CHAPTER - 3

QUESTIONNAIRE SENTIMENT ANALYSIS THROUGH POLARITY

3.1 Introduction

Sentiment analysis is a method for discovering emotions and assessing attitudes towards specific entities, such as individuals, concepts, or activities. It involves the identification of people's viewpoints and the categorization of their sentiments as favorable, unfavorable, or neutral. Sentiment analysis methods leverage methods of machine learning and natural language processing (NLP) to discern, extract, and consolidate insights and opinions from vast volumes of text.

NLP aims to create new computational capabilities surrounding human language, including information retrieval, machine translation, text creation, and other tasks [137]. Text sentiment analysis is a crucial application of NLP that involves classifying sentiment implications and attitudes in subjective texts with sentiment overtones. Applications for sentiment analysis have spread to practically every conceivable industry, including consumer goods and services, healthcare, finance, and social events [138]. Governments can successfully understand public opinion trends by assessing the sentiment orientation in these comments and providing a basis for policy making. Businesses can use customer feedback information to generate more precise marketing tactics. Regular consumers can learn what other users think of a product before making smarter purchasing selections [139].

Sentiment analysis alternatively referred to as opinion mining, aims to grasp a user's stance and viewpoints by exploring, examining, and extracting subjective content that encompasses their opinions, preferences, and sentiments. This intricate procedure encompasses various additional stages such as data processing, transformation, and dimensionality reduction [140]. It also entails model training. The Figure 3.1 illustrates the schematic representation of sentiment analysis.

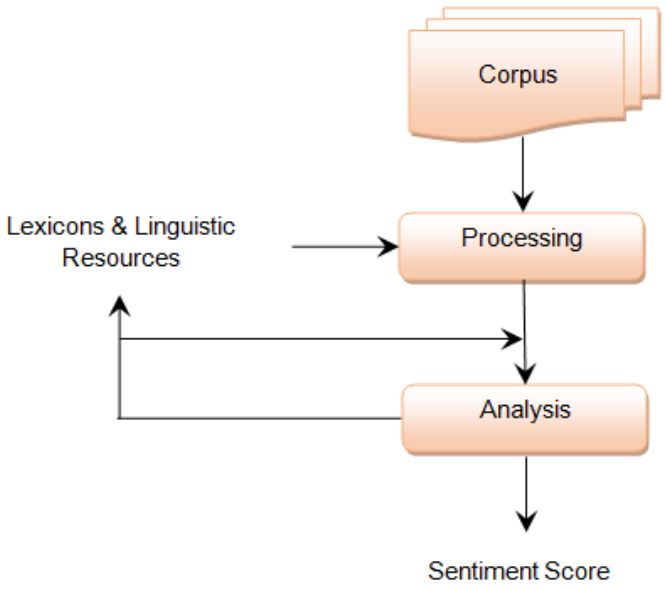

Figure 3.1 Sentiment Analysis Architecture [141]

The educational community has paid much attention to sentiment analysis and opinion mining [1]. Research on sentiment analysis in education encompasses multiple perspectives, including those of educational institutions, teachers, instructors, learners, students, and decision-makers. By examining students' opinions and behaviors towards specific educational programs, platforms, academic establishments, and instructors, sentiment analysis is commonly employed to enhance administration, instruction, and evaluation processes.

The application of sentiment analysis is explored in studying the connection between learners' emotions and dropout rates in enormous open online courses, as well as the correlation between academic achievement, retention, and student's emotional responses [142]. After every academic term, students are typically required to complete post-course questionnaires so that instructors can learn more regarding their knowledges. This process enables educators to examine student evaluations and improve learning experiences. The survey includes both closed- and open-ended inquiries. Closed-ended queries, like Likert scale queries, attempt to quantify student

evaluations using numerical ratings. Responding to open-ended questions, students may write comments or thoughts expressing their unique viewpoints and perspectives.

This chapter identifies the student's emotions using sentiment analysis through closed-ended questions. Students' feedback is collected through a questionnaire with predefined response options, and their emotions are analyzed utilizing a sentiment analysis algorithm based on polarity. The results were assessed utilizing emotion variance and sentiment scores. Based on the sentiment ratings, students expressed a range of positive feelings, including unhappiness, moderate happiness, and joy.

3.2 Closed-Ended Questionnaire and Sentiment analysis

A survey is a form of research instrument comprising a set of questions targeted at a specific demographic for response. Its purpose is to collect relevant data to facilitate informed decision-making based on your research. The primary objective of a survey is to obtain information from the participants. Even when the investigator is not there to record those comments directly, it can be a relatively cheap, rapid, and effective technique to get a lot of data.

A survey is merely a collection of inquiries crafted to solicit statistically significant data from the respondents. It's frequently regarded as a crucial tool for conducting surveys. A survey is a procedure that involves using a questionnaire to pose questions, gather answers, and then examine them to get a conclusion. A survey's crucial analysis and evaluation components set it apart from a questionnaire.

Questionnaires can be broadly classified into two main types: structured and unstructured. The quasi-structured questionnaire, commonly employed in social science research, is a blend of both types.

- Structured questionnaires commonly include pre-coded questions with well-defined skip patterns. They are widely utilized in most quantitative data collection endeavors. These well-organized questionnaires provide various advantages, including reduced discrepancies, ease of administration, consistent responses, and straightforward data handling.
- Unstructured questionnaires often contain open-ended and ambiguous opinion-type inquiries. If the question is not in the form of an interrogative sentence, the moderator or enumerator may need to provide an explanation of its meaning. Such questionnaires are commonly employed during focus group sessions.
- Not every question can have pre-coded alternatives with nearly certain options. In some typical surveys, certain answer choices are left as 'others' and respondents are requested to specify. Most questions are structured, which is a conventional and efficient approach. However, it is acceptable to include some unstructured questions that may have challenging responses to fully enumerate. A quasi-structured questionnaire adopts this type of format.

In unstructured, open-ended questionnaires, the qualitative questions are more straightforward, enabling the target audience to openly share their thoughts and emotions. As there are no predefined answer choices, respondents can freely express their ideas, resulting in the capture of more comprehensive data. However, since individuals often hold different viewpoints, dealing with qualitative questions can be challenging from the researcher's standpoint, making the processing of qualitative data difficult.

Often, open-ended queries are asked during the interview to gather information about the respondents' personal life, such as their values, relationships with their families, and motives and to elicit opinions or viewpoints regarding a subject by asking questions. Despite its numerous

advantages, interpreting and consolidating results into reports can be challenging for open-ended surveys. The extensive number of questions may lower the response rate, and there is a possibility of obtaining unrelated data through the open-ended questionnaire.

Structured closed-ended questions offer several options for respondents to choose from, and they are asked to select one or more options. Close-ended questions are commonly used in conducting quantitative research. These questions are targeted and intentionally planned to elicit as many comments and details as possible from responses. Open-ended questions are frequently used to gather feedback on a service, policy, or other topic and to gather the information that may be quickly organized into options. Hence, this study focuses on structured closed-ended questionnaires. Sentiment analysis is the process of evaluating emotional expression through language by obtaining dynamic information, analysing and interpreting it, and categorising text. Sentiment analysis can assist educators in the educational sector in quickly identifying students' thoughts about a course and making required and timely changes to the lesson plan to improve the level of instruction and learning. The fundamental goal of this research is to is to evaluate learners' feelings, self-worth, and effectiveness using closed-ended questionnaires. Through sentiment analysis, it becomes possible to recognize students' emotions, academic performance, behavior, and other related factors.

3.3 Proposed Methodology

Sentiment analysis is an analytical investigation that assesses person’s perceptions of various entities, such as people, events, objects, concepts, activities, and their attributes. Its aim is to automatically uncover the underlying viewpoints regarding a specific entity. Sentiment analysis is mostly employed in business contexts, including public relations, marketing analysis, and product reviews [143]. In addition, sentiment analysis is used in the education study to ascertain a student's emotions. Therefore, sentiment

analysis in education can improve learning and performance, decreasing study abandonment, the teaching process, and course satisfaction.

The standard definition of emotion refers to an individual's mental state, encompassing their attitudes, feelings, and behaviors. Presently, determining the public sentiment in specific situations has become straightforward, as opinions can be extracted from freely obtainable data on sites like Instagram, Twitter, Youtube, and Facebook. As a novel area of research in Natural Language Processing (NLP), emotion detection on platforms like Reddit, Twitter, and others is gaining popularity. Additionally, this approach finds applications in various fields, comprising education (to detect student dissatisfaction) and healthcare (as an aid for psychoanalysis) [144]. This research introduces "Quest_SA," a method that employs Sentiment analysis with polarity assessment in a questionnaire with predefined response options to uncover students' affective traits. Figure 3.2 illustrates the structural design of the proposed Quest_SA.

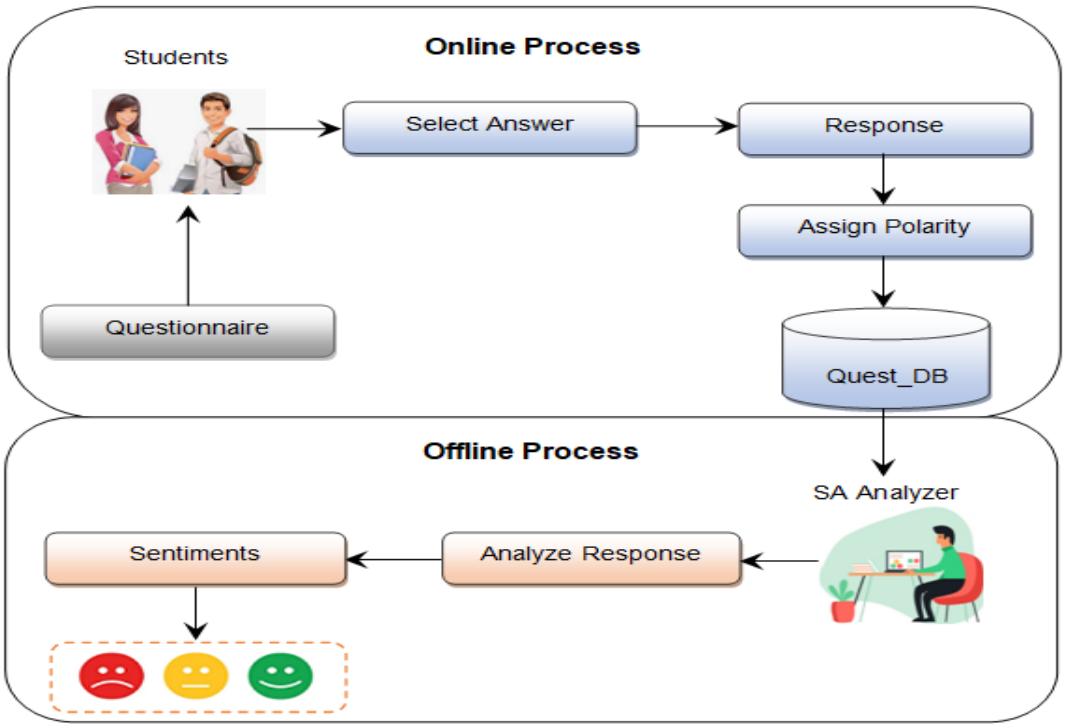

Figure 3. 2 Architecture of "Quest_SA"

The proposed Quest_SA is divided into 2 stages: online and offline stages. In the online phase, students are invited to participate in a questionnaire with predefined response options. The questionnaire covers seven types of inquiries, including EI,Personality,Well being, SE,SDS,GSE,PANA. Example questions are listed in Table 3.1.

Table 3. 1 Example Questions

Inventory	Example Questions	Scale
Emotional intelligence	1. I promptly recognize when I become angry. 2. I can swiftly reframe unfavorable situations. 3. I consistently find ways to motivate myself for challenging tasks. 4. I am consistently capable of understanding things from the other person's perspective. 5. I excel at listening attentively.	Not at all Rarely Sometimes Often Very Often
Eysenck personality	1. Is your mood frequently fluctuating? 2. Do you pay significant attention to what others think? 3. Would you describe yourself as a talkative person? 4. If you make a commitment, do you always honor it, even if it proves inconvenient? 5. Do you sometimes feel utterly miserable without any apparent reason?	Yes No
Self-determination scale	1. A. I consistently sense that I make my own choices. B. At times, I feel like my decisions are not entirely my own. 2. A. Occasionally, my emotions appear foreign to me. B. Generally, my emotions feel like they belong to me. 3. A. I willingly opt to do what needs to be done. B. Although I fulfill my responsibilities, I sometimes feel like it's not entirely under my control.	1 2 3 4 5
General Self-efficacy	1. I can consistently find solutions to challenging problems with sufficient effort. 2. When faced with opposition, I can discover the necessary means to achieve my objectives. 3. Staying committed to my goals and accomplishing them comes naturally to me. 4. I have confidence in my ability to handle unexpected events effectively. 5. My resourcefulness enables me to navigate unforeseen situations adeptly.	Not at all or almost imperceptibly A bit To a moderate extent Significantly To an extreme degree

Inventory	Example Questions	Scale
Rosenberg's self-esteem	1. In general, I feel satisfied with who I am. 2. Occasionally, I believe I am not good enough. 3. I recognize that I possess several positive qualities. 4. I am capable of performing as well as the majority of people. 5. I feel there isn't much to boast about in my life.	Completely agree Agree Disagree Completely disagree
Positive and Negative affect schedule	1. Curious 2. Troubled 3. Enthusiastic 4. Unsettled 5. Powerful	Not at all or almost imperceptibly A bit To a moderate extent Significantly To an extreme degree
Oxford Happiness	1. I am not especially satisfied with the way I am. (R) 2. I am profoundly intrigued by other people. 3. I find life to be highly rewarding. 4. I hold very affectionate feelings toward nearly everyone. 5. I seldom wake up feeling fully rested. (R)	Completely disagree Somewhat disagree Slightly disagree Slightly agree Somewhat agree Completely agree

An emotional intelligence (EI) questionnaire serves as a self-assessment tool. Emotional intelligence encompasses five essential traits: self-awareness, self-control, inspiration, empathy, and interpersonal abilities, as described in [145].

The Eysenck personality test, marked by a self-assessment [146], comprises 48 thoughtfully distributed questions, with 12 questions dedicated to each personality trait: neuroticism, extraversion, psychoticism, and a lying scale. Participants respond with simple binary choices (Yes or No) in response to each query. For every item with a binary value of 1 or 0 is assigned, leading

to a scoring range from 0 to 12, allowing a comprehensive evaluation of the individual's traits.

The primary purpose of the self-determination scale is to examine the conduct of individuals who possess self-determination. This scale is considered to represent a relatively persistent aspect of people's personalities, revealing two key elements: (1) heightened understanding of their emotions and self-perception, and (2) a feeling of mastery and authority over their behaviors.

The GSE measure is composed of ten items in a psychometric assessment aimed at evaluating optimistic self-beliefs when confronting life's different challenges.

The RSE scale, developed by Rosenberg, is utilized to measure an individual's self-esteem. To maintain consistency, the scores the negative test items are reversed during analysis, ensuring that both positive and negative aspects hold equal importance. The final test result typically falls within the range of 10 to 40. A score below 14 indicates low self-esteem and highlights the need for support.

There are several terms on the Affects, Both Positive and Negative Schedule that convey different emotions and sentiments. There are several terms on the Affects, Both Positive and Negative Schedule that convey different emotions and sentiments. The total is determined by adding the sum of the 10 positive variables and the 10 negative variables.

The Oxford Happiness Questionnaire was created by psychologists at Oxford University [147]. The reverse score is indicated in this questionnaire by the sign (R). For instance, if a student responds with the number "1," they should cross it out and write the number "6."

Each question's scale value is transformed into a polarity value. The polarity of the survey is shown in Table 3.2. These polarity values are used by

the sentiment analyzer to forecast the students' emotions during the offline phase. The positive and negative orientations of the answers determine the polarity values. Positive Likert scale responses are scored positively, while negative ones are scored negatively.

Table 3. 2 Polarity Score

Questionnaire	Scale	Standard Values	Proposed Polarity Values
Emotional intelligence	Not at all Rarely Sometimes Often Very Often	1 2 3 4 5	-2 -1 0 1 2
Eysenck personality	Yes No	1 0	1 -1
Self-determination scale	1 2 3 4 5	1 2 3 4 5	-2 -1 0 1 2
General Self-efficacy	Almost Imperceptibly or Not at All A Bit To a Moderate Extent Considerably Extremely	1 2 3 4 5	-2 -1 0 1 2
Rosenberg's self-esteem	Completely Agree Agree Disagree Completely Disagree	4 3 2 1	2 1 -1 -2
Positive and Negative affect schedule	Almost Imperceptibly or Not at All A Bit To a Moderate Extent Considerably Extremely	1 2 3 4 5	-2 -1 0 1 2
Oxford Happiness	Completely Disagree Somewhat Disagree Slightly Disagree Slightly Agree Somewhat Agree Completely Agree	1 2 3 4 5 6	-3 -2 -1 1 2 3

The sentiment analyzer uses the polarity value to anticipate the students' feelings during the offline phase. The determination of these polarity values depends on whether their responses have positive or negative orientations. Positive responses on the Likert scale receive a positive score, whereas negative responses are assigned a negative score.

3.4 Experimental Results

The efficiency of the suggested sentiment analysis is assessed in this section. The study involved surveying a total of 1,000 students, utilizing seven diverse types of questionnaires, which encompass EI ,EPQ,EH,GSE,PANAS and SDS. Based on a conventional and polarity-based evaluation, the gathered responses were examined. The acquired findings are then calculated and assessed using accuracy and MAE (mean absolute error).

Table 3.3 contains information about the questionnaire, along with a display of the computation for the formative outcome. For each questionnaire, the standard computation is carried out using the pertinent polarity value from Table 3.2, as indicated in Table 3.3. A negative polarity implies a low scale, a positive polarity indicates a high scale, and a result of zero indicates an intermediate scale. A negative score, for instance, suggests poor self-esteem whereas a good score, for Rosenberg's self-esteem, indicates high self-esteem.

Table 3. 3 Questionnaire details

Questionnaire	No of Questions	Scale	Score Calculations	Result	Result with Polarity
Emotional intelligence links an individual's cognitive processes with their emotional processes.	15	1- Not at all 2-Rarely 3-Sometimes 4-Often 5-very often	Aggregate all scale values for each item	Result<34 is Low Result is between 35 and 55 is Average Result > 56 is High	-ve Score is low 0 is Average +ve Score is high
Eysenck personality measures the personality domain	48	1-Yes 0-No	Sum all scale values of Psychoticism Extroversion and Neuroticism	The highest value of Psychoticism, Extroversion, and Neuroticism	The positive value of psychoticism, extroversion and neuroticism
The self-determination scale evaluates variations among individuals regarding the degree to what extent they incline to act with self-determination.	10	1 2 3 4 5	Obtain the sum of all scale values and then divide it by 10.	Result < 3 is Low Result > 3 is High	+ve Score is high -ve score is Low
General Self-efficacyis a self-report measure of self-efficacy	10	1-Very Slightly or Not at All 2-A little 3-Moderately 4-Quite a bit 5-Extremely	Compute the total	Result < 20 is Low Result > 20 is High	-ve Score is low +ve Score is high
Rosenberg's self-esteem measures global self-worth	10	4-Strongly agree 3-Agree 2-Disagree 1-Strongly disagree	Compute the sum	Result<14 is Low Result between 15 and 25 is	-ve Score is low 0 is normal +ve Score is

Questionnaire	No of Questions	Scale	Score Calculations	Result	Result with Polarity
				Normal Result>26 is High	high
The Positive Afferct and Negative Affect Schedule is a self-report measurement of emotional impact.	20	1-Almost Imperceptibly or Not at All 2-A Bit 3-To a Moderate Extent 4-Considerably 5- Extremely	Add up all the positive (PS) and negative (NS) affect scale values.	If PS > NS, the result is positive. If NS > PS, the result is negative. If PS = NS, the result is inconclusive.	Positive score is positive Negative score is negative
The Oxford Happiness questionnaire is utilized to forecast the individual's happiness score.	29	1- Completely Disagree 2-Somewhat Disagree 3-Slightly Disagree 4-Slightly Agree 5-Somewhat Agree 6-Completely Agree	Add up all the scale values and then divide the sum by the total number of questions.	Between 1 and 3 is Unhappy Between 3 and 4 is Moderately Happy Between 4 and 6 is Happy	-ve Score is Happy 0 is Moderately Happy +ve Score is Unhappy

Table 3.4 displays the standard emotional intelligence scores and the presented polarity-based outcomes. Additionally, it presents a comparison of all possible results. The observations suggest that the difference in outcome percentages between the established and suggested assessments is negligible. The emotional intelligence scale's scores are shown as low, moderate, and high in Figure 3.3.

Table 3. 4 EIQ Results

EIQ Results			
Standard		**Proposed**	
Category	**Count**	**Category**	**Count**
Low	10	Low Average High	10 0 0
Average	290	Low Average High	10 265 15
High	700	Low Average High	0 0 700

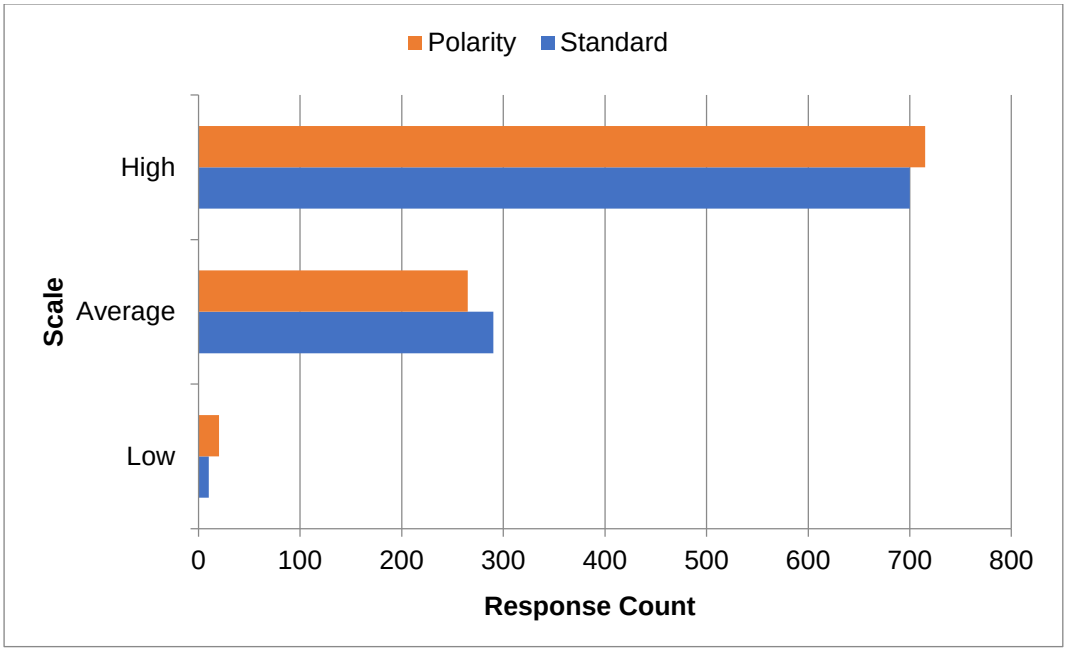

Figure 3. 3 EIQ (standard vs polarity).

The accuracy ratings for each response are included in Table 3.5 along with the Mean Absolute Error (MAE) and Emotional Intelligence (EI) scores. Again, fewer responses (between 200 and 400) result in less inaccuracy.

Table 3. 5 EIQ Results -MAE and Accuracy

Responses Count	MAE	Accuracy (%)
200	7	95
400	9	97
600	12	97
800	15	97
1000	17	98

The standard Eysenck personality scores and suggested polarity-based outcomes are shown in Table 3.6. The questionnaire scale values are depicted in Figure 3.4. For all scale values, the standard and polarity assessments produce identical findings.

Table 3. 6 EPI Results

EPI Results			
Standard		**Proposed**	
Category	**Count**	**Category**	**Count**
Psychoticism	35	Psychoticism Extroversion Neuroticism	35 0 0
Extroversion	665	Psychoticism Extroversion Neuroticism	0 665 0
Neuroticism	300	Psychoticism Extroversion Neuroticism	0 0 300

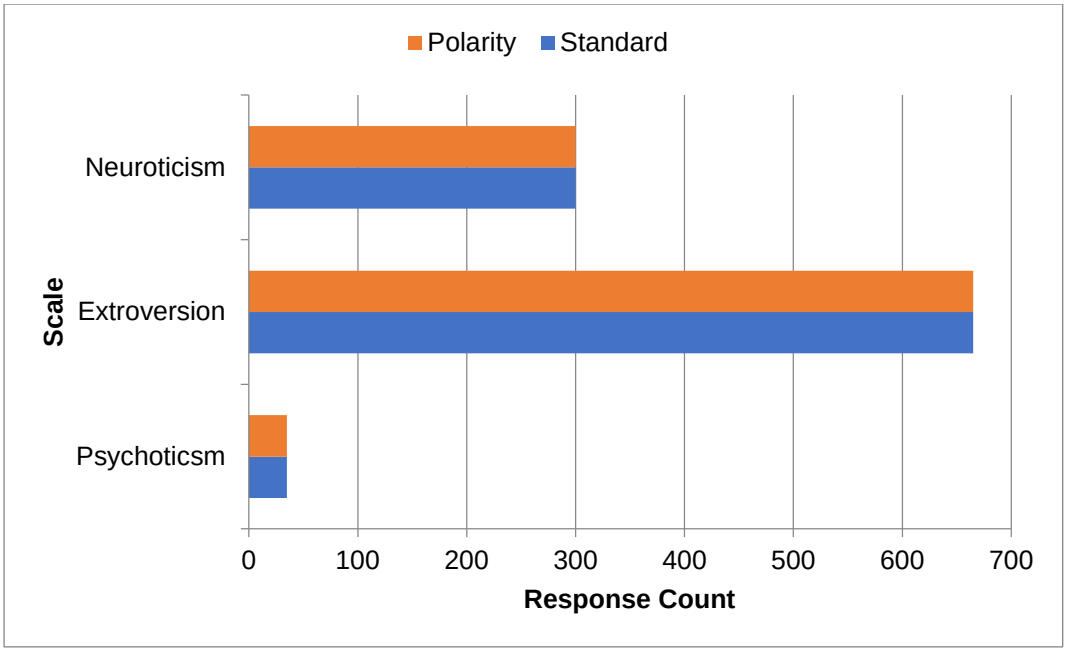

Figure 3. 4 EPI Scale (Standard vs Polarity).

Table 3.7 compares MAE and EPI accuracy for different amounts of replies. Once again, the data show no errors and 100% accuracy across a wide range of response numbers.

Table 3. 7 EPI Results -MAE and Accuracy

Responses Count	MAE	Accuracy (%)
200	0	100
400	0	100
600	0	100
800	0	100
1000	0	100

The conventional Self-Determination Scale (SDS) scores and suggested polarity-based outcomes are shown in Table 3.8. Figure 3.5 depicts the low and high SDS questionnaire scale results.

Table 3. 8 SDS Results

SDS Results			
Standard		**Proposed**	
Category	**Count**	**Category**	**Count**
Low	240	Low High	215 25
High	760	Low High	20 740

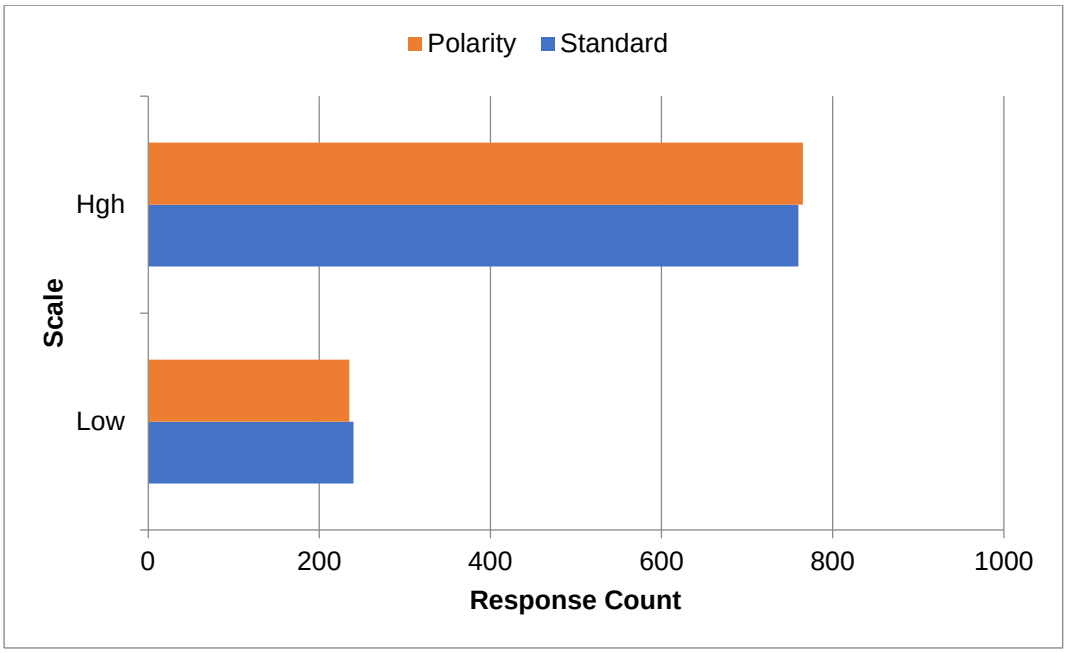

Figure 3.5: SDS (Standard vs. Polarity).

Table 3.9 compares the Mean Absolute Error (MAE) and accuracy for different numbers of responses for SDS.

Table 3. 9 SDS Results -MAE and Accuracy

Responses count	MAE	Accuracy (%)
200	5	96
400	10	96
600	10	95
800	5	96
1000	2	94

The standard General Self-Efficacy (GSE) scores and the suggested polarity-based outcome are shown in Table 3.10. The GSE questionnaire scale values are depicted in Figure 3.6 as low and high.

Table 3. 10 GSE Results

GSE Results			
Standard		**Proposed**	
Category	**Count**	**Category**	**Count**
Low	30	Low High	30 0
High	970	Low High	25 945

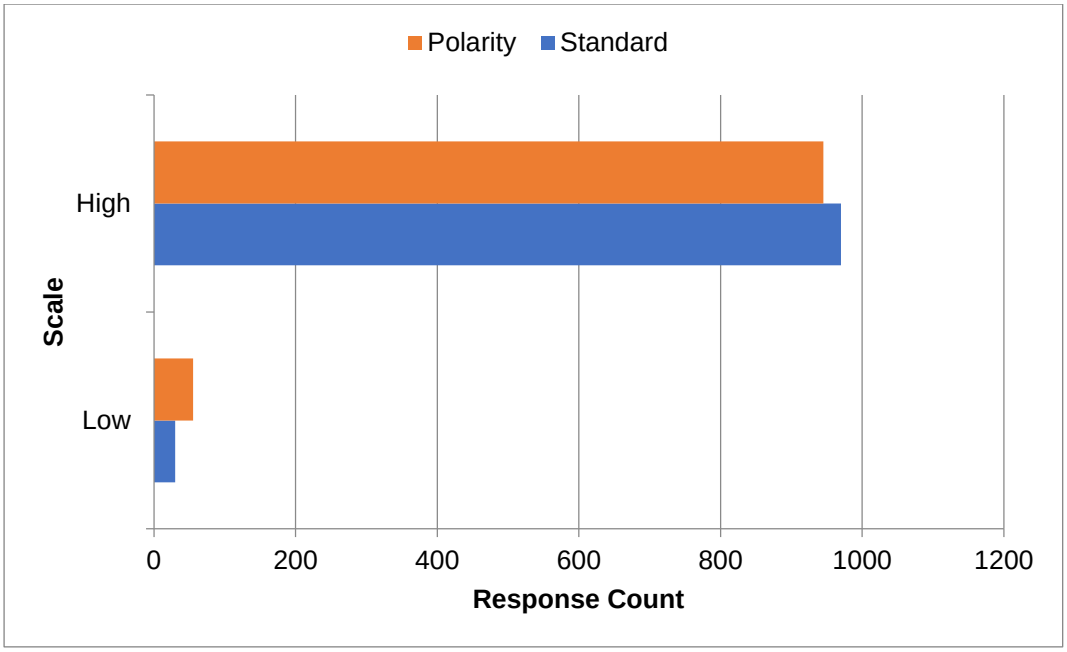

Figure 3. 6 GSE Scale (standard vs polarity).

Table 3.11 compares the General Self-Efficacy (GSE) accuracy and mean absolute error (MAE) for various response counts.

Table 3.11 GSE Results -MAE and Accuracy

Responses Count	MAE	Accuracy (%)
200	5	97
400	15	96
600	20	97
800	25	97
1000	25	98

The standard Rosenberg's self-esteem scores and the suggested polarity based outcome are shown in Table 3.12. The RSE questionnaire scale values are depicted in Figure 3.7: low, normal, and high.

Table 3. 12 RSE Results.

RSE Results			
Standard		**Proposed**	
Category	**Count**	**Category**	**Count**
Low	0	Low Normal High	0 0 0
Normal	145	Low Normal High	15 120 10
High	855	Low Normal High	15 20 820

Figure 3. 7 RSE Scale (standard vs polarity).

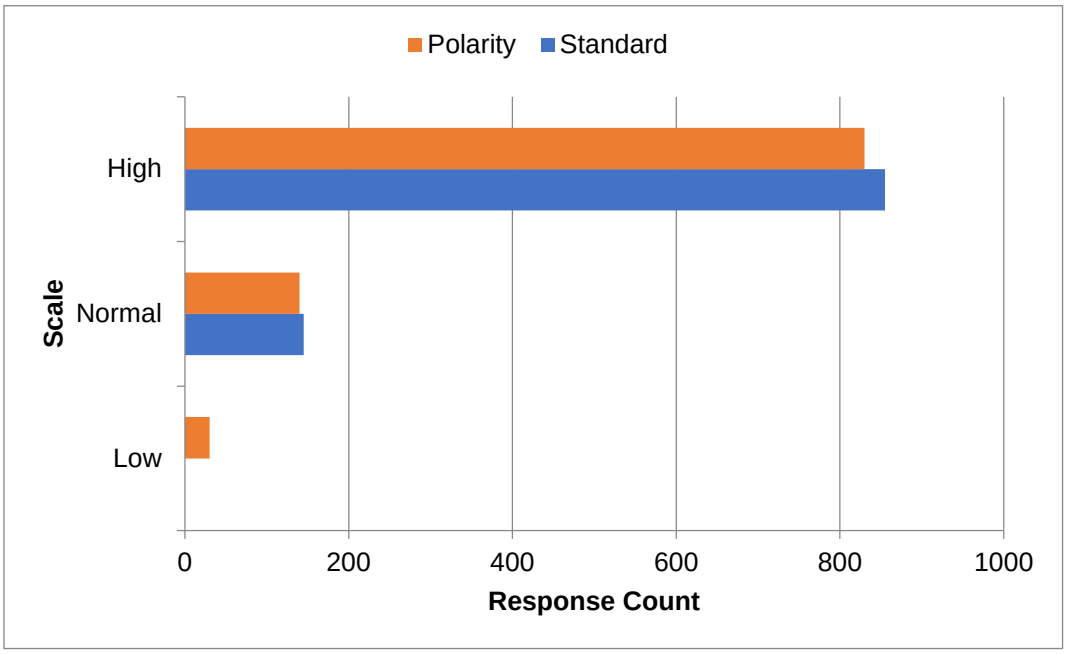

Table 3.13 compares the Mean Absolute Error (MAE) and accuracy of Rosenberg's self-esteem (RSE) for varied response numbers.

Table 3. 13 RSE Results -MAE and Accuracy.

Responses Count	MAE	Accuracy (%)
200	5	91
400	10	93
600	13	93
800	17	93
1000	20	94

Table 3.14 displays the standard Positive and Negative Affect Schedule (PANAS) ratings as well as the suggested polarity-based conclusion. For all scale values, the standard and polarity assessments produce identical results. The PANAS questionnaire scale values are depicted in Figure 3.8 as positive, negative, and neutral.

Table 3.14 PANAS results

PANAS Results			
Standard		**Proposed**	
Category	**Count**	**Category**	**Count**
Positive	805	Positive Negative Neutral	805 0 0
Negative	125	Positive Negative Neutral	0 125 0
Neutral	70	Positive Negative Neutral	0 0 70

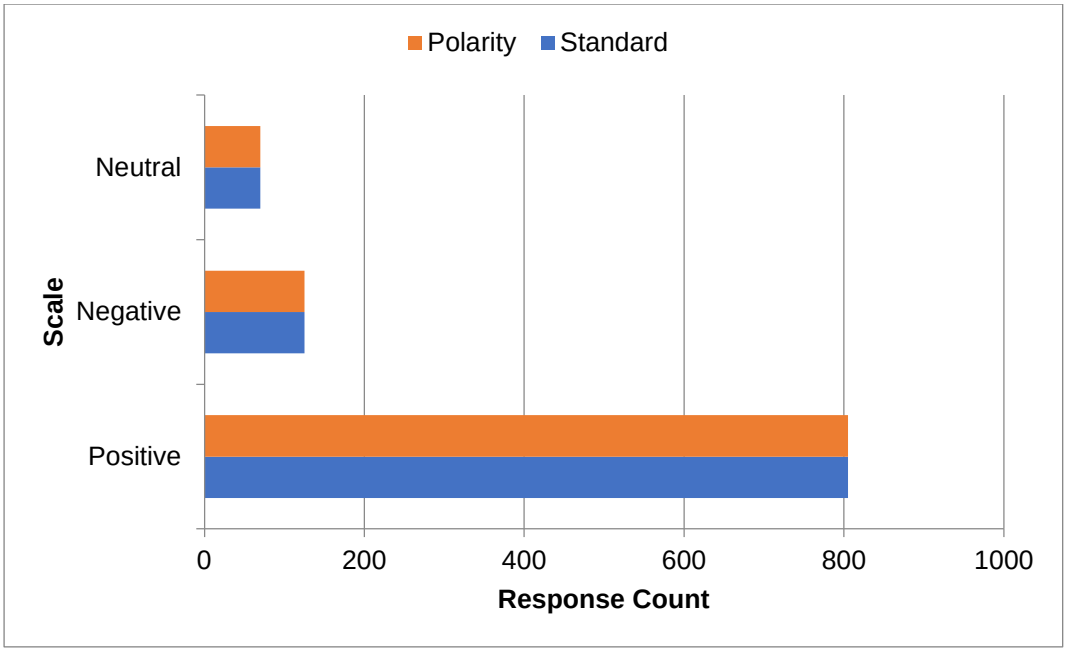

Figure 3. 8 PANAS Scale (standard vs polarity).

Table 3.15 compares the Mean Absolute Error (MAE) and accuracy of the Positive and Negative Affect Schedule (PANAS) for various answer counts. Once again, the data show no errors and 100% accuracy across a wide range of response numbers.

Table 3. 15 PANAS Results -MAE and Accuracy

Responses Count	MAE	Accuracy (%)
200	0	100
400	0	100
600	0	100
800	0	100
1000	0	100

Table 3.16 presents the standard and suggested polarity-based results of the Oxford Happiness Questionnaire. Figure 3.9 illustrates the scale values of the questionnaire, which include happy, moderately happy, and unhappy. Finally, Table 3.17 provides the MAE and accuracy of the Oxford Happiness Questionnaire.

Table 3. 16 OHQ Results

OHQ Results			
Standard		**Proposed**	
Category	**Count**	**Category**	**Count**
Happy	250	Happy Moderately Happy Unhappy	200 35 15
Moderately Happy	745	Happy Moderately Happy Unhappy	50 690 5
Unhappy	5	Happy Moderately Happy Unhappy	0 0 5

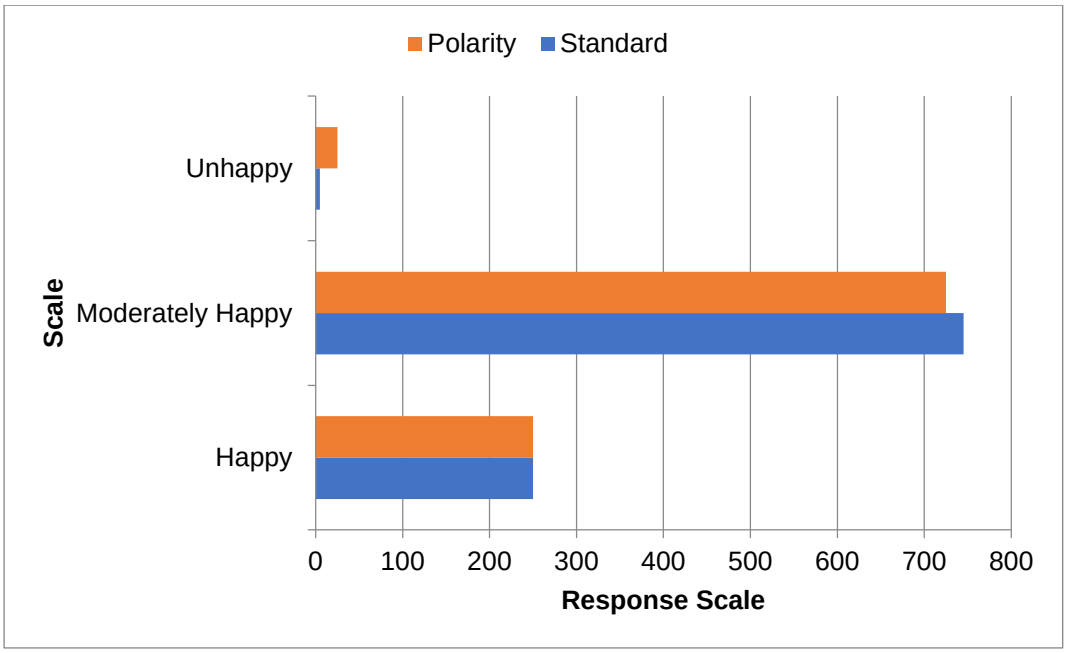

Figure 3. 9 OHQ Scale (standard vs polarity).

Table 3. 17 OHQ Results-MAE and Accuracy

Responses Count	MAE	Accuracy (%)
200	3	88
400	7	87
600	7	87
800	13	87
1000	13	89

3.5 Summary

This chapter described the sentiment analysis procedure for survey responses. The major objective was to develop a method for using closed-ended responses to analyse questions and students' emotions. The Quest SA tool for assessing students' emotions and questionnaire sentiments is described in this chapter. This study uses a closed-ended questionnaire to collect the responses from the students, and polarity-based sentiment analysis is used to determine the students' moods. Seven real-time questions are used to assess how well the work is being carried out: EIQ, EPI, OHQ, SDS, GSE, RSE, and PANAS. When compared to existing questionnaire evaluation methods, Quest_SA, which was suggested effectively anticipates students' emotional states with equivalent accuracy.