Ministry of Science and Higher Education
of the Russian Federation
ITMO University

Faculty of Digital Technology

Educational program Big Data and Machine Learning

Subject area (major) 01.04.02 Applied Mathematics and Informatics

REPORT

on practical training *Research Internship*

Task topic : Multimodal analysis of webinars and online classes
Student : Pasala Krishna Chaitanya, J4134c
Head of Practice from the trainee's host organization : Oleg Olegovich Basov, Deputy
Dean, Faculty of Information and Communication Technologies.

Practice completed with grade: ⎯⎯⎯⎯⎯⎯
Commission member signatures:

⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯ *full name*
(signature)
⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯ *full name*
(signature)
⎯⎯⎯⎯⎯⎯⎯⎯⎯⎯ *full name*
(signature)

Date ⎯⎯⎯⎯⎯

St. Petersburg

2020

1

ABSTRACT

Student satisfaction is one of the important factors to be considered in the education system. Due to the COVID-19 pandemic and to make quality education available for a wide range of people, the current higher education system has undergone several changes. The classes and webinars being conducted through video conferencing systems such as zoom have gained unprecedented popularity. So, according to statistics, more than 97% of students switched to online education in the United States as it is a safe and feasible option to continue their education during the pandemic situation. At the same time, only 31% of students consider this format to be of the same quality as full-time, and only 5% of students find online learning to be of higher quality than offline. That is, 63% of students receiving online education are not satisfied with its current quality. According to the many types of research on student-teacher communication, non-verbal communication plays a crucial role in student satisfaction. Therefore, the study is conducted to find the direct correlation between time characteristics such as loudness, silence duration, change in intonations, tempo, and hand gestures of instructors during online classes and student satisfaction levels. As part of the study, recent online classes and webinars from MIT, Harvard, Stanford, IIT, ITMO, and few prestigious universities are analyzed. The quality of the classes is rated by the students as high or low based on their satisfaction level. Based on the analysis, successful and unsuccessful models are compiled. Based on the obtained models and the time parameters illustrated, assessment and recommendation systems are offered to improve the quality of both individual characteristics and the general concept of classes.

CONTENT

SYMBOLS AND ABBREVIATIONS

E-Learning — Electronic Learning.

ICT — Information and Communications Technology.

AV — Audiovisual.

TCP — Transmission Control Protocol.

UDP — User Datagram Protocol.

IP — Internet Protocol.

MMR — Multimedia Router.

TLS — Transport Layer Security.

HTTPS — Hypertext Transfer Protocol Secure.

CPU — central processing unit.

I/O — Input-Output.

INTRODUCTION

The COVID-19 pandemic forced most of the universities and schools worldwide to remain closed for students and working staff safety. So, to continue with the curriculum and to provide uninterrupted education, most of the universities have switched to synchronous E-learning mode. Development of the technology has provided different options for distance education are quickly adopted to promote additional values in the educational area [1]. Synchronous E-learning is a method of education where teacher and students are separated in space by utilizing software-based services to bridge the gap between them [2]. It is the all-inclusive blending of ICT gadgets and modern telecommunication equipment into the education system [3]. The E-learning platform is a software-based module that integrates diverse management devices, communication, evaluation, checking, etc. with an objective that could provide technological help to teachers as well as the targeted students to improve the variety of phases of the teaching-learning process, also the educational procedure in its nature being distant, classroom-based or having a combined nature where they merge the two modes in various extents [2].

Synchronous E-learning can happen with the help of video conferencing platforms. Videoconferencing-based educational systems can enable students in one or many locations, to watch and listen to an instructor in real-time, and maintain interactivity with those remote locations. It is a mature technology that can be used in the distance learning process to provide instructor-student or student-student interaction [4]. For most of the educational systems during the pandemic, It is the best way to deliver a better experience to the students. Most of the video conferencing applications are user-friendly and easy to operate while establishing the connection between users. Virtual classrooms can be created using video conferencing systems.

The advancement of digital technology and Internet services made people connect through video conferencing in the modern world. It has increased the flexibility of geographical locations and also increased cost and time efficiency. But still, users face few constraints due to the bandwidth limitations, high development, and maintenance costs, and lack of skills to use [5]. Among the popular video conferencing applications Zoom video communications is considered to be more helpful in the teaching-learning process. It solves most of the connectivity and user issues by its unique architecture and user-friendly interface. Even in the low bandwidth, zoom provides hassles communication to its users [6].

In the education industry, the way instructors communicate with the students plays a crucial role in student satisfaction and academic performance. According to the research conducted it is assumed that 65% of the meaning of conversational messages comes from non-verbal communication and only 35% through verbal communication [7].

Therefore, verbal and nonverbal communication is an aspect that needs to be emphasized for every teacher. The teacher's ability in applying both can help improve the quality of learning in the class. It has an impact on the academic and non-academic performance of students. Instead, if the teachers are not able to communicate effectively in transferring knowledge, skills, and attitudes, it will be hard to achieve the expected objectives. This shows that students who understand and can do the tasks very well also depend on how the teachers' style of verbal and nonverbal communication. Verbal communication includes a short talk, discussion, ideas, and so forth [8]. Nonverbal communication skills include body language, the use of eye contact, hand gestures, tone, change in intonations, and speed of speech which cues for creating emotional connections and understanding the meaning of the words [9, 10]. Then the two communication skills must always be honed by teachers, to create quality learning.

1   Video conferencing

Video conferencing is a technology for receiving and transmitting audio-video signals in real-time by multiple users regardless of their location over the internet [11]. The video conferencing connection process can be divided into two steps, compression and transfer.

— During compression, the devices capture analog AV input through the camera and microphone connected to the device of the sender. The data collected is in the form of continuous waves of frequencies and amplitudes. Due to the constraints in the size of data captured, sometimes, it cannot be transferred over a normal network. So, codecs must be used to compress the data into digital packets. This enables the captured AV input to travel faster over the internet [12].

— During the transfer phase, the digitally compressed data is sent over the digital network to all the receiving devices. Once it reaches the endpoint, the codecs decompress the data. The codecs convert it back the digital packets into analog audio and video. This enables the receiving screen and speakers to correctly view and hear the AV data of the sender [12].

1.1   Connection protocol

A network connection protocol is a set of rules that determine how data is transmitted between different devices in the same network. Essentially, it allows connected devices to communicate with each other, regardless of any differences in their internal processes, structure, or design. There are different types of protocols, commonly used connection protocols are

a) TCP is a connection-oriented protocol that computers use to communicate over the internet. It is a communications protocol, using which the data is transmitted between systems over the network. In this, the data is transmitted in the form of packets. It includes error-checking, guarantees the delivery, and preserves the order of the data packets. It establishes a connection between a sender and receiver before data can be sent [13].

b) UDP is a connectionless protocol of communication proposed at the transport layer, that works just like TCP, but here it is assumed that error-checking and recovery services are not required. Instead, UDP continuously sends datagrams to the recipient whether they receive them or not. It does not establish a connection before sending data. TCP uses a re-transmission technique to overcome erroneous packets, which introduces additional delays [14]. So, TCP is slower than UDP.

1.2   Transcoding

Due to users heterogeneous streaming devices, networks, and their preferences, input Audio-video data need to be transcoded to meet the characteristics of the user devices as initial data is captured in a particular format, spatial resolution, frame rate, and bit rate [15]. Transcoding is the process of converting an audio or video file from one encoding format to another to increase the number of compatible target devices a media file can be played on.

2   Zoom video communications

Zoom Video Communications, Inc. is a cloud-based video conferencing software platform used for live video, audio, and text communications across various devices and operating systems. Zoom has several products as zoom chats, webinars, conference rooms...etc. One of the keys differentiates that facilitates the ease and reliability of the cloud platform is Zoom's connection process. Zoom's connection process ensures that whenever someone attempts to access the platform there is an optimized path to Zoom's geographically distributed and highly available infrastructure [6].

2.1   Zoom components

Zoom Meeting Zone   is a logical association of servers that are typically physically co-located that can host a Zoom session. A Zoom Meeting Zone and its associated servers may be located within one of Zoom's global data centers or can be located within an organization's network if running Zoom's on-premise solution. The primary components of a Meeting Zone are Multimedia Routers and Zone Controllers [6].

Zoom Zone Controller   is responsible for the management and orchestration of all activity that occurs within a given Zoom Meeting Zone. Deployed in a highly available configuration, these systems track the load on all servers with the Zone and help broker requests for new connections into the zone [6].

Zoom Multimedia Router   is responsible for hosting Zoom meetings and webinars. These servers ensure that the rich offering of voice, video, and content are properly distributed between all participants in a given session [6]. The MMR architecture is aimed at providing quality of service to multimedia traffic in a local area environment while retaining a compact and simple design.

Zoom HTTP Tunnel   service is an integral part of Zoom's network resiliency strategy. It is in various public clouds and Zoom data centers, these servers offer a connection point to clients who are unable to connect to the Zoom platform through other network channels. Once a tunnel is established between the Zoom Client and Zoom HTTP Tunnel, the client can access the Zoom Meeting Zone across the various data centers [6].

2.2   Zoom connection process

The process of connecting to the Zoom session happens in different layers:

— After receiving the request to join a video conference, the zoom application contacts the zoom web infrastructure to obtain the required data to connect to the conference [6]. The zoom cloud connection starts at the base with the Intelligent Transport

Layer (responsible for end-to-end communication over a network), which decides if UDP, TCP, TLS, or HTTPS on the client layer is the best for connectivity based on different proxy settings and the need to go through firewalls. Reliable UDP is used for content screen sharing. Zoom will detect client Geo-location and assign closed resources. It may impose private links for cross international countries [16].

— The next layer is the Reactive Quality of Service Layer. This layer reacts to real-time network and device status. This layer not only monitors the client's bandwidth, packet loss, latency, and Jitter but also collects the client's CPU usage, memory, and network I/O. It notifies the up layer to take the best action within adaptive technology [16].

— Zoom uses Zoom's Adaptive Codec, which happens in the Session Layer. The multiple layers around this codec optimize the video frame rate and resolution and provide superior quality and reliability for various network environments and different devices. Zoom uses multiple streams, allowing the application to toggle between streams to ensure that the best quality video gets delivered to end-users. Because of Zoom's compression technology, the system can operate well in an environment with up to 45% packet loss. In these instances, Zoom will prioritize audio over video, because audio is more crucial in business discussions and collaboration. Zoom's multi-stream technology handles bandwidth adjustments for the end-user to improve their quality based upon their ability to receive data. The client can automatically downsize the received video if there's a bad network environment [16].

— The Distributed Conference Layer utilizes subscription information for switch technology with no transcoding or mixing. Because, as the streaming traffic increases, transcoding becomes a computationally expensive and time-consuming operation. Therefore, streaming service providers have to store numerous transcoded versions of a given video to serve various display devices [15]. Traditional conferencing services always transcode the streams and mix them. Zoom's switching method uses less CPU and memory, giving the system a massive scalability difference. One server can handle 2000 clients. Zoom uses globally-distributed technology based on user Geo-location and optimized network path. The meeting participants are always connected to a nearby data center and assigned to the least loaded server. Meeting participants will be also aggregated to the same server if they are in the same place [16].

— The Meeting Server is the MMR, and MMRs are grouped in a "Meeting Zone." Zone Controllers manage all of the MMRs and report their status to the Global Cloud Controller for each Meeting Zone. The Meeting Zones are duplicated for each data center with the same architecture and can easily add more zones on-the-fly for added capacity in each region. The three layers (the MMR, Zoom Controller, and Global Cloud Controller), are used to balance resources in different locations. If just two participants are in a meeting. Zoom will utilize peer-to-peer connections for excellent speed and reliability. All

of this allows Zoom to maintain meeting services availability of 99.99% up-time and deliver the most reliable video service [16]. Zoom uses a mix of cloud technologies, their own data centers to help deliver the service. And also use AWS, Oracle Cloud Infrastructure, and Azure as well as their global data center network of co-location.

— During some cases, If there is no response from any of the Zone Controllers after a few seconds, the Zoom client will attempt to connect using the HTTP Tunnel. To ensure multiple paths for a successful connection, these servers are housed in both public clouds and Zoom data centers. This connection is attempted over SSL. The Zoom client will ping multiple HTTP tunnels and the first to respond is used [6].

3 Non verbal communication

Nonverbal communication is the process of sharing meaning interpretive and multi-media communication type. It is natural, reliable, and spontaneous [8]. In simple it is a process of conveying messages without using written or spoken words. Nonverbal communications help to maintain learner's focus throughout the class via a continuum of instruction to achieve an equilibrium between communication and rapport [10]. Using nonverbal signals, the following information can be encoded such as physical attributes of the speaker, nonarbitrary referential meaning to objects and environmental states [17]. The sender of non-verbal communication may be unaware that he or she is sending a non-verbal message. Consequently, he or she may not be aware of the impact that the non-verbal message may have. Since the non-verbal message may be transmitted unconsciously, the sender will not have planned it. Therefore, a non-verbal message can be more reliable than an oral or written message [7]. The types of nonverbal communication associated with student satisfaction are eye contact, paralinguistics, gestures, physical appearance, and proxemics [7].

According to [7], the instructor physical appearances are not associated with student satisfaction levels. As the study deals with digital education eye contact and proxemics are not considered. So, the relation between gestures and paralinguistics is considered in the study to find their association with student satisfaction during online webinars and classes.

3.1 Gestures

Gestures are movements and signals using body parts used to communicate messages with or without the conjunction of words. Gestures include movements of hands, face, head nods, and other body parts. Gestures refer to the interlocutor rather than to the topic of conversation, and they help maintain the conversation as a social system [18]. Gestures are made predominantly when the person is talking, and they are temporally synchronized with the speaker's verbal syntax occurring at precisely the same moment as the relevant part of speech [19]. Gestures are not treated as physical movements but rather as referential acts; they convey meaning, depict events, and represent ideas. They specify and often clarify verbal references, and they can denote meanings that may not be in the accompanying words. Gestures work together to convey semantic information in a highly integrated manner. Gestures are particularly suited for solving the speaker's problem of involving the listener without disrupting the topical verbal narrative. Sometimes, interactive gestures are specialized to serve their function efficiently without requiring words [18].

Interactive gestures would be particularly affected by the loss of visual availability (i.e., fewer in the partition than in the face-to-face condition) because unwitnessed

gestures could not serve the function of including and involving the other person in the conversational system. Individuals talk and gesture to convey information to others, so the absence of another person physically would be expected to decrease the overall rate of gesturing. The rate of interactive gestures was higher for partners interacting face-to-face than for those who could not see each other [18].

3.2 Paralinguistics

Paralinguistics is part of vocal communication that refers to every element and nuance of the speech that does not involve the verbal message of the speech. Paralinguistics is defined as the discipline dealing with those phenomena that are modulated onto or embedded into the verbal message, be this in acoustics (vocal, non-verbal phenomena) or linguistics (connotations of single units or bunches of units) [20]. Besides the linguistic (verbal) information conveyed by speech, the paralinguistic (non-verbal) information, such as intentions, attitudes, and emotions expressed by the speaker, also convey important meanings in communication. Therefore, to realize smooth communication between humans and spoken dialogue systems (such as robots), it becomes important to consider both linguistic and para-linguistic information. Para-linguistic communication can be much more subtle than other forms of nonverbal communication.

[21] out that although what we say; the words, phrases, and sentences are important to communication, the way we use language can be even more important than our words as sources of information. Para-language, sometimes called vocalic refers to any cues created in the process of verbalizing other than the words themselves. This includes sniff and sneezes, pitch, rate of speech, nasality, pauses, and even silence. With spoken language, loudness, rate of speaking, tone, pitch variation, and uses of pauses can have a major influence on whether and how one can react to the words.

Research reveals that when a listener or receiver attempts to judge where the speaker or source is on a contempt-affection continuum, they rely more on the tone of the voice than on the content of the message. Paralinguistic features of the human voice are used to identify a person's age, gender, emotions, and health condition [22].

3.2.1 Change in intonations

The term intonation refers to a means for conveying information in a speech that is independent of the words. Central to intonation is variation in speaking pitch, and intonation is often thought of as the use of pitch over the domain of the utterance [23].

Variation of pitch within speakers, whether within or across utterances, is known as intonation, and largely reflects the syntactic and semantic structure of the utterance. Intonational contours are more abstract than absolute pitches and are usually described

as sequences of relative units of highs, and Lows. Languages vary in the intonational contours they use for particular types of utterances.

Intonation can also, for instance, help to regulate turn-taking in conversation, since there are intonational mechanisms speakers can use to indicate that they have had their say, or, conversely, that they are in full flow and do not want to be interrupted.

3.2.2 Tempo

The tempo is defined as the speed of speaking, that is speaking rate, which is best measured by the rate of syllable succession [24]. The speech tempo of a person may vary based on contextual and emotional factors, and the tempo is also dependent on the language is spoken and the dialects of the language. People may interpret speech tempo as an indicator of the importance of what is being said [25].

3.2.3 Loudness

Loudness is a psychological term used to describe the magnitude of an auditory sensation [26]. It is evident that loudness is dependent not alone upon the intensity of the sound but also upon their physical composition.

Loudness is the perceptual attribute of sound intensity and is that attribute of auditory sensation in terms of which sounds can be ordered on a scale from quiet to loud. The perception of loudness is a psychological attribute that depends on intensity, but also on several other parameters, like frequency selectivity, bandwidth, and duration of the sound [27].

4 Experiment and Results

The experiment is conducted to find the direct correlation between time characteristics such as gestures, and paralinguistic features of the instructors during online classes and student satisfaction levels. Based on the obtained results, assessment and recommendation systems are offered to improve the quality of both individual characteristics and the general concept of classes.

4.1 Sample data

As part of the experiment, recent online classes and webinars from MIT, Harvard, Stanford, Oxford, IIT's, ITMO, and few prestigious universities during the period 2019-2020 are collected. The study included 215 webinars. Out of the sample, 192 units are of English medium and 23 units are in Russian medium. The data is extracted from free sources such as YouTube and university websites available on the internet. From the obtained data, characteristics such as loudness, tempo, change in intonations, instructor visual presence, and duration of hand gestures performed are extracted and analyzed.

4.2 Student satisfaction level

The quality of the classes is rated by the viewers on a scale of 1 to 5 based on their satisfaction. Based on the rating system, '1' implies all the participants are dissatisfied, whereas '5' implies all the participants are satisfied with the class or webinar. As the data is collected from the public domain, most of the students, who rated the video may be interested in the topic or subject. From Fig. 4.1, the frequency of the rating is aligned right toward '5'. Therefore the distribution of rating is left-skewed. According to the collected sample data, the median is considered a central tendency due to its skewness. The median satisfaction level is 4.82.

4.3 Correlation and dependence

Correlation is used to find the association between two variables and estimate the strength of their relationship. Correlation provides information about not only the strength but also the direction of a relationship [28]. The most important fact is that correlation does not imply causation. The correlation between two variable is quantified with a number which varies between -1 to 1, negative values mean that variables are inversely related [29]. The study is conducted to find the association between individual nonverbal characteristics and student satisfaction levels, using the collected sample data.

In the experiment, the Pearson correlation coefficient is used to determine the strength of association between the time characters tics and the student's satisfaction level. Pearson correlation coefficient just efficiently characterizes the linear correlation between the attributes of the normal distribution [30]. A Pearson coefficient also ranges
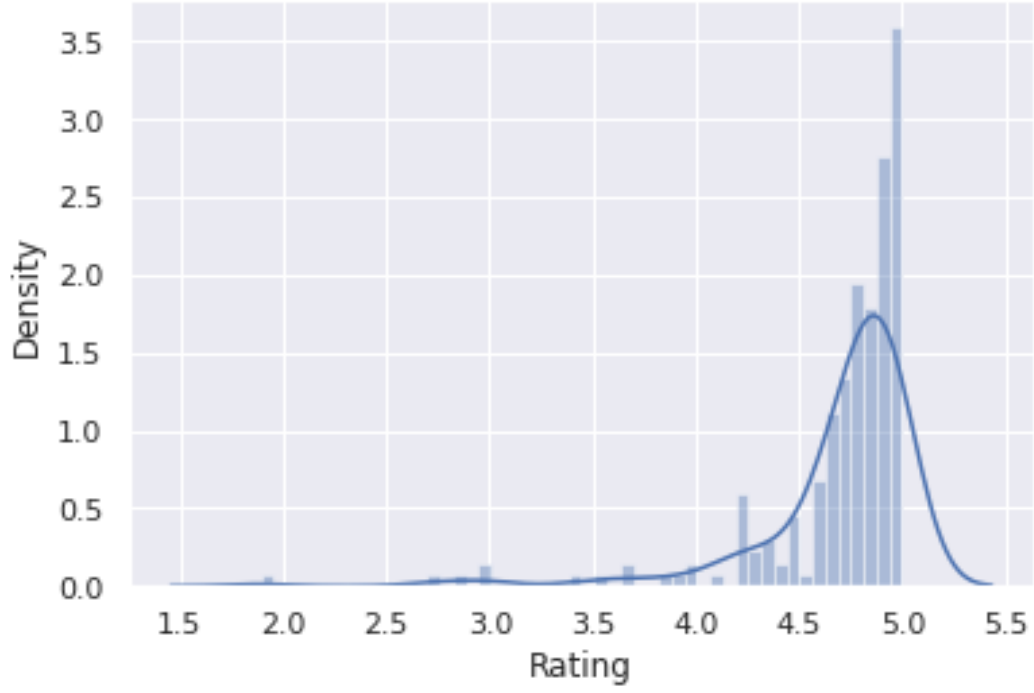
Fig. 4.1 — Frequency distribution of students rating

from –1 to +1. It can be interpreted as describing anything between no association ($\rho = 0$) to a perfect monotonic relationship ($\rho = \pm 1$) [28].

Any relationship or association between two variables should be assessed not just for the strength and direction, but also by whether the relationship is "significant", based on the p-value [31]. The significance level of $\alpha = 0.05$ is chosen for the experiment to check the reliability of correlation analysis.

Table 4.1 — Pearson correlation matrix between variables

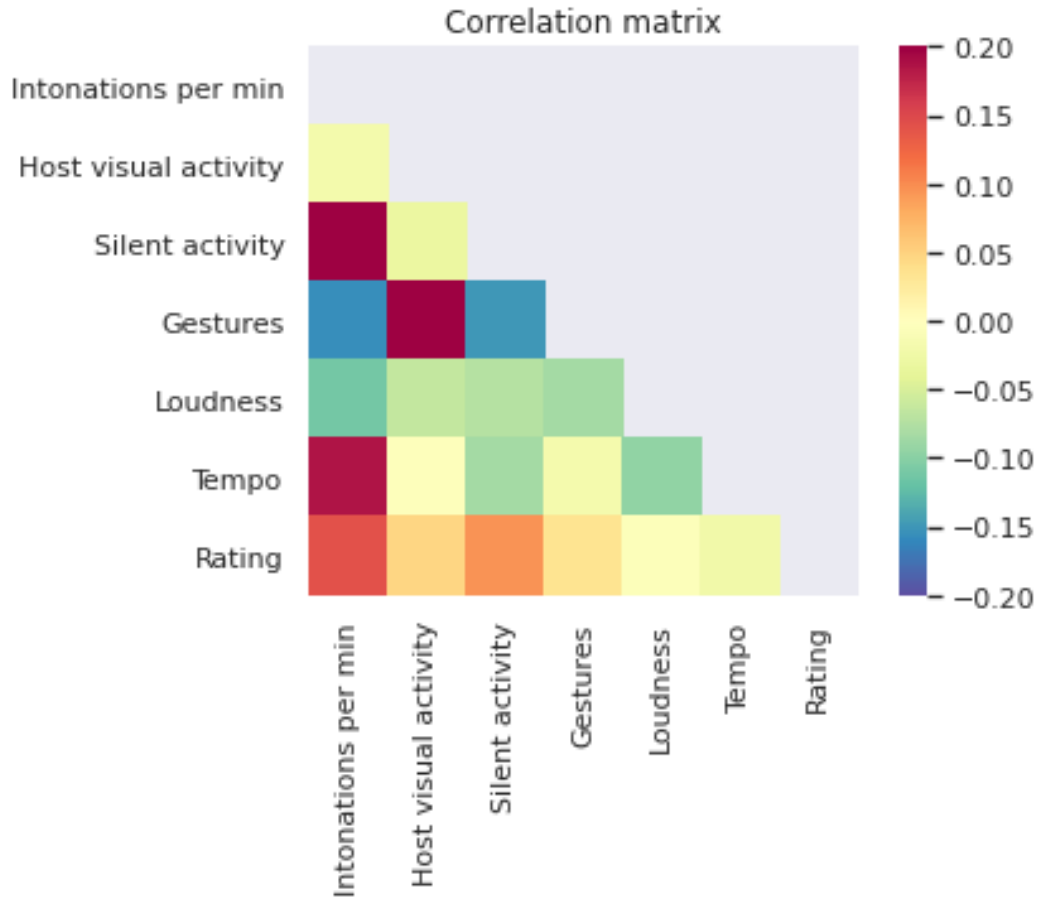| Index	Gestures	Change in intonations	Silence	Tempo	Loudness	Visual presence	Rating
Gestures	1	-0.156	-0.148	-0.01	-0.083	0.75	0.04
Change in intonations	-0.156	1	0.75	0.186	-0.112	-0.02	0.15
Silence	-0.148	0.75	1	-0.08	-0.07	-0.03	0.1
Tempo	-0.01	0.186	-0.08	1	-0.09	0	-0.02
Loudness	-0.083	-0.112	-0.07	-0.09	1	-0.06	0
Visual presence	0.75	-0.02	-0.03	0	-0.06	1	0.05
Rating	0.04	0.15	0.1	-0.02	0	0.05	1

Fig. 4.2 — Heatmap of correlations between the variables

4.3.1 Gestures

According to the sample data collected, the correlation between gestures performed during online lectures and student satisfaction level is 0.036, which is weak. And the corresponding p-value for the sample is 0.6, which is greater than the significance level 0.05. Therefore the relation is not significant enough. So, it is assumed that there is no significant relation between gestures performed during online classes and student satisfaction.

From the sample data it can be observed that the host is visually present in the online class or webinars about 30 percent of the duration only, as, during most of the lecture, the instructor used digital presentation in form of slides to interact with the audience. From the table 4.1, the correlation between host visual presence and gestures performed is 0.75, which is strong. So, due to the loss of visual availability of the host, gestures are not showing much effect on students' satisfaction level.

4.3.2 Paralinguistic

Among the paralinguistic features, From the table 4.2, it is clear that silence, tempo, and loudness are not significantly related to satisfaction level during online classes. But there is a significant relationship between change in intonations and students' satisfaction. As the Pearson coefficient between the variables is 0.14, and the p-value is 0.036, which is less than the significance level 0.05. So, it is assumed that there is a significant positive relationship between change in intonations and satisfaction.

Table 4.2 — The table represents the reliability and pearson correlation of paralinguistic features with student satisfaction

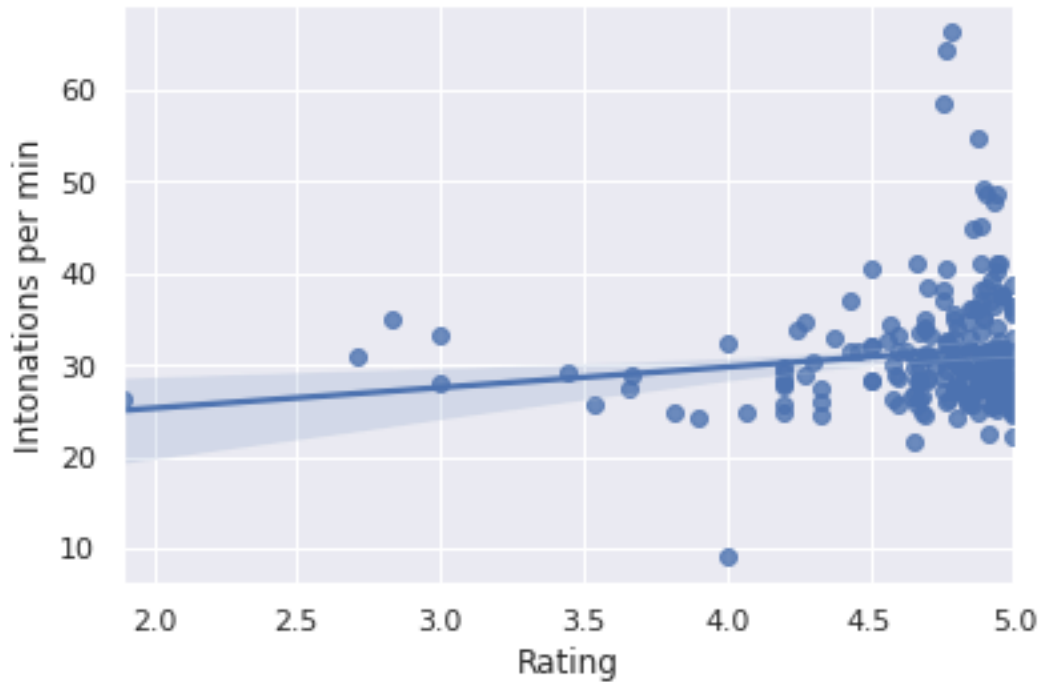| Paralinguistic feature	Correlation	p-value	Significance
Change in intonations	0.15	0.036	Significant
Silence	0.1	0.15	Not significant
Tempo	-0.02	0.75	Not significant
Loudness	0	0.94	Not significant

Fig. 4.3 — Scatter plot and regression line between change in intonations and students rating

4.4   Limitations

One of the major limitations of the study conducted is that the majority of the data collected is from free sources like YouTube. So, the data may differ from the real-time online lectures conducted by the universities. And most of the people who participated in the lecture or webinar are interested in the topic or course. So, the ratings given by the participants may be biased. Due to the small sample size of 215 units, the results may not reflect the whole population. There are few other factors such as internet connectivity, geographical locations, students' mood, instructors' behavior towards students which also affect the student's satisfaction level. But these all factors are neglected in this study case. In further research, these limitations will be avoided by collecting proper data from universities and by continuous student's feedback.

CONCLUSION

Based on the experiment conducted, it is assumed that there is a significant positive association between instructors' change in intonations in speech and students' satisfaction during online classes. This implies that any effort to enhance students' satisfaction should be preceded by improvement in para-linguistic features. This result is also supported by previous researches conducted based on nonverbal communications.

The study offers sufficient depth of insights into the specific areas that require attention and improvements during online classes. Despite many limitations and disadvantages, online education is growing everyday. So, to provide proper education and satisfy students, educational departments should try to overcome these kinds of situations, the instructor should try to show more variations in pitch based on the emotion carried by the sentence and its meaning. Mostly, due to the unavailability of visual presence of the instructor during online classes or webinars, para-linguistic features should be prioritized.

REFERENCES

1. *Malinovski, Toni.* Impact of Different Quality of Service Mechanisms on Students' Quality of Experience in Videoconferencing Learning Environment / Toni Malinovski, Vladimir Trajkovik, Tatjana Vasileva Stojanovska // *Turkish Online Journal of Distance Education.* — 2018. — 07. — Vol. 19. — Pp. 24–37.

2. *Singh, Satendar.* Mechanising E-learning for equiping start-up entrepreneurs / Satendar Singh, Syed Zakir Hussain // *Materials Today: Proceedings.* — 2020. `http://www.sciencedirect.com/science/article/pii/S2214785320361617`.

3. *Eze, Sunday.* The utilisation of e-learning facilities in the educational delivery system of Nigeria: a study of M-University / Sunday Eze, Vera Chinedu-Eze, Adenike Bello. — 2018. — 09. — Vol. 15. — Pp. 1–20.

4. Cloud-based videoconferencing learning environments in Higher Education: The growing pains of synchronous learning in an asynchronous world. / Deborah Clawson, Michael Korns, Scott Decker, David Piper // Proceedings of Society for Information Technology Teacher Education International Conference 2016 / Ed. by Gregory Chamblee, Lee Langub. — Savannah, GA, United States: Association for the Advancement of Computing in Education (AACE), 2016. — March. — Pp. 178–184. `https://www.learntechlib.org/p/171670`.

5. *Kimiloglu, Hande.* Perceptions About and Attitude toward the Usage of E-Learning in Corporate Training / Hande Kimiloglu, Meltem Özturan, Birgul Kutlu // *Computers in Human Behavior.* — 2017. — 02. — Vol. 72.

6. *Inc., Zoom Video Communications.* Zoom Connection Process / Zoom Video Communications Inc. — 2020. `https://zoom.us/docs/doc/ZoomConnectionProcessWhitepaper.pdf`.

7. *Binti Sarah Ahmad, Sabir Binti Farah Merican Merican Isahak Nazrenee Intan Zani Rosliza Administration Business Teknologi Universiti Kedah Mara Surat Peti Darulaman Kedah.* Non-verbal communication and its effect on student satisfaction / Sabir Binti Farah Merican Merican Isahak Nazrenee Intan Zani Rosliza Administration Business Teknologi Universiti Kedah Mara Surat Peti Darulaman Kedah Binti, Sarah Ahmad. — 2012. — Vol. 51. — Pp. 10644–10649.

8. *Wahyuni, Akhtim.* The power of verbal and nonverbal communication in learning / Akhtim Wahyuni // 1st International Conference on Intellectuals' Global Responsibility (ICIGR 2017) / Atlantis Press. — 2018.

9. Nonverbal Communication // *Dental Abstracts.* — 2020. — Vol. 65, no. 6. — P. 376. `http://www.sciencedirect.com/science/article/pii/S0011848620305513`.

10. *Scott-Parker, Bridie.* Nonverbal communication during the learner lesson with a professional driving instructor: a novel investigation / Bridie Scott-Parker // *Transportation research part F: traffic psychology and behaviour.* — 2017. — Vol. 47. — Pp. 1–12.

11. *of Engineering, McGraw-Hill Concise Encyclopedia.* Videotelephony | Article about Videotelephony by The Free Dictionary. — 2002. `https://encyclopedia2.thefreedictionary.com/Videotelephony`.

12. *Rouse, Margaret.* What is Video Conferencing? — 2020. `https://searchunifiedcommunications.techtarget.com/definition/video-conference`.

13. *Pandya, Pramod.* Transmission Control Protocol/Internet Protocol Packet Analysis / Pramod Pandya // Computer and Information Security Handbook. — Elsevier, 2013. — Pp. e205–e218.

14. *Zaidi, S.* Enhanced user datagram protocol for video streaming in VANET / S. Zaidi, S. Bitam, A. Mellouk // 2017 IEEE International Conference on Communications (ICC). — 2017. — Pp. 1–6.

15. Performance Analysis and Modeling of Video Transcoding Using Heterogeneous Cloud Services / X. Li, M. A. Salehi, Y. Joshi et al. // *IEEE Transactions on Parallel and Distributed Systems.* — 2019. — Vol. 30, no. 4. — Pp. 910–922.

16. *Inc., Zoom Video Communications.* Innovation Creates the Difference in User Experience / Zoom Video Communications Inc. — 2019. `https://zoomgov.com/docs/doc/Zoom_Global_Infrastructure.pdf`.

17. *Frühholz, Sascha.* Nonverbal Auditory Communication–Evidence for Integrated Neural Systems for Voice Signal Production and Perception / Sascha Frühholz, Stefan R Schweinberger // *Progress in Neurobiology.* — 2020. — P. 101948.

18. Interactive Gestures / Janet Bavelas, Nicole Chovil, Douglas Lawrie, Allan Wade // *Discourse Processes - DISCOURSE PROCESS.* — 1992. — 10. — Vol. 15. — Pp. 469–489.

19. *Kita, Sotaro.* How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. / Sotaro Kita, Martha W Alibali, Mingyuan Chu // *Psychological review.* — 2017. — Vol. 124, no. 3. — P. 245.

20. Paralinguistics in speech and language—State-of-the-art and the challenge / Björn Schuller, Stefan Steidl, Anton Batliner et al. // *Computer Speech & Language.* — 2013. — Vol. 27, no. 1. — Pp. 4–39.

21. *Brent D Ruben, Lea Stewart.* Communication Human Behavior | Higher Education / Lea Stewart Brent D Ruben. — 2020. `https://he.kendallhunt.com/product/communication-human-behavior`.

22. Using acoustic paralinguistic information to assess the interaction quality in speech-based systems for elderly users / Humberto Pérez-Espinosa,

Juan Martínez-Miranda, Ismael Espinosa-Curiel et al. // *International Journal of Human-Computer Studies.* — 2017. — Vol. 98. — Pp. 1–13.

23. *Aarts, Bas.* The handbook of English linguistics / Bas Aarts, April MS McMahon, Lars Hinrichs. — Wiley Online Library, 2006.

24. *Wood, S.* Speech tempo (Introduction), year = 1973, url = https://medium.com/av-transcode/what-is-webrtc-and-how-to-setup-stun-turn-server-for-webrtc journal=Dept. of Linguistics, Lund University. .

25. *Quené, Hugo.* On the just noticeable difference for tempo in speech / Hugo Quené // *Journal of Phonetics.* — 2007. — Vol. 35, no. 3. — Pp. 353–362.

26. *Fletcher, Harvey.* Loudness, its definition, measurement and calculation / Harvey Fletcher, Wilden A Munson // *Bell System Technical Journal.* — 1933. — Vol. 12, no. 4. — Pp. 377–430.

27. *Schenkman, Bo N.* Detection, thresholds of human echolocation in static situations for distance, pitch, loudness and sharpness / Bo N Schenkman, Vijay Kiran Gidla // *Applied Acoustics.* — 2020. — Vol. 163. — P. 107214.

28. *Schober, Patrick.* Correlation coefficients: appropriate use and interpretation / Patrick Schober, Christa Boer, Lothar A Schwarte // *Anesthesia & Analgesia.* — 2018. — Vol. 126, no. 5. — Pp. 1763–1768.

29. *Akoglu, Haldun.* User's guide to correlation coefficients / Haldun Akoglu // *Turkish journal of emergency medicine.* — 2018. — Vol. 18, no. 3. — Pp. 91–93.

30. *Xu, Honghui.* Dependent evidence combination based on shearman coefficient and pearson coefficient / Honghui Xu, Yong Deng // *IEEE Access.* — 2017. — Vol. 6. — Pp. 11634–11640.

31. *Gogtay, NJ.* Principles of correlation analysis / NJ Gogtay, UM Thatte // *Journal of the Association of Physicians of India.* — 2017. — Vol. 65, no. 3. — Pp. 78–81.