Aspect-Based Sentiment Analysis Of Students Regarding The Indian Education System

Anupama G R
Department of Mathematics
Vellore Institute of Technology
Vellore, India

Prerana H Sulibhavi
Department of Mathematics
Vellore Institute of Technology
Vellore, India

S R V Prasad Bhuvanagiri
Department of Mathematics
Vellore Institute Of Technology
Vellore, India

***Abstract*— Education is intended to prepare students for life in all aspects: physically, intellectually, emotionally and spiritually. It should prepare students to overcome life's challenges with self-confidence and courage, make the appropriate decisions, and broaden their horizons. One of the primary purposes of education is to enable individuals to live a valuable and meaningful life. The primary beneficiaries of the education system are students. Students' opinions and feelings are significant sources of information that can be used to analyse their behaviour towards a course, topic, or teacher. In addition, they can be used to reform policies and institutions to improve their performance. So, we aim to analyse the Indian education system based on students' sentiments through a survey. We will discuss the different aspects of the education system, including course content, examinations, learning methods, stress and their future.**

***Keywords*— NLP, aspect-based Sentiment analysis, tokens, lemmatisation, Exam Pattern; Course; Preference; Stress; Learning Methods- Flipped Classroom, Active Learning, Project-Based, Hands-on training, Gamified learning.**

I. INTRODUCTION

Education prepares students for life in all aspects: physically, intellectually, emotionally, and spiritually. In addition to preparing students to face life's challenges with self-confidence, courage, and wise decisions, it should also broaden their horizons and help them achieve their highest potential throughout their lifetime. Students are the main beneficiaries of education. Students' attitudes towards the system are significant, as they impact their performance and society as a whole. In addition, sentiment analysis can be used to identify aspects that require attention and improvement. To gain a better understanding of students' feelings about different aspects of the system, we surveyed the age group of 18 to 28. Based on their views, we performed sentiment analysis. We have used statistical tools to understand the relationship between aspects and derive concrete conclusions [6].

Human language is characterised by various irregularities, including homonyms, homophones, sarcasm, idioms, metaphors, grammar and usage exceptions, variations in sentence structure, etc. For natural language-driven applications to be practical, programmers must teach them to recognise and understand natural language accurately.

NLP analyses different aspects of human language, including syntax, semantics, pragmatics, and morphology, to understand its structure and meaning. NLP tools use text vectorisation to convert text into something a machine can understand. Then machine learning algorithms develop algorithms to identify relationships between inputs and outputs (tags) by studying training data and expected outputs (tags). For the machine to make predictions regarding anonymous data, statistical analysis is used to construct a "knowledge bank" [7].

In sentiment analysis, reviews assess people's attitudes towards a particular topic. An opinion is generally a combination of individuals' feelings, beliefs, thoughts, sentiments, and desires. Based on the student comments, this study analyses their view using sentiment analysis approaches and categorises them as positive, negative or neutral. Depending on the context, sentiment analysis can be conducted at the level of words, sentences, or documents. Considering many records, more than manual sentiment analysis is needed. Therefore, automatic data processing is necessary. Sentiment analysis from text-based or document-level corpora is employed using natural language processing (NLP).

Survey sentiment analysis is a method for identifying the sentiment behind survey responses and is most often used in the analysis of open-ended surveys. It is crucial to measure a respondent's emotional connection to education. Happy people are more likely to succeed, so the more connected they feel, the more comfortable they are [3].

II. RELATED WORKS

In the domain of education, the sentiment analysis approach allows us to understand the hidden knowledge to be explored. Open-ended questions in the evaluation process provide students with the opportunity to voice their opinions [6]. A majority of the literature focuses on quantitative data (numerical ratings) while ignoring qualitative data (student comments). Qualitative data enables efficient feedback that helps us to make better decisions. The tools and aspects of the education system have been extracted from a few related works.

This paper aims to draw insights from responses to a survey conducted on students about the Indian education system [1]. Based on sentiment analysis, we are attempting to analyse the different aspects of the system. We have adopted the workflow described in the paper [4], which is shown in figure 1. Following data pre-processing, we performed unsupervised sentiment analysis [7]. A lexicon-based analysis of the reviews was conducted using Textblob. We used the help of different online sources to implement the

same [10,12,14,15]. For suggestive methods, we have used methods adopted by Imperial College London [17].
This paper represents a combination of different ideas that aim to understand and analyse the Indian educational system in light of student sentiment.

III. Methodology

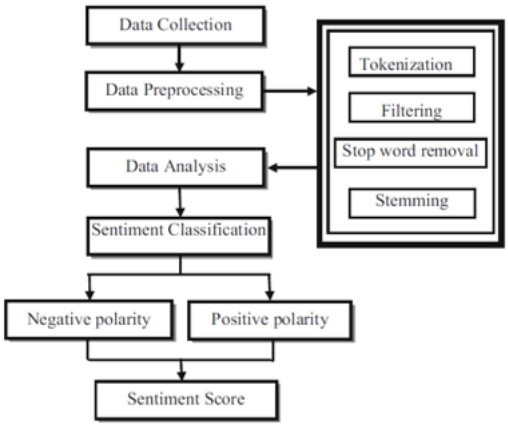

Figure 1. Flow chart of the methodology for the analysis [4]

A. Data Collection

Data were collected by conducting a survey. A questionnaire was framed, and responses were collected through google forms and personal interviews. The sampling technique adopted is convenience sampling. The population under consideration is respondents who had passed higher secondary school between the age of 18 to 28. The questionnaire included both open-ended and closed-ended questions. It considers different aspects such as course selection, examination, stress, learning methods and plans for the future.

B. Data Cleaning/Data Preprocessing

During this phase, the sentimental analysis system prepares collected data for further processing. This involves the following steps. [2,4]

- **Lowercasing**: We convert each word to lowercase to remove case sensitivity
- **Removal of irrelevant content** (Remove punctuation and emojis): All the punctuation marks, numbers and emojis are removed to obtain only words.
- **Removal of stop words**: A stop word is a commonly used word such as ‘the’, ‘a’, ‘an’, or ‘in’ Removal of stop words reduces the dataset size and thus reduces the training time due to the fewer number of tokens involved in the training.
- **Tokenization**: The students’ reviews are split into words or tokens using the tokenise function in Python.
- **Lemmatization**: usually refers to doing things properly with the use of a vocabulary and morphological analysis of words, typically aiming to remove inflectional endings only and to return the base or dictionary form of a word, which is known as the *lemma* [8,9,10].

IV. Data Analysis

To facilitate analysis of student reviews about the course, stress, examination, learning methods, and future plans. The following approaches have been used.

A. Visualizations

Data visualisation is the graphical representation of information and data. In our analysis, we have used Bar charts, pie charts, tree maps, Word Clouds, Frequency tables, multiple bar charts, and network-weighted graphs.

B. Text Blob

The Text-Blob function returns the polarity and subjectivity of a sentence. Polarity is defined as a feeling between [-1,1]. -1 represents a negative feeling, whereas 1 represents a positive one. The polarity of negative words is reversed. A fine-grained analysis is made possible by Text-Blob's semantic labels. Subjectivity can be defined as the amount of personal opinion in the text that is accompanied by factual information. A text with a higher subjectivity contains subjective opinion rather than an accurate description. By the polarity and subjectivity score, the text is either positive, neutral, or negative. Based on the polarity and subjectivity score, the text is classified as positive, neutral, or negative [13,14,15].

C. Word Cloud

A word cloud is a collection, or cluster, of words depicted in different sizes. The more prominent and bolder the word, the more often it appears within a given text and the more significant it is. A word cloud is a powerful tool for visualising what your audience thinks about something. They are easy to read, quick to produce, and easy to understand [11].

D. Vectorisation

As part of Natural Language Processing, Word Embeddings or Word Vectorization help predict words, find word similarities, and extract semantic information. Mapping words or phrases make these maps from the vocabulary to corresponding vectors of real numbers. The conversion of words into numbers is called vectorisation. We used Word2Vec to convert words into vectors [11].

E. Cosine Similarity

Mathematically, this function measures the cosine of the angle between a pair of vectors (item1, item2) projected in an N-dimensional vector space. By using it, you can predict the similarity between the two documents. The smaller the angle, the higher the similarity, according to a cosine [9].

F. Clustering

Clustering refers to grouping objects so that they are more similar to one another than objects from different clusters. Clusters are formed based on criteria such as smallest distances, the density of data points, graphs, or statistical distributions [11].

V. Definitions

A. *Flipped Classroom:*

The reverse of traditional lectures. The lecture materials are read at home, self-learnt, problems are solved, and doubts are cleared in classrooms [16].

B. *Active Learning:*

It's group-based learning. Active learning methods ask students to engage in their learning by thinking, discussing, investigating, and creating [16].

C. *Project-based Learning:*

Learning through doing projects and engaging them with real-life problems [16].

D. *Hands-on Learning:*

Hands-on learning means learning by doing [16].

E. *Gamified Learning:*

Learning through gamification involves incorporating game-based elements like point scoring, peer competition, teamwork, and score tables to drive engagement and assimilate new information [16].

VI. Result And Calculations

Based on our analysis, we obtained the following results:

A. ASPECT 1:COURSE/PROGRAM CHOSEN

Table 1. Frequency count of respondents' feelings towards their course/program.

Are you happy with program what you studied /studying?	Frequency
Yes	106
No	25
Maybe	15

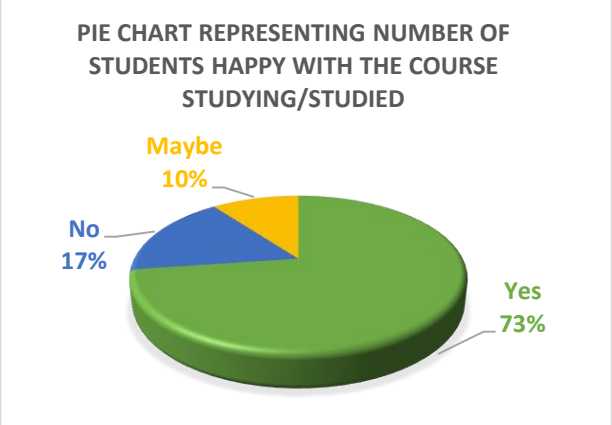

Figure 2. Pie Chart representing the number of students who are happy with the course studying

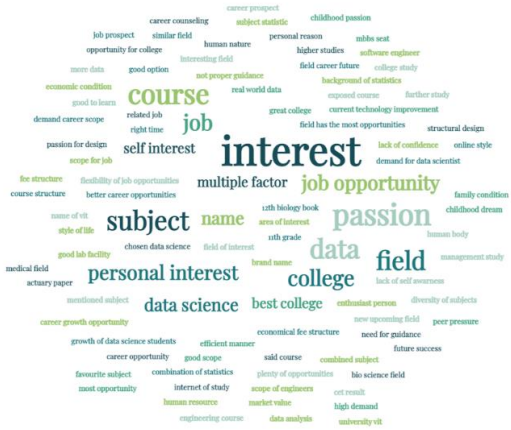

Figure 3. Cloud chart highlights the reasons behind the choice of the course.

From figure 2 we observe that 73% of the respondents are happy with their current course, 17% are unhappy, and 10% are unsure. Figure 3 highlights the reasons behind the choice of course. The most common reasons are Interest, passion, subject or course, and job opportunities class most specified reasons.

Table 2. Frequency count of respondents' course preference.

Is your chosen course/program your first preference? If not, what was it?	FREQUENCY	PERCENTAGE
Yes	97	67.12328767
No	49	32.87671233
	146	100

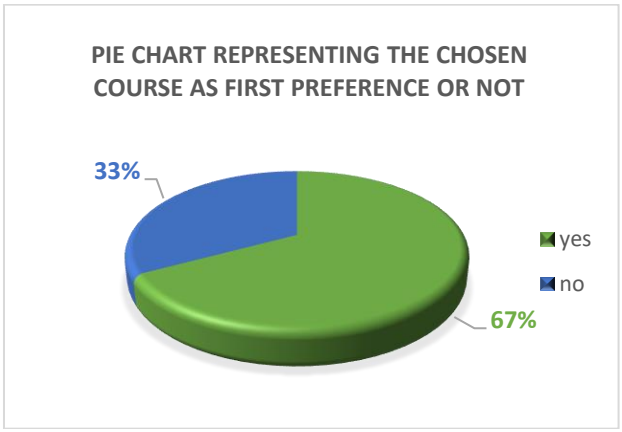

Figure 4. Pie Chart representing the chosen course is their first preference or not.

From figure 4 we observe that 67% of the respondents' chosen course is their first preference and 33% is not.

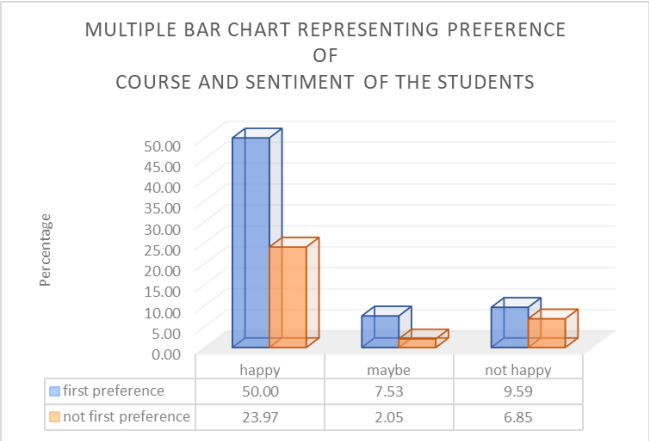

Figure 5. Multiple bar chart representing a preference for the course and sentiment of the students.

From Figure 5 we observe that 50% of the respondents are happy with their first choice of course, while 9.5% are unhappy with it. Also, 24% of the respondents are satisfied with their second choice of course, while 7% are not sure about the choice.

Table 3. Frequency count of respondents' feelings towards the course based on their preference.

	happy	maybe	not happy	
first preference	73	11	14	
not first preference	35	3	10	
				146

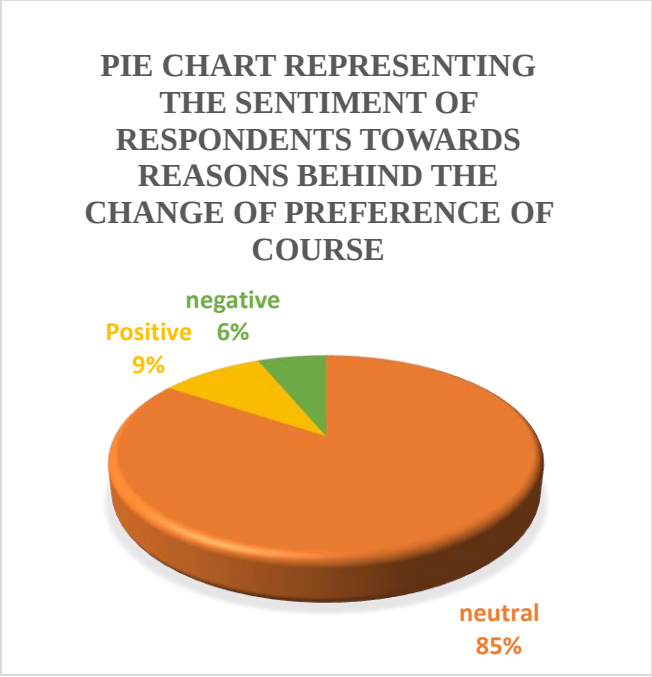

Figure 6. Pie Chart representing the sentiment of respondents towards reasons behind the change of course.

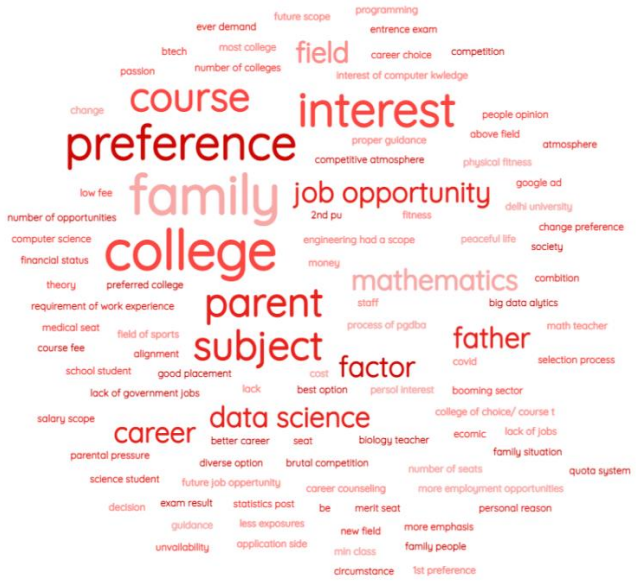

Figure 7. Word Cloud represents the factors for change of course

Figure 6 we observe that 9% of the respondents have a positive reason for the change of preference and 6% have a negative reason and 85% of the total respondents have a neutral reason. Figure 7 highlights the main factors for change of the course such as family, course preferred not available, job opportunities and student interest.

Based on the above graphs, we can conclude that the majority of respondents are satisfied with their course. Upon analysing the reasons for their choice, we find that it is a combination of personal interest in the subject, passion, curiosity, course structure, colleges, employment opportunities, and future prospects. In the case of respondents who are not satisfied, the reasons include economic or family circumstances, lack of confidence and self-awareness, competition, and lack of availability of seats in the preferred course or college.

In light of their preferences, let's compare their sentiments. The results of the survey indicate that 10% of respondents are dissatisfied with their first preference. This may be due to a lack of self-awareness, inadequate guidance, family conditions, the quota system, economic status, or other personal reasons. According to our survey, 24% of respondents are satisfied with their change of course. The reasons given are career counselling, proper guidance, passion for the subject, and current technological advancements.

Our recommendation is to provide individualized career counselling to high school students. As a result, their interests and passions will be incorporated into their careers, allowing them to reach their full potential. This will help expose them to different fields rather than limiting them to

mainstream occupations. In turn, this will reduce competition and provide better opportunities for everyone.

B. ASPECT 2:STRESS

Table 4. Frequency and percentage of respondents score of stress

	COUNTS	Percentage
2	3	2.05
3	6	4.11
4	7	4.79
5	21	14.38
6	16	10.96
7	16	10.96
8	34	23.29
9	15	10.27
10	28	19.18
	146	

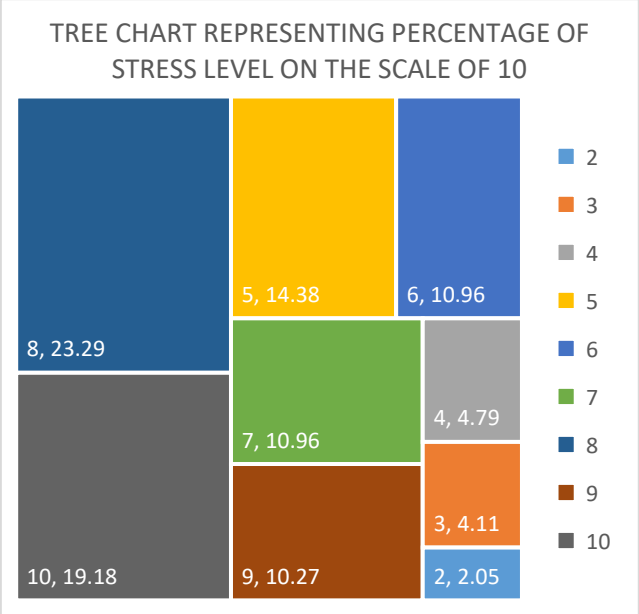

Figure 8. Tree chart representing the percentage of stress levels on a scale of 10.

52% of the respondents feel a pressure higher than a scale of 8, while 26% feel less stressed (less than 5).

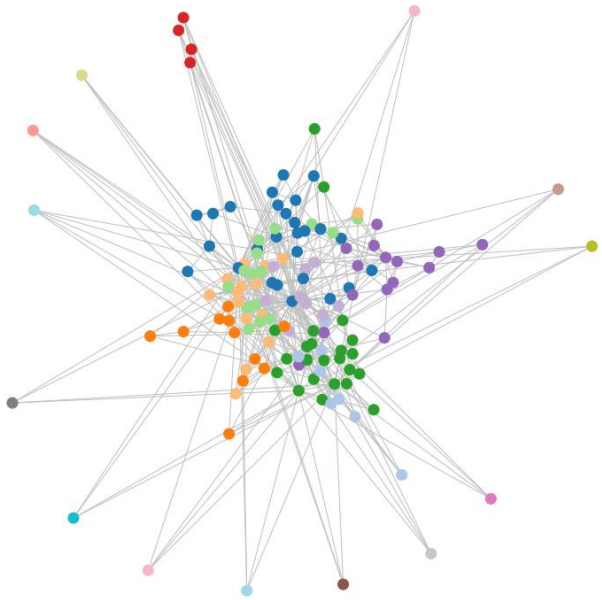

Figure 9. Cluster diagram representing the different clusters of mechanisms that student adopts to cope with their stress.

Figure 9 represents the different mechanisms that students adopt to cope with their stress. Nodes of the same colour belong to the same cluster. Concepts behind the clusters are represented in the form of a word cloud in figure 10[12,11].

Figure 10 Word Cloud represents the activities to reduce stress.

From Figure 10 we infer that the common mechanisms that students adopt to cope with stress are listening to songs, spending time with friends and family, meditation, hobbies, watching movies/series, sports and me-time or self-care.

Regardless of their educational qualifications, 50% of respondents face high pressure. So, to understand how they handle such high pressure, we have analysed their coping mechanisms. It appears that distractions, outdoor activities, calming ourselves, and talking to loved ones are some of their main mechanisms.

We recommend that schools and colleges provide equal emphasis on extracurricular activities. Several respondents report feeling better after spending time or talking with loved ones. Spending quality time together with family should be a priority. In this regard, we come to mental health. Talking and discussing the problems that students face can itself be a coping mechanism. It is therefore important to encourage therapy sessions in colleges and families. A calm mind and healthy body increase the efficiency of an individual. Meditation, yoga, and exercise are all effective ways of achieving this.

To maintain a high level of productivity, quality breaks should be emphasised. These actions should be viewed as coping mechanisms, not as guilt-inducing. Studies show that taking breaks between studying (anywhere between 5–60 minutes) increases your productivity, energy, and focus.

As research shows, social media does not work well as a "purposeful break". Instead, engage in activities that allow you to remove your mind from the everyday grind and allow you to breathe deeply, laugh, move your body, be creative, or "zone out" on purpose. Taking part in these types of activities will help you re-energize and regain your focus.

C. ASPECT 3: EXAMINATION PATTERN

Table 5. Frequency count of respondent's sentiment towards examination pattern.

Sentiment towards examination pattern	Frequency
Neutral	87
Positive	41
Negative	18
	146

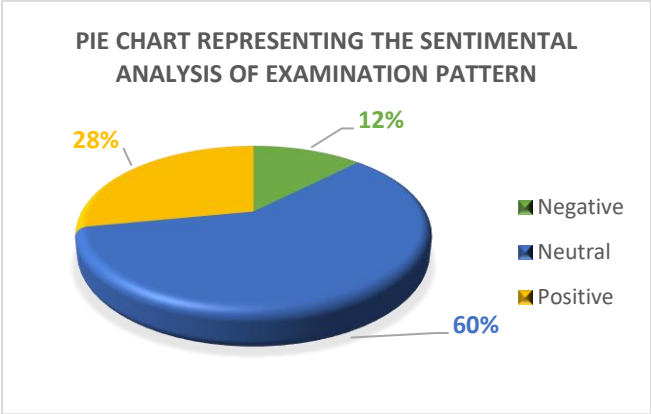

Figure 11. Pie chart representing the sentiment analysis of examination pattern.

From Figure 11 we infer that 60% of the respondents have a neutral feeling towards the examination pattern, 28% feel optimistic about it, and 18% feel negative about it.

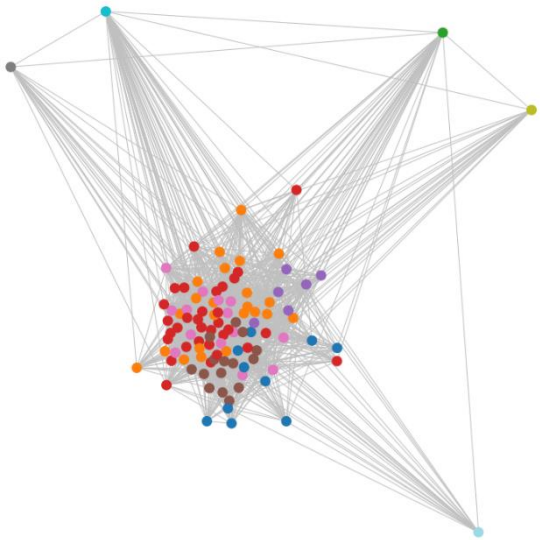

Figure 12. Cluster Diagram representing the different perceptive of students towards the examinations.

Figure 13. Word cloud represents the views towards the examinations.

Table 6. Frequency count of respondents' sentiment towards university entrance examination concerning education qualification.

	Sentiment towards university entrance examination			
Education Qualification	**Negative**	**Neutral**	**Positive**	**Grand Total**
HSC	5		3	8
PG	12	32	33	77
UG	15	24	20	59
Grand Total	**32**	**56**	**56**	**144**

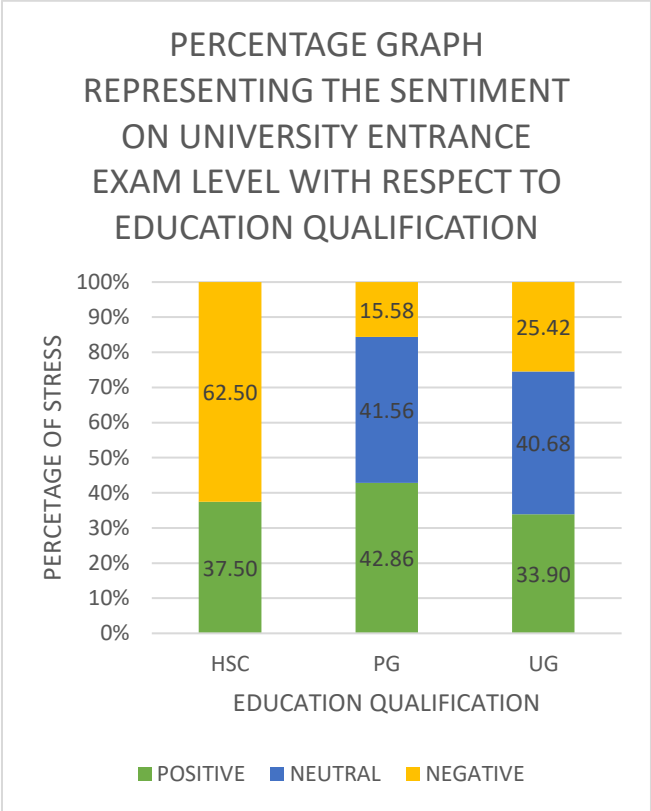

Figure 14. Percentage Graph representing the sentiment on university entrance exam level concerning education qualification.

Figure 14 infers that 62.5% of higher secondary school students have a negative feeling towards university entrance exams. While 43% of the postgraduate students feel positive, and 41% of undergraduate students think neutral[12,11,].

D. ASPECT 4: LEARNING METHODS

Table 7. Frequency count of respondent's score towards their comfort level when a friend or instructor teaches

How comfortable are you when a friend or an instructor teaches you?	
Scale	**Frequency**
2	1
3	3
4	3
5	7
6	12
7	23
8	42
9	25
10	30
Grand Total	**146**

Table 8. Frequency count of respondent's score towards their comfort level when a friend or instructor teach

How comfortable are you when you self-learn?	
Scale	**Frequency**
1	1
3	3
4	5
5	7
6	12
7	23
8	37
9	39
10	19
Grand Total	**146**

Table 9. Number of respondents who are comfortable being taught by an instructor or friend on a scale greater than 7

INSTRUCTOR/FRIEND					
7	23				23
8		42			42
9			25		25
10				30	30
					120
					82.19178

Table 10. Number of respondents who are comfortable being self-taught on a scale greater than 7

SELF LEARNING					
7	23				23
8		37			37
9			39		39
10				19	19
					118
					80.82192

82% of respondents are comfortable when an instructor/friend teaches, while 81% are comfortable with self-learning. So, we can infer that both ways are important and preferred by students.

TABLE 11. Representing the percentage of ranking of different learning methods.

Learning Methods/Ranks	1	2	3	4	5
Flipped Classroom	28.57	11.90	10.32	15.87	33.33
Active Learning	15.87	34.92	22.22	19.84	7.14
Project Based	22.22	18.25	30.16	19.84	9.52
Hands On Learning	18.25	22.22	22.22	22.22	15.08
Gamified Learning	15.08	12.70	15.08	22.22	34.92

Table 11 we observe that 29% of the respondents ranked flipped classrooms 1.16% of the respondents ranked active learning 1.22% of the respondents ranked project-based learning 1. 18% of the respondents ranked hands-on learning as 1. 15% of the respondents ranked gamified learning as 1.

According to respondents, current university examinations lack practical applications and are largely theoretical. There is not much application and it is mostly memory-based. In their view, there is a gap between the course material and its application in industry. To resolve this, we have taken the suggested methods from the paper [17] to incorporate different teaching methods. Imperial College London has experimented with different methodologies based on their course.

One prominent example is the flipped classroom, in which the student is asked to do research and acquire basic knowledge about the subject matter and answers and discussions are conducted in class. Another type is active learning, which involves group learning. Discussions, questions, and teamwork are all part of the learning process. As a result, students remain engaged and active throughout the lesson. Students learn by experimenting and exploring on their own through hands-on learning and project-based learning. To develop their skills, they are given different projects to work on. Besides enhancing their skills, this will also give them a glimpse of the industry. In gamified learning, game-based elements like point scoring, peer competition, teamwork, and scores are incorporated to engage students and help them assimilate new information.

Lecturers can adopt these methods according to what is favourable for their subject, helping students to better understand it. In addition, they can suggest alternative testing methods to traditional ones for examining students [17].

E. ASPECT 5: FUTURE PLANS

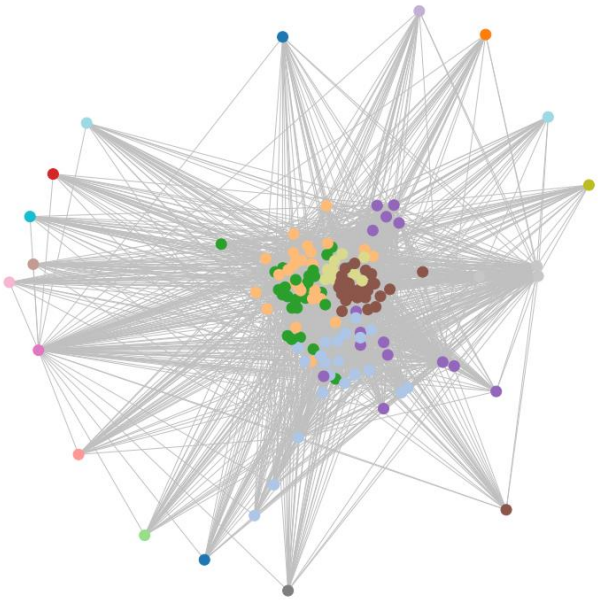

Figure 15. Cluster Diagram representing the different perceptive of students toward their future plans.

Figure 16. Word Cloud represents their near future plans

Figure 17. Word Cloud represents the future plans after their education.

The respondents wanted to gain knowledge, be happy and have a peaceful life, a promising career, earn money and improve society with their education [11,12,16].

The main focus of the respondents for their near future is to finish their higher education or get into a decent job and develop their career. So, each respondent is looking at their next big step. Let us now look at another interesting aspect of their education: their purpose. Most respondents desire to make society a better place, gain knowledge and earn well. Taking a closer look at the definition of success, we find answers such as luxurious living, financial independence, fulfilling self-expectations, career growth, achieving their full potential, and living a peaceful and fulfilling lifestyle. Interestingly, the current generation is interested not just in monetary gains, but also in the development of society as a whole. The common perspective of respondents about current university examinations is they are theoretical and lack practical implications. It is mainly memory based and not much of an application. They find a lag between the course and industrial application.

VII. DISCUSSION AND CONCLUSIONS

Let us summarise this paper from the perspective of an Indian student. The decision to choose a career starts in high school. So when students are provided with proper guidance and are

made aware of different options, they can choose what is right for their interests and capabilities. This will help them reach their full potential and also reduce unnecessary competition in mainstream professions. As stress is an inevitable part of life, we can handle it by incorporating extracurricular activities into the curriculum. In addition, we can encourage students, society, and parents to care for their mental health.

To enhance the course outcome and bridge the gap between the curriculum and industry we can incorporate different learning methods as proposed in the paper. Accordingly, the examination pattern can be revised to test the students' skills and learning in order to enable them to evaluate their skill set and improve it.

Since education aims at the overall development of students, improving every aspect of the education system at every stage can help in building a better future for the student and society as a whole.

REFERENCES

[1] Rani, S., & Kumar, P. (2017). A Sentiment Analysis System to Improve Teaching and Learning. Computer, 50(5), 36–43. doi:10.1109/mc.2017.133

[2] Mujahid, M.; Lee, E.; (2021) Rustam, F.; Washington, P.B.; Ullah, S.; Reshi, A.A.; Ashraf, I. Sentiment Analysis and Topic Modeling on Tweets about Online Education during COVID-19. 11, 8438.

[3] Naveen Kumar Laskari1, Suresh Kumar Sanampudi,(2016). Aspect-Based Sentiment Analysis Survey, March

[4] Verma, B., & Thakur, R. S. (2018). Sentiment Analysis Using Lexicon and Machine Learning-Based Approaches: A Survey. Proceedings of International Conference on Recent Advancement on Computer and Communication, 441–447. doi:10.1007/978-981-10-8198-9_46

[5] Jain, P. K., Pamula, R., & Srivastava, G. (2021). A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. Computer Science Review, 41, 100413. doi:10.1016/j.cosrev.2021.100413

[6] Chauhan, G. S., Agrawal, P., & Meena, Y. K. (2018). Aspect-Based Sentiment Analysis of Students' Feedback to Improve Teaching–Learning Process. Smart Innovation, Systems and Technologies, 259–266. doi:10.1007/978-981-13-1747-7_25

[7] Kastrati, Z., Dalipi, F., Imran, A. S., Pireva Nuci, K., & Wani, M. A. (2021). Sentiment Analysis of Students' Feedback with NLP and Deep Learning: A Systematic Mapping Study. Applied Sciences, 11(9), 3986. doi:10.3390/app11093986

[8] M. Young(1989), The Technical Writer's Handbook. Mill Valley, CA: University Science.

[9] Sentiment Analysis For Brand Building; https://www.revuze.it/blog/sentiment-analysis/#getting_started_with_sentiment_analysis:_the_four_main_steps;last_accessed:07-nov-2022

[10] Lemmatization in Natural Language Processing (NLP) and Machine Learning; https://towardsdatascience.com/lemmatization-in-natural-language-processing-nlp-and-machine-learning-a4416f69a7b6; last_accessed:07-nov-2022

[11] What is a Word Cloud and Why Should I Use One?; https://www.participoll.com/what-is-a-word-cloud-and-why-should-i-use-one/; last_accessed:07-nov-2022

[12] How to cluster similar sentences using TF-IDF and Graph partitioning in Python; https://towardsdatascience.com/how-to-cluster-similar-sentences-using-tf-idf-and-graph-partitioning-in-python-5fb70627a4e0; last_accessed:07-nov-2022

[13] Sentiment Analysis in Python: TextBlob vs Vader Sentiment vs Flair vs Building It From Scratch; https://neptune.ai/blog/sentiment-analysis-python-textblob-vs-vader-vs-flair; last_accessed:07-nov-2022

[14] Rule-Based Sentiment Analysis in Python; https://www.analyticsvidhya.com/blog/2021/06/rule-based-sentiment-analysis-in-python/?; last_accessed:07-nov-2022

[15] Sentiment classification using NLP With Text Analytics; https://www.analyticsvidhya.com/blog/2021/09/sentiment-classification-using-nlp-with-text-analytics/; last_accessed:07-nov-2022

[16] Sentiment Analysis of Survey Responses; https://monkeylearn.com/blog/sentiment-analysis-of-survey-responses/; last_accessed:07-nov-2022

[17] Innovative teaching for world class learning Learning and Teaching Strategy; https://www.imperial.ac.uk/media/imperial-college/about/leadership-and-strategy/vp-education/public/LearningTeachingStrategy.pdf; last_accessed:07-nov-2022