Reviewing the Brain wave and Visualizing Massive Open Online Courses MOOCs Data on the brain of Students

Beeram Nithin Reddy[1],Jahnavi Dasari[2], Boinapally Srikar Rao[3], Rajkumar Rajasekaran[4]
[1]School of Information Technology and Engineering, Vellore Institute of Technology
[2]Geethanjali College of Engineering and Technology, Hyderabad
[3]Lovely Professional University
[4]School of Computer Science and Engineering, Vellore Institute of Technology

Abstract

MOOCs, the upcoming platform for online education, are free of charge courses uploaded by teachers from varying backgrounds of expertise for the benefit of students, no matter where they are or what topic they wish to study. However, inspite of the increasing growth curve, concerns have risen pertaining to the level of attention students pay to the videos or the comfort level they have while studying as compared to the education received in the presence of a teacher, one on one. This paper aims at reviewing the brain wave (Attention, Mediation, Raw waves, Delta waves, Alpha waves, Beta waves, Gamma waves) data collected for 10 college students and then using basic visualization techniques like histograms, line graphs etc. to draw conclusions regarding attention span, calmness and overall effectiveness of MOOCs in the field of education.

Keywords: MOOC, EEG, Delta waves, Alpha waves, Beta waves, Gamma waves, Attention, Mediation, histogram, line graphs.

Introduction

In this day and age of advancing educational platforms, Massive Open Online Courses, abbreviated as MOOCs are the next big thing and are being vastly researched upon. This development in the field of online and distance learning is one of the most radical recent ones. MOOCs are essentially not-for-credit online courses that are freely available and allow anyone who is interested to enroll. There is no limit to the number of people who are allowed to enroll and a vast multitude of topics are covered, ranging from Artificial Intelligence and Data Visualization to Languages and Home Sciences.

These videos are now being included in the official teachings in MIT and Harvard

MOOCs accounts for a lot of factors that differentiate them from the regular teaching practices like the diversity, informality in the way teaching takes place but most importantly the vast amount of data that is generated by students as online footprint with respect to these courses that can then be used for analysis of how students respond to online environments, what impact a change in the instructional environment brings etc. MOOCs also prove to be more beneficial for students in thesense that the same platform/"class" has students with varying degrees of knowledge, expertise and questions. This results in informative discussion forums and exchange of information that otherwise may not have been explicitly mentioned in the MOOC.

Debates arise, due to the inherent belief that the shaping a student can receive in the presence of an educated human cannot be paralleled by nine education alternatives, where the students essentially work by themselves, in the absence of traditional supervision.

As far as ratings go, the number of students enrolling remain on an all-time high. However, the course completion rates are still low. That is mostly because a lot of students enroll in the course to view a specific topic instead of the whole course. However, as researched, this is not counted as a dropout, or failure on the part of the MOOC course, because the aim of the course is not to have students complete it, but to provide assistance in whichever topic the student needs.

Concerns that have been sparked regarding the attention span students have and how much material they actually absorb from online tutorials are being researched upon. This paper looks into the same.

A Study on "MOOC scoring algorithm based on Chinese University MOOC learning behavior data" proposed an algorithm based on big data of learning behavior. Information entropy theory is used to take conversion ratio into account. The algorithm aims at helping the learners understand the learning behavior and state and hence, keep their interests engaged. The data used is from 4 courses, all from Chinese National Great Open Online Courses (2017). The algorithm studies all activities carried out by the students like viewing videos, doing exercises, participating in discussion forums, completing quizzes and exams etc. It was concluded that the passing rate of the courses under consideration was low, indicating that most people view MOOCs for specific topics, not for the certification. Hence, MOOCs are being used as more of an aid than an original source of knowledge. It was also concluded that the amount of time spent watching a video is directly related to the knowledge gained. The comprehensive scored were viewed using a radar graph. The formula for calculating the process score was derived in the paper itself. It was concluded that over the years the number of people passing the courses has increased and that showing this data to the users helps them understand their learning curve and helps maintain interest [1].Another study aims at arriving at conclusions regarding the benefits of integrating the use of MOOCs in regular classroom based teaching while also pointing out the possible hindrances that can be faced on the same. A balanced student workload was used and qualitative approach was adopted to analyse the learning diaries maintained by students. The conclusions were predominantly positive. Sourcing several MOOCs and allowing the students to choose which one they wish to study promotes positive motivation. The workload however should be carefully examined so that students face reasonable expectations. The outcomes of the course should be kept at the forefront to ensure a productive amalgamation of online and offline learning. Overall, the integration of MOOCs with classroom based studies is a beneficial step for educating students [2].

A research on "Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories" proposed using clustering as a means to study how engaged the learners are, with their MOOC courses. The dataset used is a MOOC offered by Graz University of Technology. An attempt is made to classify students on the basis of the levels of engagement in the course. The conclusion to the study was that the addition of intrinsic factors is needed to improve the future of MOOCs, because extrinsic factors alone are not enough to push students to be more committed to the MOOCs they enroll for. Focus should be on increasing motivation and making MOOCs more interactive and engaging [3].

A study on "Deciphering the attributes of student retention in massive open online courses using data mining techniques" aimed at using classification algorithm (decision tree) to declare a well-defined rationale for important attributes. The dataset involves three MOOC courses, and the application of data mining techniques on online learners, in order to analyse the in-course behavior exhibited by the participants. The paper also aims at using data mining results to point out the attributes that can help reduce attrition rate and conduct an analysis on the dropout rates (often significantly high in MOOCs) by analyzing different cohort behavior. The conclusion of the study

was that factors like the course that the learner is viewing, the birth year of the viewer or the age of the viewer are not attributes considered to have a significant say not the dropout rate associated with the course. They do not play a helpful role in determining the learning approach taxonomy or in analyzing the cohort behavior and its impacts on the dropout rate. On the other hand, factors like Days Active, Number of Chapters, Total number of Events etc. seem to be more influential when it comes to the analysis in question. Another conclusion was that, in order to increase the number of people who stay back for the entire MOOC course, the course providers can observe other courses that have a higher rateof retention. By studying the fuzziness amongst the enrollees, the course providers can effectively increase their own retention rates.
A report on “Massive Open Online Courses (MOOCs): Data on higher education” reported the use and benefits of Massive Open Online Courses (MOOC) in the field of higher education ranging from 2012 to 2017. Further, assessing and categorization of the MOOCs was done on the basis of multiple factors like the publication journal, researchers, year of publication, release date, theoretical content, and methodology used etc. There were five dynamics that were assessed in relation to improving the use of MOOCs which include intention to use, engagement, motivations, interaction and satisfaction [5].

Methodology

1) Collection of authentic Brain Wave Data

The dataset used for visualization has been taken from Kaggle.com[16]. The title of the dataset is “Confused student EEG brainwave data”, uploaded by Haohan Wang. It is a collection of EEG signal data from 10 college students, recorded while they were asked to watch MOOC video clips. The video clips were extracted form educational videos available online that were assumed to be understandable and not confusing for the college students. These include videos like ones based on basic algebra and geometry. Videos that count as confusing like ones based on Quantum Physics were also taken. 20 videos in total, 10 in each category were taken. The students were made to wear a single channel wireless MindSet that captured the activity of their frontal lobes while the students watched the videos. The students were also asked to manually rate whether they felt confused after watching the video on a scale of 1 to 7. All this data was recorded.

Subject ID	VideoI D	Attenti on	Mediati on	Raw	Delta	Theta	Alpha1	Beta1	Gamma 1	predefi nedlabe l	user-definedl abeln
0.00E+00	0.00E+00	5.60E+01	4.30E+01	2.78E+02	3.02E+05	9.06E+04	3.37E+04	2.79E+04	3.32E+04	0.00E+00	0.00E+00
0.00E+00	0.00E+00	4.00E+01	3.50E+01	-5.00E+01	7.38E+04	2.81E+04	1.44E+03	2.75E+03	5.29E+03	0.00E+00	0.00E+00
0.00E+00	0.00E+00	4.70E+01	4.80E+01	1.01E+02	7.58E+05	3.84E+05	2.02E+05	3.63E+04	5.72E+04	0.00E+00	0.00E+00
0.00E+00	0.00E+00	4.70E+01	5.70E+01	-5.00E+00	2.01E+06	1.29E+05	6.12E+04	1.15E+04	5.00E+04	0.00E+00	0.00E+00
0.00E+00	0.00E+00	4.40E+01	5.30E+01	-8.00E+00	1.01E+06	3.54E+05	3.71E+04	4.53E+04	4.48E+04	0.00E+00	0.00E+00
0.00E+00	0.00E+00	4.40E+01	6.60E+01	7.30E+01	1.79E+06	1.77E+05	5.94E+04	1.51E+04	3.38E+04	0.00E+00	0.00E+00
0.00E+00	0.00E+00	4.30E+01	6.90E+01	1.30E+02	6.35E+05	1.22E+05	9.01E+04	3.62E+04	6.29E+04	0.00E+00	0.00E+00
0.00E+00	0.00E+00	4.00E+01	6.10E+01	-2.00E+00	1.61E+05	1.21E+04	1.96E+03	1.28E+03	3.27E+03	0.00E+00	0.00E+00

Table 1 : Sample of the dataset (Alpha2, Beta2, Gamma2 are not shown) (original dataset)

Su bje ctI D	Video ID	Attenti on	Mediat ion	Raw	Delta	Theta	Alpha1	Beta1	Gamma1	pred efine dlab el	us er- de fin edl ab eln
1	0	56	43	278	301963	90612	33735	27946	33228	0	0
2	0	0	40	35	-50	73787	28083	2240	3687	0	0
3	0	0	47	48	101	758353	383745	62107	130536	0	0
4	0	0	47	57	-5	2012240	129350	17084	62462	0	0
5	0	0	44	53	-8	1005145	354328	88881	99603	0	0
6	0	0	44	66	73	1786446	176766	26157	33669	0	0
7	0	0	43	69	130	635191	122446	65072	53019	0	0
8	0	0	40	61	-2	161098	12119	809	3186	0	0
9	0	0	43	69	17	492796	120998	68242	88403	0	0

Table 2 : Converted Value Dataset

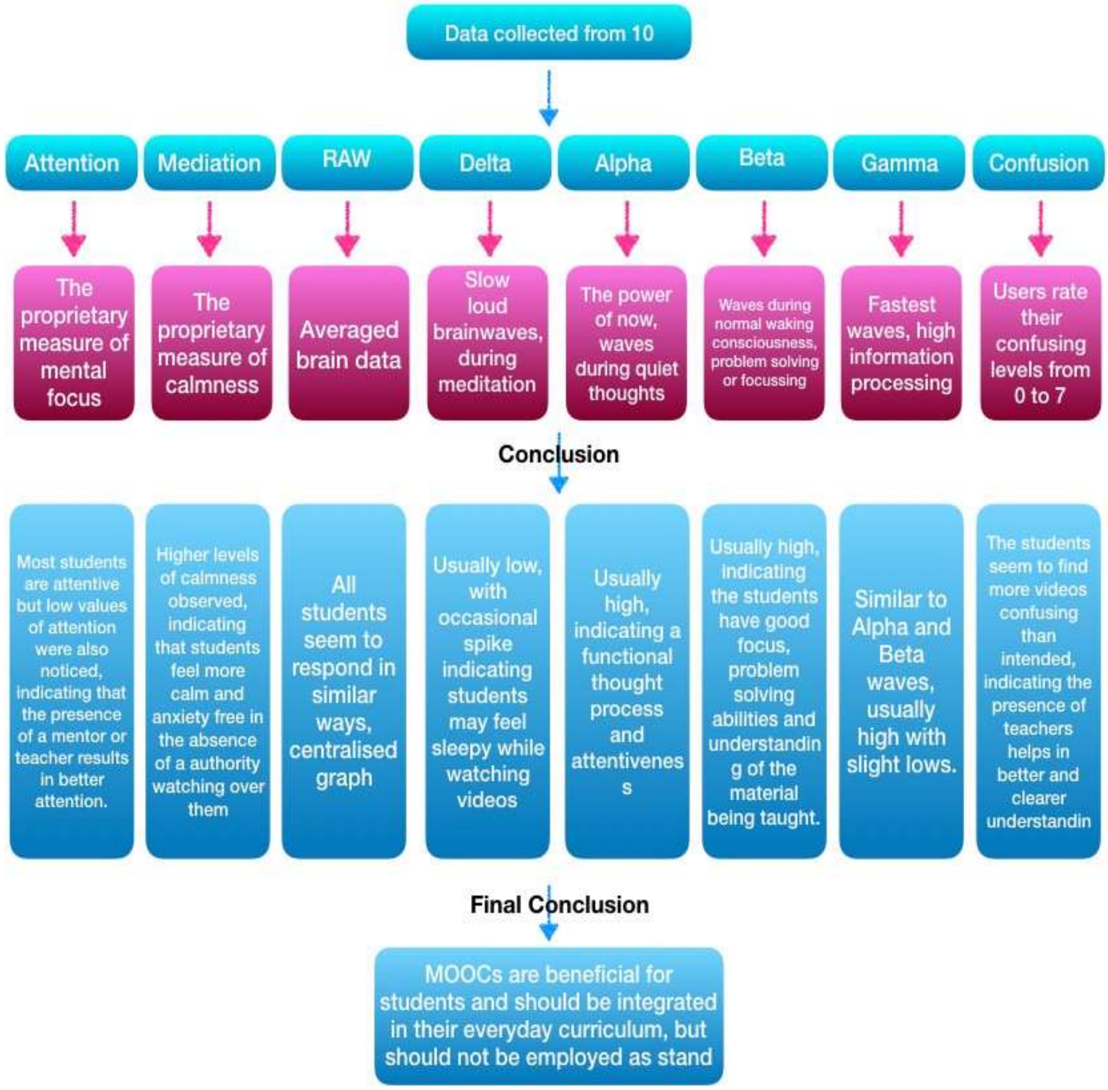
2) Workflow

The workflow pertaining to this paper started with the search of an appropriate topic that was relevant to the students as well as had a substantial role to play in the future. Since MOOCs are being widely researched upon currently and play a very important role in the education of students, now and possibly in the fore-seeable future, the topic was chosen as study of impact of MOOCs on students.

The next task was to collect research papers on topics similar to the one chosen, for research purposes. A thorough study of 15 such papers was done and the literature review was formulated. Following this, the collection of dataset was conducted. The dataset chosen in an open source dataset available on kaggle.com.

Next, the visualization process was conducted. Multiple different plots for the various fields in the dataset including Attention, Mediation, RAW brain waves etc were plotted (ranging from cluster dendrogram to histogram). Based on the visualization and the inference drawn, histograms and line graphs were decided to be the best visuals. After plotting, inferences were drawn and noted down for every graph. For the individual line graphs, the dataset was broken down into sub datasets, each then plotted separately. Conclusions were drawn based on the graphs and the research paper was concluded after all the references were correctly mentioned and all the citations were rightfully marked.

3) BLOCK DIAGRAM



Results and discussions

Graphs were plotted on the basis of the data obtained from the dataset and inferences were drawn accordingly. Firstly, a histogram was plotted on the basis of the Attention data. As understood from the histogram plotted for attention, which refers to the proprietary measure of mental focus, most of the students had a considerably good attention span, referring to the medium height bars near the middle of the graph. The values of very low and very high concentration correspond to lesser number of occurrences which is acceptable as human nature.

However some values of low attention seem to have decent number of occurrences.

Next, a histogram for Mediation was plotted.

The histogram for Mediation, which is proprietary measure of calmness shows a very symmetrical distribution. This indicates that majority of students have appropriate, normal levels of calmness while viewing the video clips. Cases of anxiety or unnatural calmness tending to sleepiness are very few.

Then, histograms of the brain waves were plotted: Raw signals, Delta waves, Theta waves, Alpha waves, Beta waves and Gamma waves in order.

The histogram for the brainwaves also shows the distribution of values, though they're not as varied as attention and mediation, meaning that at the minute electricity level, all the students seem to respond in similar ways.As an addition, the beta brainwaves for the student with student ID 0 were plotted using line graph and points. The graphs for all waves for all students were not plotted because that would amount to lot of graphs which is redundant from inference point of view.

The graphs show the variation in the brainwaves of the student as reported that 10 clips were selected for this experiment. The same can be plotted for the other students as well. Beta brainwaves were chosen to be plotted because these are the brain waves that are studied when the human's attention is directed towards tasks that are cognitive in nature. These waves represent the "fast" activity section of the brain when the subject is alert, engaged in something of a problem solving nature or decision making or is being attentive in a general sense.For the same student, alpha and gamma waves were also plotted, to understand the variation for an individual student.

Lastly, the labels of confusion and non-confusion were plotted, first as they were predefined and secondly, as they were labelled by the students. 0 for not confusing and 1 for confusing were used as normalizing values.

ISSN No : 1006-7930

Graphs

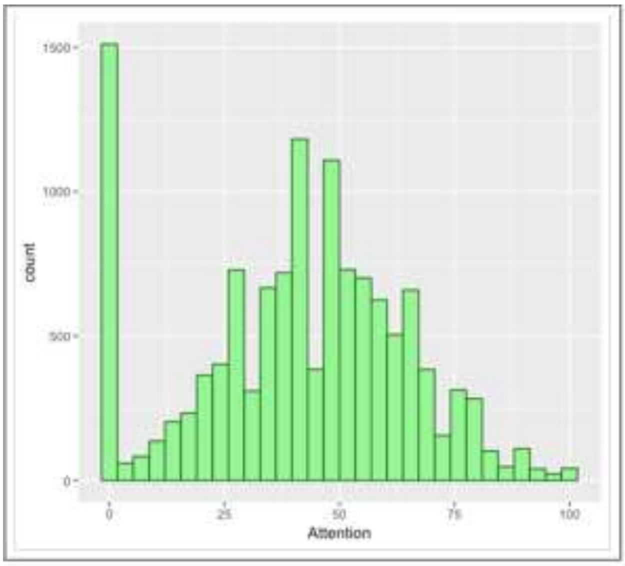


Fig 1. Histogram of Attention vs Count(Frequency)

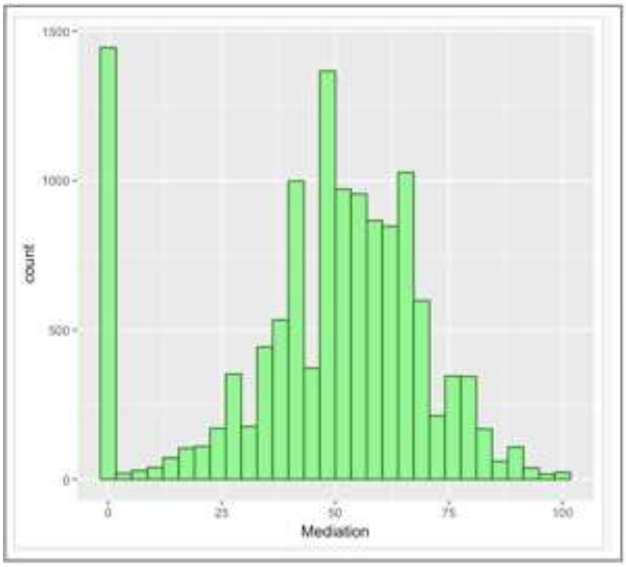
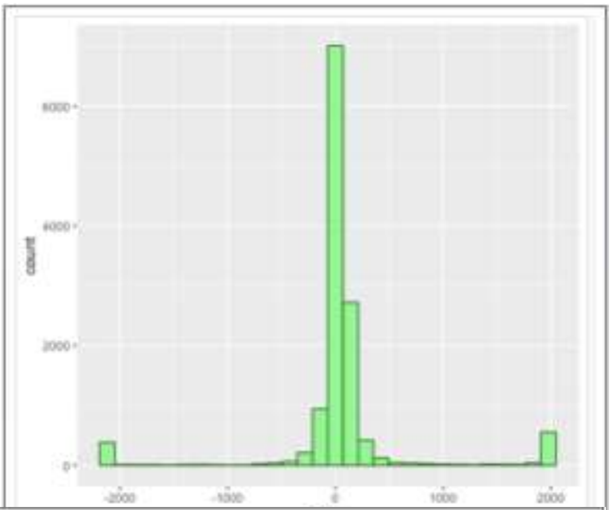
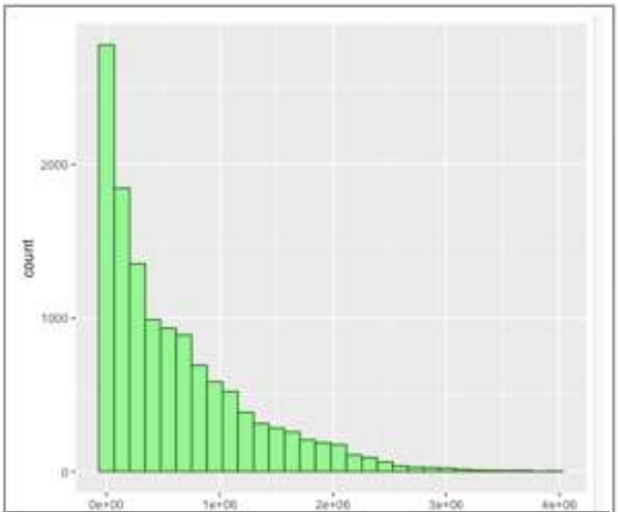


Fig 2. Histogram of Mediation vs Count(Frequency)





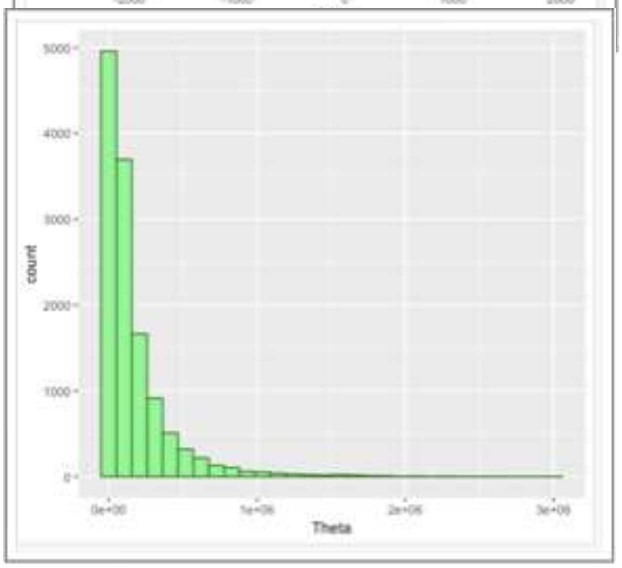


Fig 5. Histogram of Theta waves

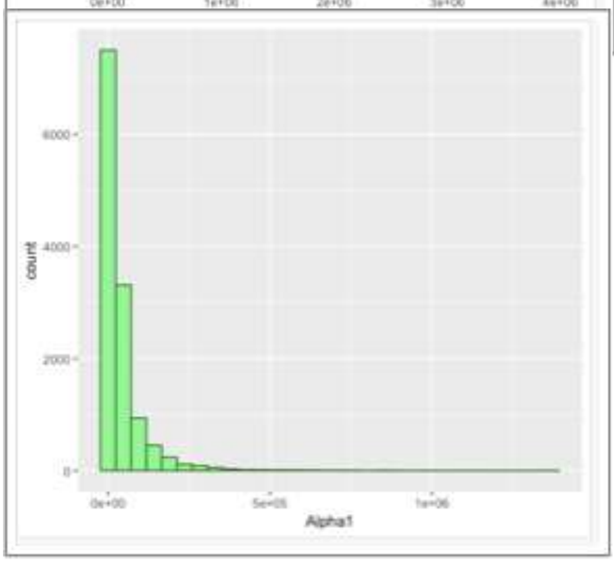


Fig 6. Histogram of Alpha waves

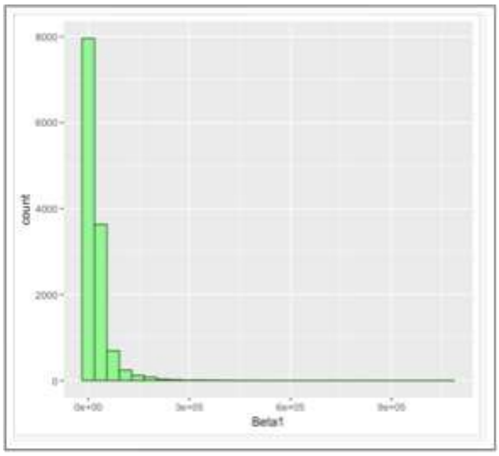


Fig 7. Histogram of Beta waves

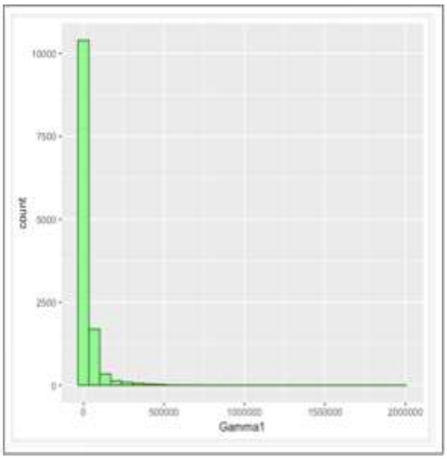


Fig 8. Histogram of Gamma waves

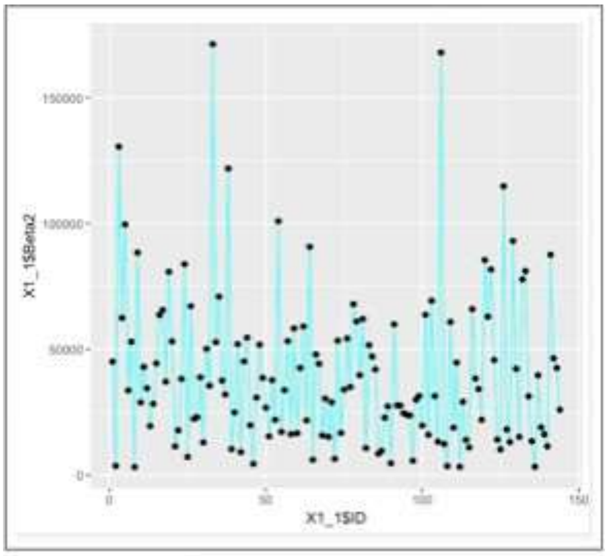


Fig 9. Student ID 0 Clip No 1

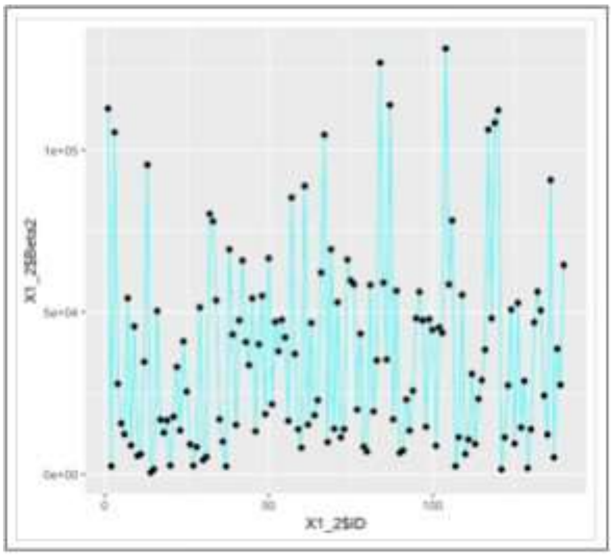


Fig 10. Student ID 0 Clip No 2

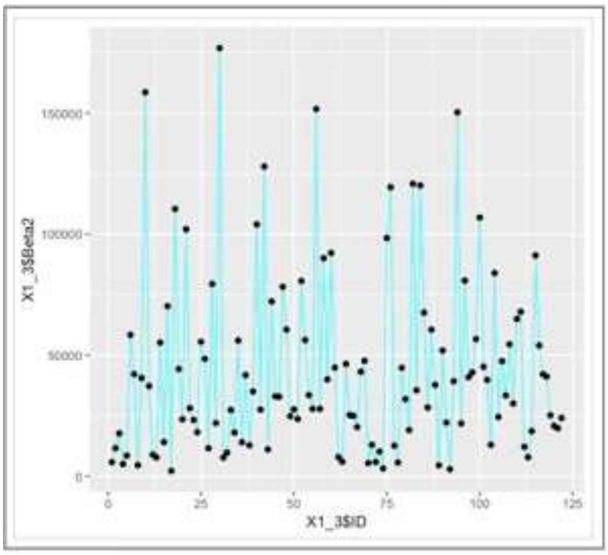


Fig 11. Student ID 0 Clip No 3

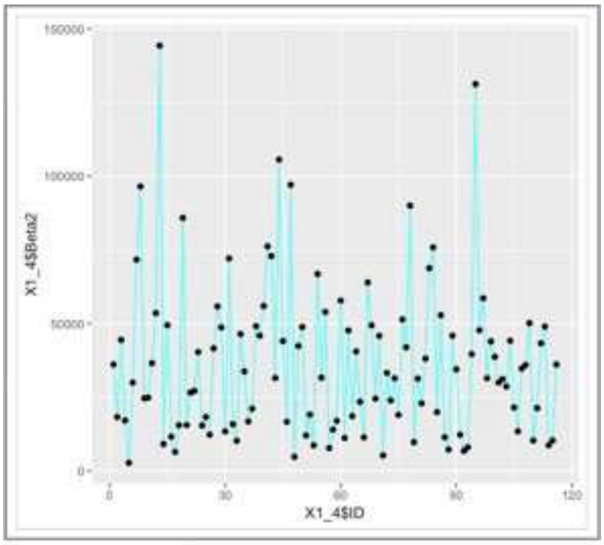


Fig 12. Student ID 0 Clip No 4

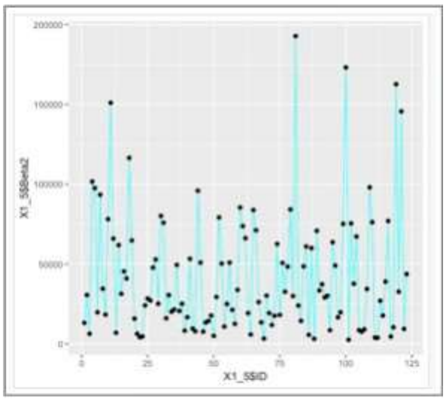


Fig 13. Student ID 0 Clip No 5

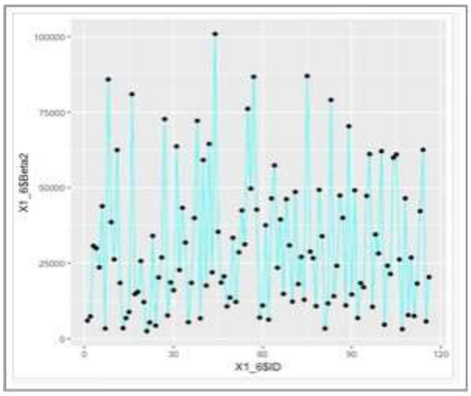


Fig 14. Student ID 0 Clip No 6

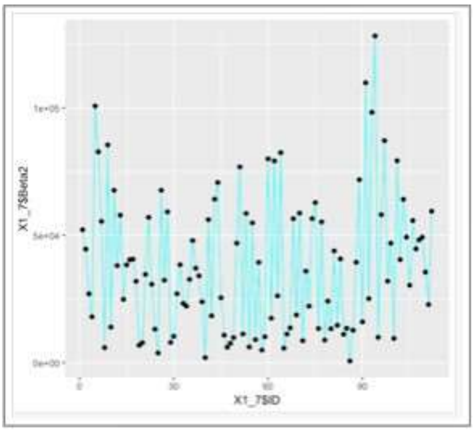


Fig 15. Student ID 0 Clip No 7

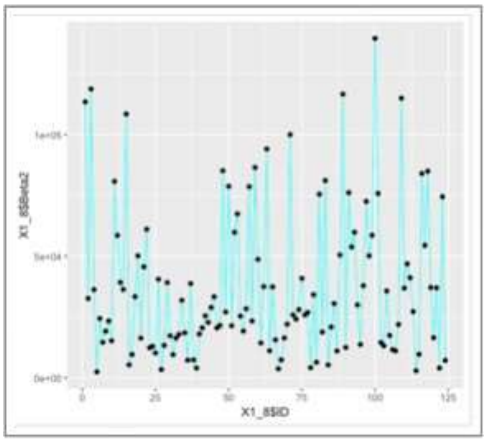


Fig 16. Student ID 0 Clip No 8

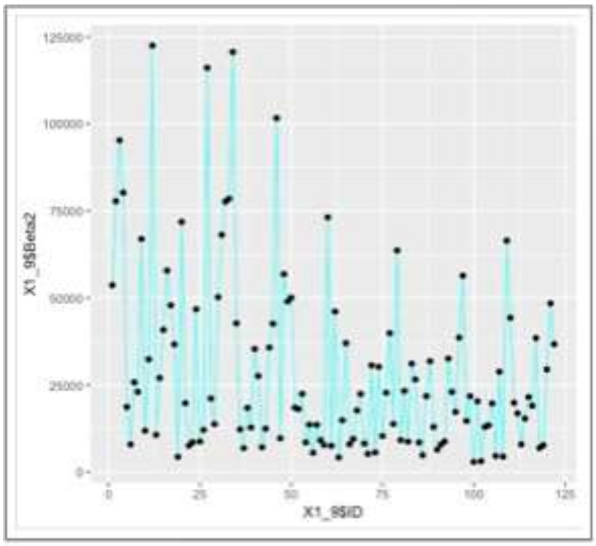


Fig 17. Student ID 0 Clip No 9

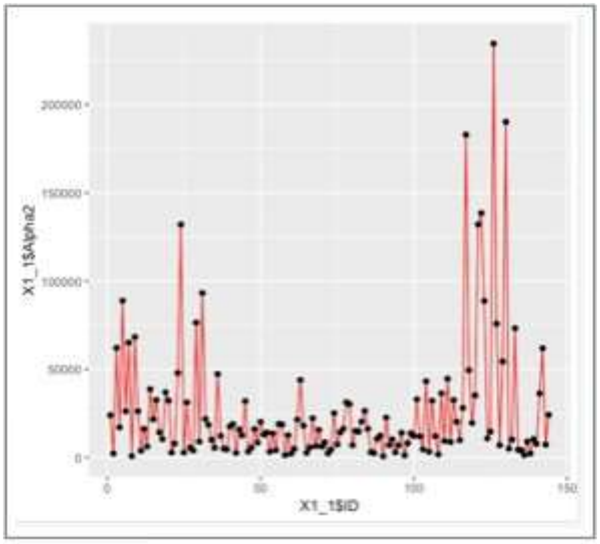


Fig 18. Student ID 0 Alpha Waves

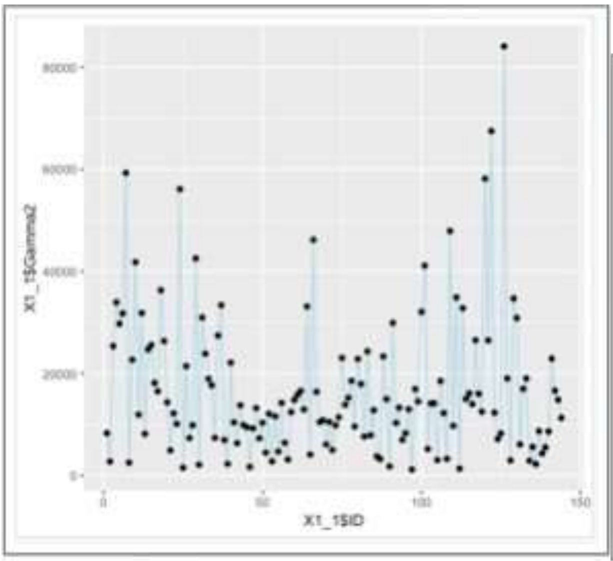


Fig 19. Student ID 0 Gamma Waves

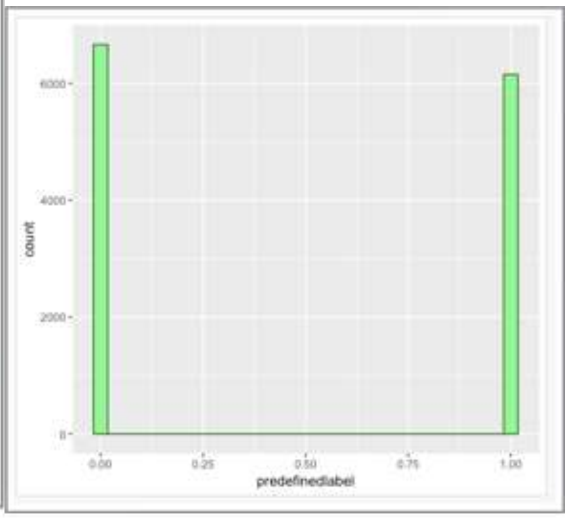


Fig 20. Predefined

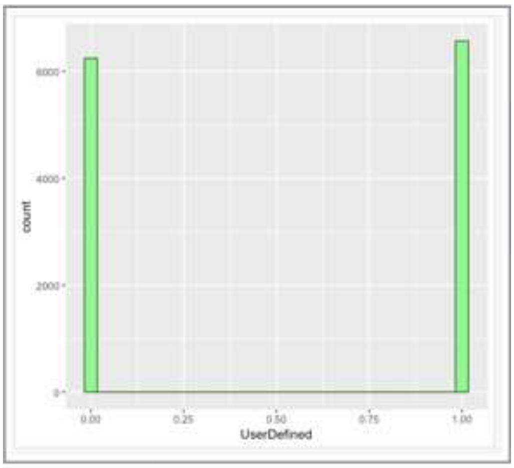


Fig 21. UserDefined

Inference

The following things were inferred from the graphs that were plotted above.

1. **Attention**: The graphs shows quite symmetrical pattern of waves, indicating that majority of student had decent levels of attention. However, a few spikes in case of very low concentration and low attention span were also noticed. This indicated that while most students maintain good levels of attention while watching videos, some of the students fall off the wagon. It was also

observed that some students paid no attention to provided online courses. This seems to be a pitfall for MOOC. While the understanding behind this could be the simple human nature to pay attention to the topics that are of particular interest to the student and disregard the rest, it can be concluded that the presence of a teacher pushes students to pay more attention, even to the topics that may not directly relevant, thereby possibly resulting in higher levels of information attained.

2. **Mediation**: Symmetry in waves indicated high level of calmness amongst the students.

   This indicates that MOOCs make students feel more comfortable as compared to classroom lectures. The presence of a teacher, especially someone the students would classify as strict or stern tends to make the students feels nervous and inattentive. This issue could be effectively tackled by MOOCs.

3. The in-depth analysis of the **brains waves** would require medical knowledge which is outside the scope of this research paper but on a surface level it can be understood that most students respond to the videos in a similar manner. The waves associated with alertness, amidst fluctuations, stay quite high while those associated with a deep level of sleep also tend to spike, indicating that simply watching videos can sometimes seem boring or make students sleepy. The sleepiness peaks when the student knows that the video series are about to end. The same peaking is observed in attentiveness so the result to this can be concluded as ambiguous, varying on an individual basis.

4. The difference in the pre-classification and user classification on the **level of difficulty** shows that although students enjoy a level of calmness with MOOCs, however, certain topics appear to be tough and difficult to understand without the assistance of a teacher.

Conclusion

From the inferences drawn from the graphs and the overall study of MOOCs, it can be concluded that majority of students are very receptive to the use of MOOCs and it will be a great addition to the regular curriculum being used as of now. The study of attention and mediation shows that students feel comfortable with this new and innovative means of education. That being said, a complete removal of teachers and physical learning is not supported because there are a lot of factors that are inculcated in students by the presence of a learned individual which simple videos cannot impart. An amalgamation of the two is be an ideal setting and the next productive step in the field of education, as is already being implemented in MIT and Harvard. A widespread adoption of the same is suggested.

References

1. “Study on MOOC scoring algorithm based on Chinese University MOOC learning behavior data” by Yong Luo, Guochang Zhou, Jianping Li, Xiao Xiao.
2. “Integrating MOOCs in traditionally taught courses: achieving learning outcomes with blended learning” by Antonia Bralić and Blaženka Divjak.
3. “Clustering patterns of engagement in Massive Open Online Courses (MOOCs): the use of learning analytics to reveal student categories” by Mohammad Khalil and Martin Ebner
4. “Deciphering the attributes of student retention in massive open online courses using data mining techniques” by Shivangi Gupta, A. Sai Sabitha·
5. “Massive Open Online Courses (MOOCs): Data on higher education” by Waleed Al-Rahmi, Ahmed Aldraiweesh, Noraffandy Yahaya, Yusri Bin Kamin, Akram M. Zeki
6. “Examining communities of inquiry in Massive Open Online Courses: The role of study strategies” by Vitomir Kovanovića, Srećko Joksimovića, Oleksandra Poqueta, Thieme Hennisb, Pieter de Vriesb, Marek Hatalac, Shane Dawsona, George Siemensa,f, Dragan Gaševićd
7. “Assessing metacognition in an online community of inquiry” by Akyol, Z., & Garrison, D. R.

8. "Your Click Decides Your Fate: Inferring Information Processing and Attrition Behavior from Mooc Video Clickstream Interactions" by T. Sinha, P. Jermann, N. Li, P. Dillenbourg.
9. "Investigating how stu- dent's cognitive behavior in MOOC discussion forums affect learning gains" by X. Wang, D. Yang, M. Wen, K. Koedinger, C.P. Rose.
10. "Participation patterns in a massive open online course (MOOC) about statistics" by L.P. Rieber.
11. "Decoding Engagement in MOOCs: An Indian Learner Perspective" by Kaveri, A., Gunasekar, S., Gupta, D., & Pratap, M.
12. "Factors leading to an effective MOOC from participants perspective" by Gamage, D., Fernando, S., & Perera.
13. "Academic motivation and self-regulation: A comparative analysis of undergraduate and graduate students learning online" by Artino, A. R., & Stephens, J. M.
14. "Gauging MOOC learners' adherence to the designed learning path" by Davis, D., Chen, G., Hauff, C., & Houben, G. J.
15. "Demographic differences in how students navigate through MOOCs" by Guo, P. J., & Reinecke, K.
16. Confused student EEG brainwave data, EEG data from 10 students watching MOOC videos by Hoahan Wang on Kaggle.com