 and Artificial Intelligence (ICIAI 2021)*. Dordrecht: Atlantis Press, 227–231.
- Aslan S, Kızıloluk S, Sert E. 2023.** Analyzing student sentiments using machine learning. *Journal of Educational Technology Research and Development* **71**(3):345–360 DOI [10.1007/s00521-023-08236-2](https://doi.org/10.1007/s00521-023-08236-2).
- Banea C, Mihalcea R, Wiebe J. 2013.** Porting multilingual subjectivity resources across languages. *IEEE Transactions on Affective Computing* **4**(2):211–225 DOI [10.1109/T-AFFC.2013.1](https://doi.org/10.1109/T-AFFC.2013.1).
- Basiri ME, Nemati S, Abdar M, Cambria E, Acharya UR. 2021.** ABCDM: an attention-based bidirectional CNN-RNN deep model for sentiment analysis. *Future Generation Computer Systems* **115**(3):279–294 DOI [10.1016/j.future.2020.08.005](https://doi.org/10.1016/j.future.2020.08.005).
- Bayer J, Bydzovská H, Géryk J, Obsivac T, Popelinsky L. 2012.** *Predicting drop-out from social behaviour of students*. Chania, Greece: International Educational Data Mining Society.

- Blikstein P, Worsley M. 2016.** Multimodal learning analytics and education data mining: using computational technologies to measure complex learning tasks. *Journal of Learning Analytics* 3(2):220–238 DOI 10.18608/jla.2016.32.11.
- Bo P, Lee L. 2008.** Opinion mining and sentiment analysis foundations and trends in information retrieval. Available at <http://www.sentimentanalysis.com>.
- Bonta V, Janardhan NKN. 2019.** A comprehensive study on lexicon-based approaches for sentiment analysis. *Asian Journal of Computer Science and Technology* 8(S2):1–6 DOI 10.51983/ajcst-2019.8.S2.2037.
- Borg A, Boldt M. 2020.** Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications* 162:113746 DOI 10.1016/j.eswa.2020.113746.
- Botchway RK, Jibril AB, Kwarteng MA, Chovancova M, Oplatková ZK. 2019.** A review of social media posts from UniCredit bank in Europe: a sentiment analysis approach. In: *Proceedings of the 3rd International Conference on Business and Information Management*. 74–79.
- Calvo RA, D’Mello S. 2010.** Affect detection: an interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing* 1(1):18–37 DOI 10.1109/T-AFFC.2010.1.
- Chikersal P, Poria S, Cambria E. 2015.** SeNTU: sentiment analysis of tweets by combining a rule-based classifier with supervised learning. In: *SemEval@ NAACL-HLT*. Denver, CO, USA: Association for Computational Linguistics, 647–651.
- Costa EB, Fonseca B, Santana MA, de Araújo FF, Rego J. 2017.** Evaluating the effectiveness of educational data mining techniques for early prediction of students’ academic failure in introductory programming courses. *Computers in Human Behavior* 73:247–256 DOI 10.1016/j.chb.2017.01.047.
- D’Mello S, Jackson T, Craig S, Morgan B, Chipman P, White H, Person N, Kort B, El Kaliouby R, Picard RW, Graesser A. 2008.** AutoTutor detects and responds to learners affective and cognitive states. In: *Workshop on Emotional and Cognitive Issues at the International Conference on Intelligent Tutoring Systems*. 306–308.
- Elbourhamy DM, Najmi AH, Elfeky AIM. 2023.** Students’ performance in interactive environments: an intelligent model. *PeerJ Computer Science* 9(9):e1348 DOI 10.7717/peerj-cs.1348.
- Ghafri HM. 2015.** Enhancing verbal skills through auditory methods: a study on visually impaired students. *International Journal of Educational Research* 45(2):95–110 DOI 10.1016/j.ijer.2015.06.005.
- Gustafsson M. 2020.** Sentiment Analysis in Educational Settings. Available at <http://www.educationsentimentanalysis.com>.
- Huspi SHH, Ali Z. 2024.** Sentiment analysis on roman urdu students’ feedback using enhanced word embedding technique. *Baghdad Science Journal* 21(2 (SI)):0725-0725 DOI 10.21123/bsj.2024.9822.
- Hutto C, Gilbert E. 2014.** Vader: a parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media* 8(1):216–225 DOI 10.1609/icwsm.v8il.14550.
- Ifenthaler D, Widanapathirana C. 2014.** Development and validation of a learning analytics framework: two case studies using support vector machines. *Technology, Knowledge and Learning* 19(1):221–240 DOI 10.1007/s10758-014-9226-4.
- Jayaprakash SM, Moody EW, Lauría EJ, Regan JR, Baron JD. 2014.** Early alert of academically at-risk students: an open source analytics initiative. *Journal of Learning Analytics* 1(1):6–47 DOI 10.18608/jla.2014.11.3.

- Jimmy J, Prasetyo VR. 2022.** Sentiment analysis on feedback of higher education teaching conduct: an empirical evaluation of methods. In: *AIP Conference Proceedings*. Vol. 2470, AIP Publishing.
- Khaiser FK, Saad A, Mason C. 2023.** Sentiment analysis of students' feedback on institutional facilities using text-based classification and natural language processing (NLP). *Journal of Language and Communication* **10(1)**:101–111 DOI [10.47836/jlc](https://doi.org/10.47836/jlc).
- Khan L, Amjad A, Afaq KM, Chang HT. 2022.** Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu text shared in social media. *Applied Science* **12(5)**:2694 DOI [10.3390/app12052694](https://doi.org/10.3390/app12052694).
- Kim Y. 2014.** Convolutional neural networks for sentence classification. *ArXiv preprint* DOI [10.48550/arXiv.1408.5882](https://doi.org/10.48550/arXiv.1408.5882).
- Kim M, Cao B, Mau T, Wang J. 2017.** Speaker-independent silent speech recognition from flesh-point articulatory movements using an LSTM neural network. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **25(12)**:2323–2336 DOI [10.1109/TASLP.2017.2758999](https://doi.org/10.1109/TASLP.2017.2758999).
- Kiritchenko S, Mohammad SM. 2017.** Capturing reliable fine-grained sentiment associations by crowdsourcing and best-worst scaling. *ArXiv preprint*. DOI [10.48550/arXiv.1712.01741](https://doi.org/10.48550/arXiv.1712.01741).
- Krizhevsky A, Sutskever I, Hinton GE. 2012.** Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*. Red Hook, NY, USA: Curran Associates, Inc., Vol. 25.
- Li D, Ahmed K, Zheng Z, Mohsan SAH, Alsharif MH, Hadjouni M, Jamjoom MM, Mostafa SM. 2022.** Roman Urdu sentiment analysis using transfer learning. *Applied Sciences* **12(20)**:10344 DOI [10.3390/app122010344](https://doi.org/10.3390/app122010344).
- Li J, Ott M, Cardie C, Hovy E. 2014.** Towards a general rule for identifying deceptive opinion spam. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 1566–1576.
- Liu B. 2015.** Many facets of sentiment analysis. In: *A Practical Guide to Sentiment Analysis*. Cham: Springer, 11–39.
- Mohammdi HM, Elbourhamy DM. 2021.** An intelligent system to help deaf students learn Arabic sign language. *Interactive Learning Environments* **31(2)**:1–16 DOI [10.1080/10494820.2021.1910318](https://doi.org/10.1080/10494820.2021.1910318).
- Mwakyaja BM. 2013.** Teaching students with visual impairments in inclusive classrooms: a case study of one secondary school in Tanzania (Master's thesis).
- Nguyen T, Phung D, Dao B, Venkatesh S, Berk M. 2014.** Affective and content analysis of online depression communities. *IEEE Transactions on Affective Computing* **5(3)**:217–226 DOI [10.1109/TAFFC.2014.2315623](https://doi.org/10.1109/TAFFC.2014.2315623).
- Office of the Higher Education Commission. 2013.** Guidelines for enhancing educational quality for students with disabilities. Thailand: Ministry of Education Available at https://en.wikipedia.org/wiki/Office_of_the_Higher_Education_Commission.
- Pekrun R, Goetz T, Titz W, Perry RP. 2002.** Academic sentiments in students' self-regulated learning and achievement: a program of qualitative and quantitative research. *Educational Psychologist* **37(2)**:91–105 DOI [10.1207/S15326985EP3702_4](https://doi.org/10.1207/S15326985EP3702_4).
- Pritchard ME, Wilson GS. 2003.** Using emotional and social factors to predict student success. *Journal of College Student Development* **44(1)**:18–28 DOI [10.1353/csd.2003.0008](https://doi.org/10.1353/csd.2003.0008).
- Rodrigues RG, das Dores RM, Camilo-Junior CG, Rosa TC. 2016.** SentiHealth-cancer: a sentiment analysis tool to help detecting mood of patients in online social networks. *International Journal of Medical Informatics* **85(1)**:80–95 DOI [10.1016/j.ijmedinf.2015.09.007](https://doi.org/10.1016/j.ijmedinf.2015.09.007).

- Saif H, Fernandez M, He Y, Alani H. 2013.** Evaluation datasets for twitter sentiment analysis: a survey and a new dataset, the STS-gold. Available at <http://www.semanticwebjournal.org/content/semantic-sentiment-analysis-twitter>.
- Schouten K, Frasincar F. 2015.** Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering* **28(3)**:813–830 DOI [10.1109/TKDE.2015.2485209](https://doi.org/10.1109/TKDE.2015.2485209).
- Shahiri AM, Husain W, Rashid NA. 2015.** A review on predicting student's performance using data mining techniques. *Procedia Computer Science* **72(14)**:414–422 DOI [10.1016/j.procs.2015.12.157](https://doi.org/10.1016/j.procs.2015.12.157).
- Sharma S, Prasetyo DB. 2022.** Machine learning approaches to sentiment analysis in education. *Computers & Education* **180**:104412 DOI [10.1016/j.compedu.2022.104412](https://doi.org/10.1016/j.compedu.2022.104412).
- Singh AK, Kumar S, Bhushan S, Kumar P, Vashishtha A. 2021.** A proportional sentiment analysis of MOOCs course reviews using supervised learning algorithms. *Ingénierie des Systèmes d'Information* **26(5)**:501–506 DOI [10.18280/isi.260510](https://doi.org/10.18280/isi.260510).
- Vijayarani S, Janani R. 2016.** Text mining: open source tokenization tools-an analysis. *Advanced Computational Intelligence: An International Journal (ACII)* **3(1)**:37–47 DOI [10.1504/IJACII.2016.10000369](https://doi.org/10.1504/IJACII.2016.10000369).
- Xie X, Luo Y. 2021.** Sentiment analysis in education: trends and applications. *Journal of Educational Data Mining* **13(1)**:78–91 DOI [10.18706/jedm.2021.013](https://doi.org/10.18706/jedm.2021.013).
- Yu LC, Lee LH, Hao S, Wang J, He Y, Hu J, Zhang X, Lai R. 2016.** Building Chinese affective resources in valence-arousal dimensions. In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 540–545.
- Yu LC, Lee CW, Pan HI, Chou CY, Chao PY, Chen ZH, Tseng SF, Chan CL, Lai KR. 2017.** Improving early prediction of academic failure using sentiment analysis on self-evaluated comments. *Journal of Computer Assisted Learning* **34(4)**:358–365 DOI [10.1111/jcal.12247](https://doi.org/10.1111/jcal.12247).
- Zhang L, Liu B. 2017.** Sentiment analysis and opinion mining. *Encyclopedia of Machine Learning and Data Mining* **1**:1152–1161 DOI [10.1007/978-1-4899-7687-1](https://doi.org/10.1007/978-1-4899-7687-1).
- Zhao Y, Mamat M, Aysa A, Ubul K. 2023.** Multimodal sentiment system and method based on CRNN-SVM. *Neural Computing and Applications* **35**:10311–10328 DOI [10.1007/s00521-023-08236-2](https://doi.org/10.1007/s00521-023-08236-2).