CHAPTER - 6

RECOMMENDER SYSTEM FOR PREDICITING STUDENTS' ACADEMIC PERFORMANCE

6.1 Introduction

Clients who have access to and are able to thoroughly evaluate user data will receive applicable recommendations from the recommendation system. [194]. As an illustration, information systems that can scrutinize prior behaviors and offer advice for ongoing issues are considered recommender systems. Additionally, in order to discover the most fitting scenario in accordance with user favorites, recommender systems endeavor to deduce their inclinations from comparable activities or that of other users who possess analogous attributes [195]. Numerous kinds of recommender systems exist, including:

The methodology of the content-based approach involves scrutinizing an object's substance and detecting trends among them [196]. Preserving insights of customers, offerings, and a particular customer's desires, a tactic known as knowledge-based [197] is employed. A method that relies on akin user favorites for constructing recommendations is Collaborative Filtering [198]. There are various challenges associated with collaborative filtering methods, including handling sparsity in datasets, improving scalability, and improving recommendation efficacy. Multiple approaches have been investigated to address the challenges having collaborative filtering as well as create top-notch suggestions.

Suggestions could be offered for commodities and amenities, as well as data like multimedia content, news, travel particulars, and even articles [199]. Recommender systems could be employed in instructive environments to steer learners via online learning pursuits through proposing written materials, books, and video content that facilitate their learning and performance improvement [200]. This chapter advocates student character with emotions.

6.2 Recommender System

Recommender systems offer tailored service sustain to customers through analyzing precedent activities and predicting their current preferences for specific products. Recommender systems leverage various information sources to forecast users' preferences for things of interest. Originally, recommender systems found application in e-commerce as a solution to the information overload resulting from Web 2.0. Later, their application was broadened to customise online commerce, virtual learning, and electronic travel. Recommender systems are a key component of modern Internet services like Netflix, Facebook, and YouTube. To assess the worthiness of an item for recommendation and gauge its usefulness, these systems are designed, in essence. [44].

The three main classifications of recommendation systems, extensively researched, are content-based, collaborative filtering (CF)-based, and knowledge-based. Each of the above categories has unique traits and appropriate application circumstances. Content-based recommendation systems [45] create user profiles based on past activity and suggest products matching those profiles. It is appropriate for promoting stuff like documents or webpages with a lot of content information. The assumption behind CF-based recommendation [46] algorithms is that users who are similar will favoursimilar goods. The Netflix competition has popularized its usage in diverse contexts. Knowledge-based recommendation [47] leverages precise data regarding users or products and operational guidelines to create consumer profiles and offer personalized suggestions. It is used in a few multifaceted cases when products like cars, houses, and insurance are not graded or are not commonly purchased. The combination of the three strategies above (hybrid model) [48] is frequently used in real-world recommender systems since each strategy benefits from the strengths of the others to make up for its shortcomings.

Recommender systems utilize past user records to predict an individual's attention in a product and deliver optimal recommendations tailored to that customer. These systems aid customers in making informed choices by providing personalized online data, as well as product and service recommendations. Originally developed to tackle the problem of information overload in e-commerce, recommender systems quickly extended their applications to encompass personalized solutions in digital government, online business, virtual learning, and electronic tourism.

The purpose of recommender systems is to assess whether recommending a product is worthwhile and to measure its usefulness. These systems are determined by a function that assesses the utility of a particular product for a consumer [201]:

$$f: U * P \rightarrow D$$

This characteristic depicts the usefulness of a specific thing (product) p ∈ P to a consumer u ∈ U. Multiple products are ranked in order to form the conclusive set of recommendations denoted as D. On this list, the usefulness of all items, which the customer has not used, is evaluated. The primary objective of most recommendation techniques is to locate products for the customer that maximizes the utility function, as presented in the formulation that follows:

$$\forall u \in U, \operatorname*{argmax}_{p \in P} f(u, p)$$

Subsequently, a summary of recommendations will be presented, comprising a choice of products arranged in ranked order. Essentially, a customer's scores of a product reflect its usefulness, denoted as ru,p, where user u rates product p. With M users and N products, the rating matrix is represented as $R \in R^{(M \times N)}$. The completion of the rating matrix or the assignment of a ranking score to each user-item pair are the two goals of making a suggestion. As a result, consumers' first-time viewings of products might be Sorted with the Top-K things being suggested. Different suggestion algorithms display a varying degree of success in predicting a product's utility for a particular buyer. Furthermore, several techniques have been explored to incorporate other data, for example product features, customer-created content, and social data.

6.3 Proposed Methodology

The suggested recommendation system for forecasting student character with feelings using closed-ended questionnaires is elaborated in this section. The two phases of the suggested recommendation method consist of an online and an offline process. Figure 6.1 illustrates the architecture of the suggested method. In the online process, learners are prompted to complete a closed-ended questionnaire.

The subsequent closed-ended questionnaires employed are: EIQ,GSE,RSE,SDS,EPI,EIQ,OHQ and PANAS. In Quest_DB, student replies are recorded. In the offline procedure, the examiner carefully examines the student responses and makes use of sentiment scores to predict the students' feelings. Student data is clustered using the reclust clustering technique. The ARM idea generates rules from the clustered data. The student's personality and emotional state are predicted using the recommendation system, which is the last step.

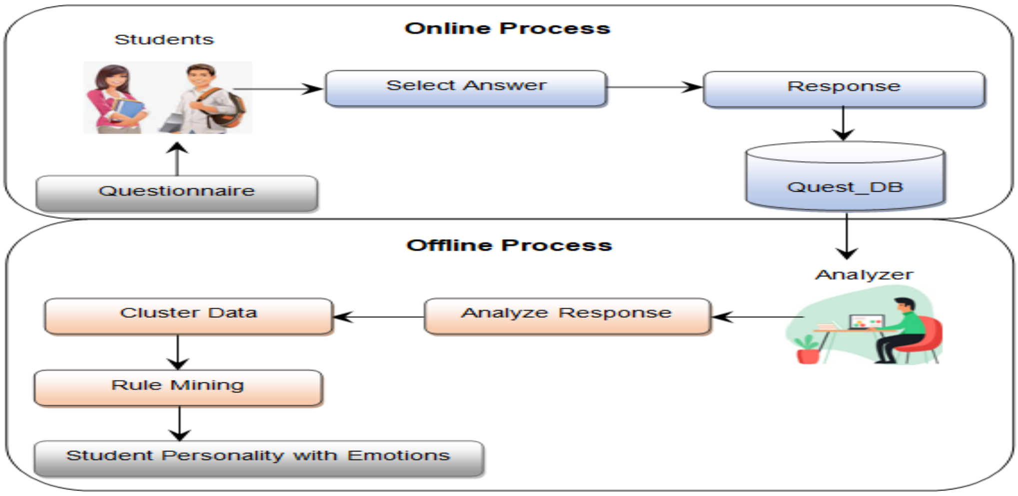


Figure 6. 1 System Architecture

Online Process

The field of emotion detection is garnering increased attention as a novel research area in Natural Language Processing (NLP) because of its capacity to take out opinions from openly accessible data on platforms similar to YouTube, Facebook, and more. Though, its applications extend beyond, encompassing various industries, for example education (for discovering learner disappointment) and healthcare (as a psychoanalytic tool) [144].

The online process employs SA based on polarity on a closed-ended questionnaire to identify students' affective traits. Subsequently, the chosen scale degree is transformed to a polarity degree for all questions. Table 6.1 displays the polarity values to the questionnaire scale.

Offline Process

The three sub-processes of the proposed offline process are forecasting of learner feelings, clustering of learner data, and suggestion utilizing ARM.

The examiner forecasts the learner's feelings during the offline phase, initially employing the polarity value. Then table 6.1 shows the questionnaire details with score calculation and results.

Table 6. 1 Questionnaire Details with Proposed score calculation

Questionnaire	No of Questions	Scale value and Scale	Proposed Polarity Values	Score Calculations	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	-2 -1 0 1 2	Aggregate all scale values for each item	-ve Score is low 0 is Average +ve Score is high
Eysenck personality measures the personality domain	48	1-Yes 0-No	1 -1	Sum all scale values 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	-2 -1 0 1 2	Obtain the sum of all scale values and then divide it by 10.	+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	-2 -1 0 1 2	Compute the total	-ve Score is low +ve Score is high
Rosenberg's self-esteem measures	10	4-Strongly agree 3-Agree 2-Disagree 1-Strongly	2 1 -1 -2	Compute the sum	-ve Score is low 0 is

global self-worth		disagree			normal +ve Score is 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	-2 -1 0 1 2	Add up all the positive (PS) and negative (NS) affect scale values.	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	-3 -2 -1 1 2 3	Add up all the scale values and then divide the sum by the total number of questions.	-ve Score is Happy 0 is Moderately Happy +ve Score is Unhappy

The student information undergoes clustering using a process called Reclust, which involves re-clustering and cluster validation. The clustering approach consists of three key steps: initial clustering, validation, and re-clustering. For the initial clustering, two modified clustering algorithms, KM and SOM, are utilized. The clustering procedure focuses on correcting inaccurately clustered data, while the re-clustering procedure is followed by the validation step to assess the clustering outcome. To enhance the overall excellence of the cluster outcomes, both the re-clustering and validation processes operate iteratively.

In the re-cluster algorithm, pair of clustering techniques, MPS-KM and SOM, are employed, and the cluster outcomes are assessed. The evaluate Cluster deploys class-based evaluation for clustering. It constructs clusters while disregarding the class. It assigns classes to clusters using the predominant class feature value in all clusters during the testing phase. Subsequently, the error of classification is computed using this task identifies the value with smallest error from pair of clustering techniques. It retrieves the data which clustered inaccurately from the assessment outcomes. Ultimately, it assesses the clustering outcome and clusters the erroneous data. For the re-clustering stage, the smallest amount of error value or the smallest amount of instances count in the data which is clustered incorrectly serves as the stopping condition.

Following clustering, ARM is implemented to produce a set of rules. Typically, support and confidence parameters are employed to determine the rules' significance. This study utilizes two support values to select the MIR and LIR to forecast . The final step involves forecast, where the equivalent class is forecasted using the created rules. It then examines the rule set's properties that relate to the expected class. The sample rules are as follows are given in figure 6.2

EIQ->High EPQ->Extroversion GSE->High OX->Happy PANAS->Positive RSE->High SDS->High Performance->Excellent
EIQ->High EPQ-> Neuroticism GSE->High OX->Happy PANAS->Positive RSE->High SDS->High Performance->Excellent
EIQ->High EPQ-> Neuroticism GSE->High OX->Moderately PANAS->Positive RSE->High SDS->High Performance->Excellent
EIQ->High EPQ-> Neuroticism GSE->High OX->Moderately PANAS->Positive RSE->Average SDS->High Performance->Excellent
EIQ->High EPQ->Extroversion GSE->High OX->Moderately PANAS->Positive RSE->High SDS->High Performance->Excellent
EIQ->High EPQ-> Neuroticism GSE->High OX->Happy PANAS->Positive RSE->Average SDS->High Performance->Excellent
EIQ->High EPQ->Extroversion GSE->High OX->Moderately PANAS->Positive RSE->High SDS->High Performance->Very Good
EIQ->High EPQ->Extroversion GSE->High OX->Happy PANAS->Positive RSE->High SDS->High Performance->Very Good
EIQ->High EPQ->Extroversion GSE->High OX->Happy PANAS-> Positive RSE->Average SDS->High Performance->Very Good
EIQ->High EPQ-> Neuroticism GSE->High OX->Happy PANAS->Positive RSE->High SDS->High Performance->Very Good
EIQ->High EPQ->Extroversion GSE->High OX->Moderately PANAS->Positive RSE->Average SDS->High Performance->Very Good
EIQ->High EPQ-> Neuroticism GSE->High OX->Happy PANAS->Positive RSE->Average SDS->High Performance->Very Good
EIQ->High EPQ-> Neuroticism GSE->High OX->Moderately PANAS->Positive RSE->High SDS High Performance->Very Good

e EIQ->High EPQ-> Neuroticism GSE->High OX->Moderately PANAS->Positive RSE->Average SDS->High Performance->Very Good

Figure 6. 2 Sample Rules mined

Derived from the created rules, guidance and suggestions are offered to the student to enhance their personality. For example,

If the student has high self-esteem and self-efficacy with a positive attitude, the student's emotion is 'Happy'.

The student has an extroverted personality if the student has high self-esteem, happy emotions with a positive attitude, and low self-determination.

6.4 Experimental Results

The efficiency of the presented experimentations is assessed in this section. The research employs seven types of questionnaires, namely GSE,RSE,SDS,EPI,EIQ,OHQ,PANAS are the seven types of questionnaires employed in the research. Responses from 1000 students were collected for analysis.

The collected responses were evaluated using both assessment methods based on standard and polarity. The results attained were then computed and assessed utilizing MAE (Mean Absolute Error) and Accuracy metrics, which are presented in Table 6.2.

Table 6. 2 Polarity Result

Questionnaire	MAE	Accuracy
EIQ	18.25	98.4
EPI	0	100
GSE	27	98
OHQ	15.76	92.3
PANAS	0	100
RSE	20	95
SDS	5	96.4

Figures 6.3 and 6.4 illustrate the comparison of MAE and accuracy in sentiment analysis across various questionnaires.

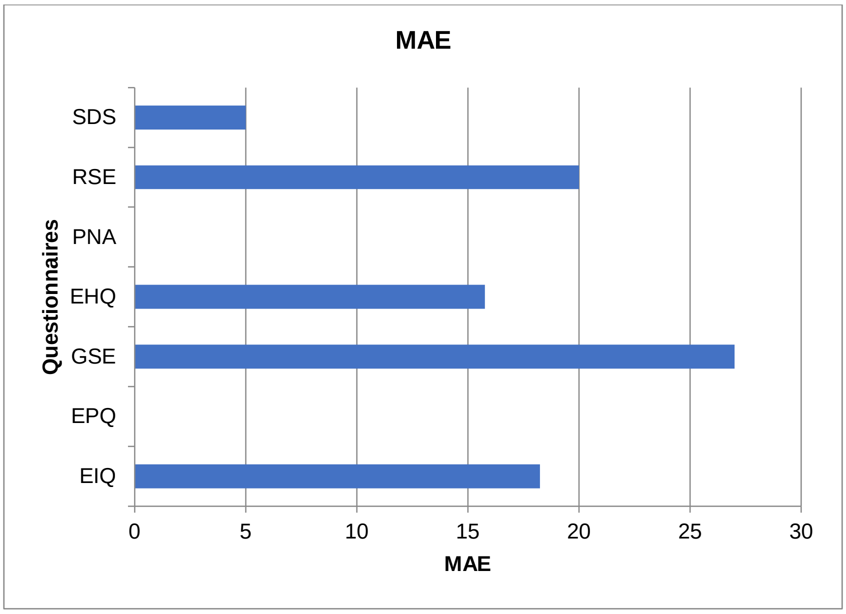


Figure 6. 3 MAE Comparison

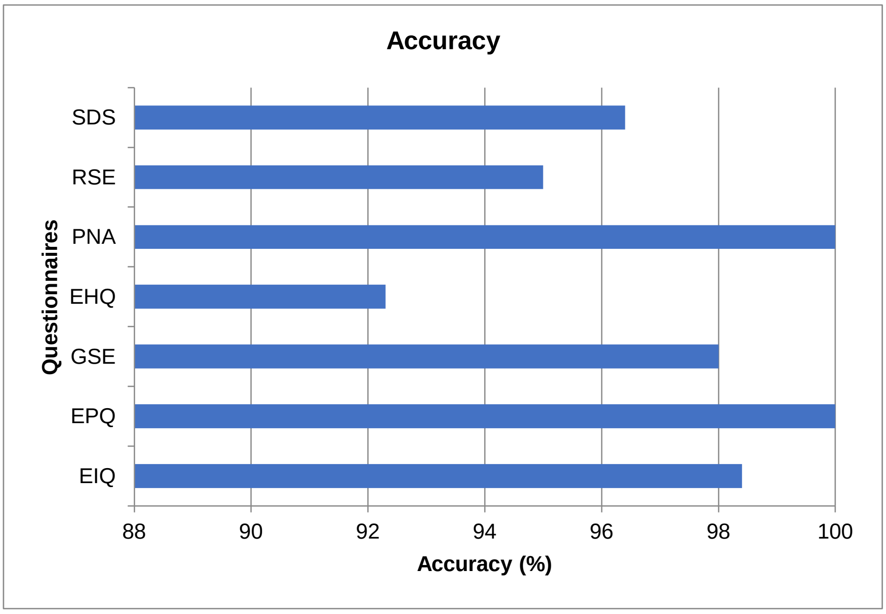


Figure 6. 4 Accuracy Comparison

Once the student's emotions are determined, the student data undergoes clustering using the Reclust algorithm. The clustering results are evaluated using several metrics, including Purity, Recall (Rec), Rand Index (RI), Normalized Mutual Information (NMI), and Precision (Pre). The assessment metrics comparison for clustering is presented in Table 6.3.

Table 6. 3 Comparing Assessment Metrics for Clustering

Data Set	Purity	RI	NMI	Pre	Rec
EIQ	0.98	0.92	0.85	0.86	0.89
EPQ	0.98	0.94	0.89	0.96	0.99
GSE	0.98	0.87	0.74	0.96	0.93
EHQ	0.98	0.94	0.88	0.95	0.99
PNA	0.96	0.86	0.78	0.96	0.93
RSE	0.97	0.86	0.75	0.91	0.91
SDS	0.97	0.89	0.79	0.96	0.97

Figure 6.5 displays the clustering assessment metrics for various kinds of questionnaires.

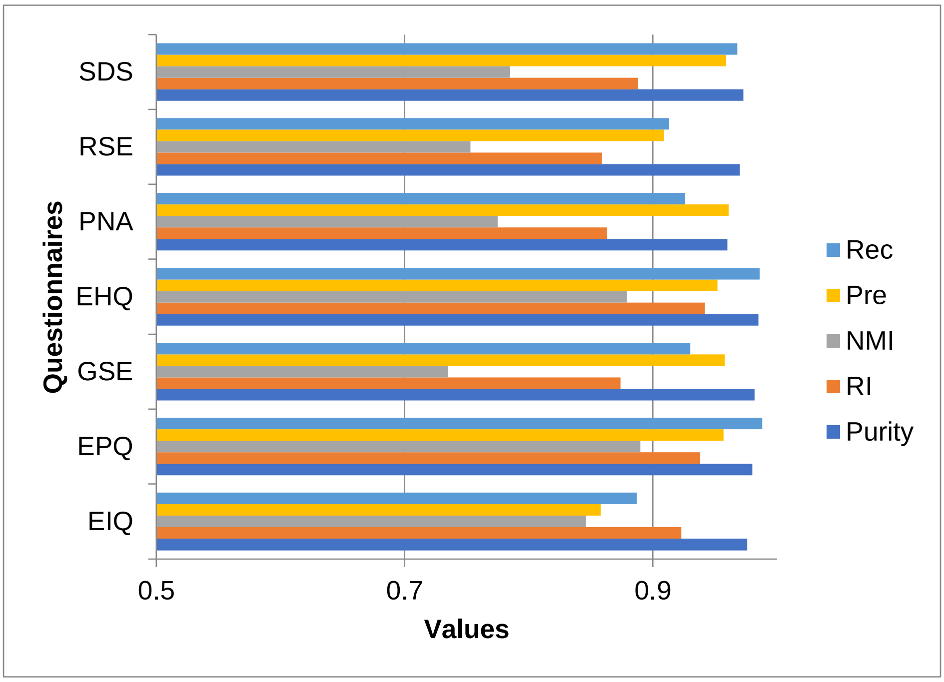


Figure 6. 5 Comparison of Assessment Metrics for Clustering

Table 6.4 presents the assessment criteria for Association Rule Mining (ARM).

Table 6. 4 Assessment Metrics for ARM

Dataset	Rules Count	Accuracy in %	Execution time in ms
EIQ	27	71	256
EPQ	7	68	133
GSE	27	79	121
EHQ	15	71	131
PNA	24	86	110
RSE	23	88.82	87
SDS	20	65.95	102

6.5 Implementation Based on the Experimental Results

The programming language Java (version 1.8) with Netbeans 8.2 is used for software implementation. On a machine having an Intel(R) Pentium processor running at 2.13 GHz and 4.0 GB of RAM, operating on a Windows 10 Operating System (64-bit), the tests are performed.Figure 6.6 shows the student home page. The student must give proper email id to enter the system.

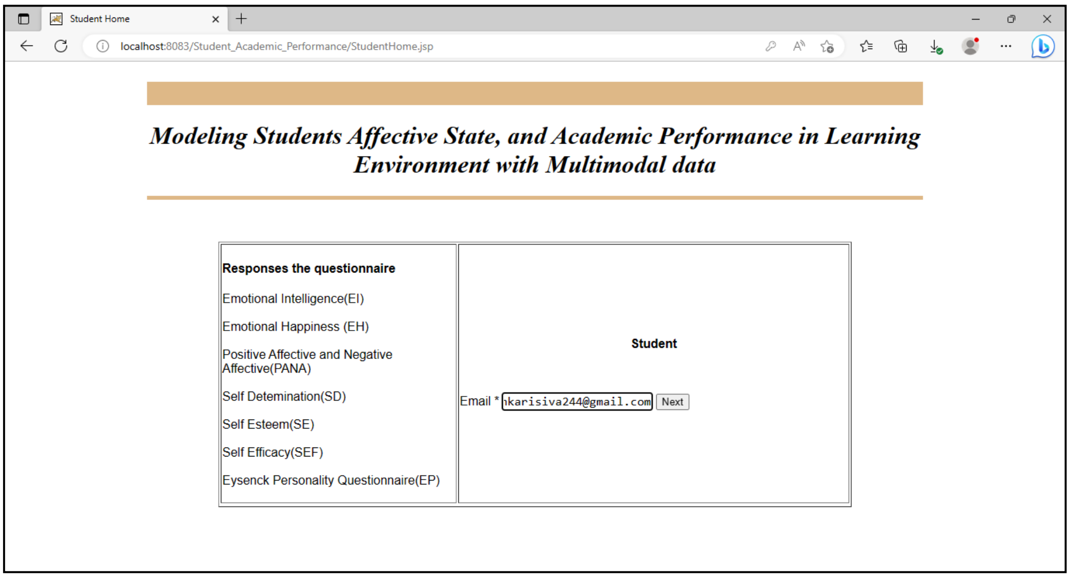


Figure 6. 6 Student Home Page

Student responses are collected through questionnaires. Figures 6.7, 6.8 6.9, 6.10, 6.11, 6.12 and 6.13 shows the questionnaire pages (EI, EP, SDS, GSE, RSE, PNA and EH).

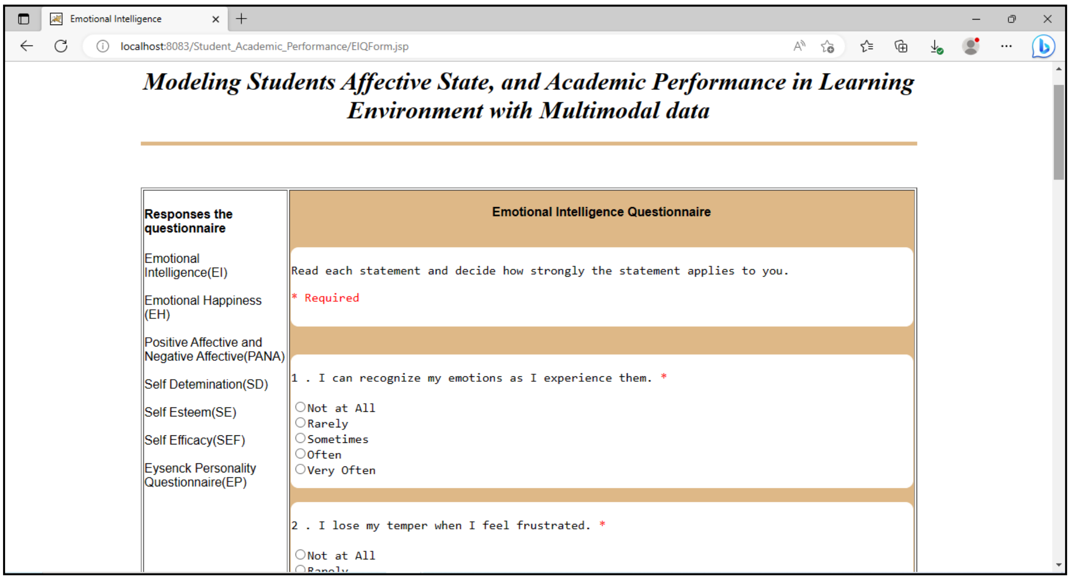


Figure 6. 7 EI Questionnaire Page

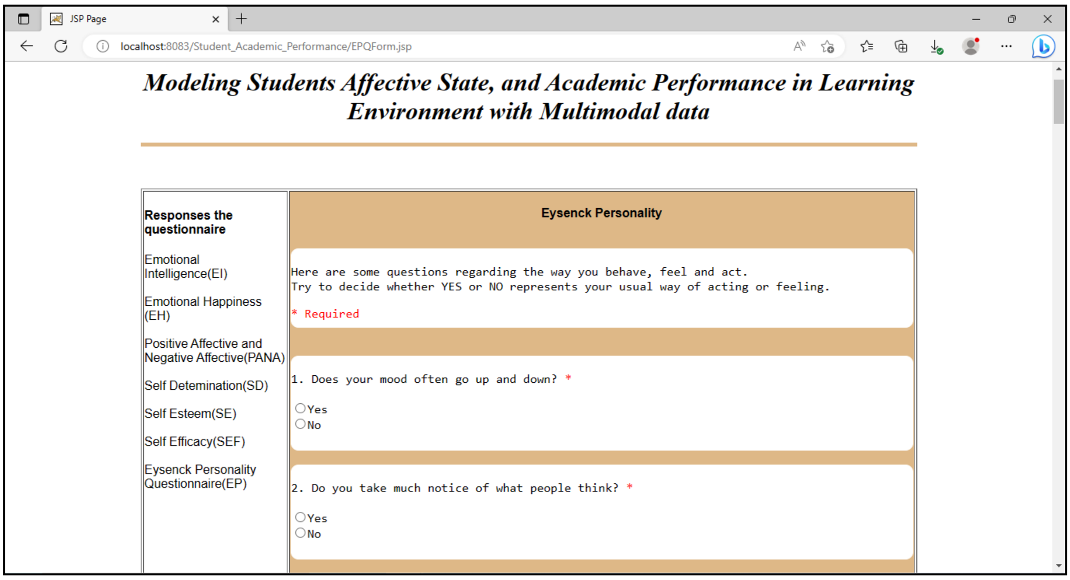


Figure 6. 8 EP Page

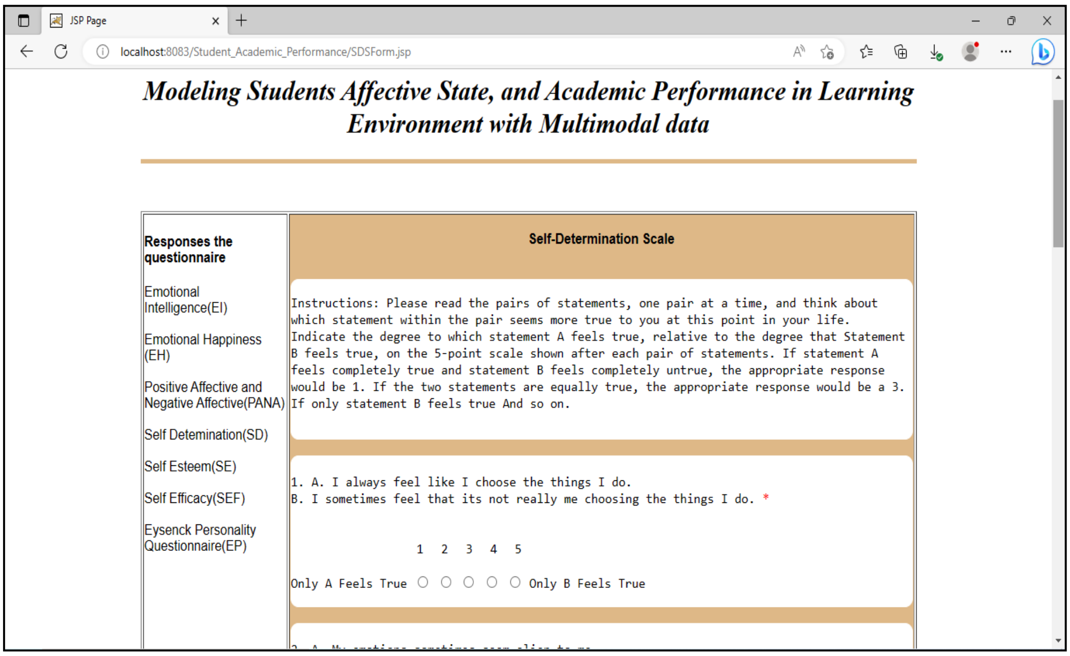


Figure 6. 9 SDS Questionnaire Page

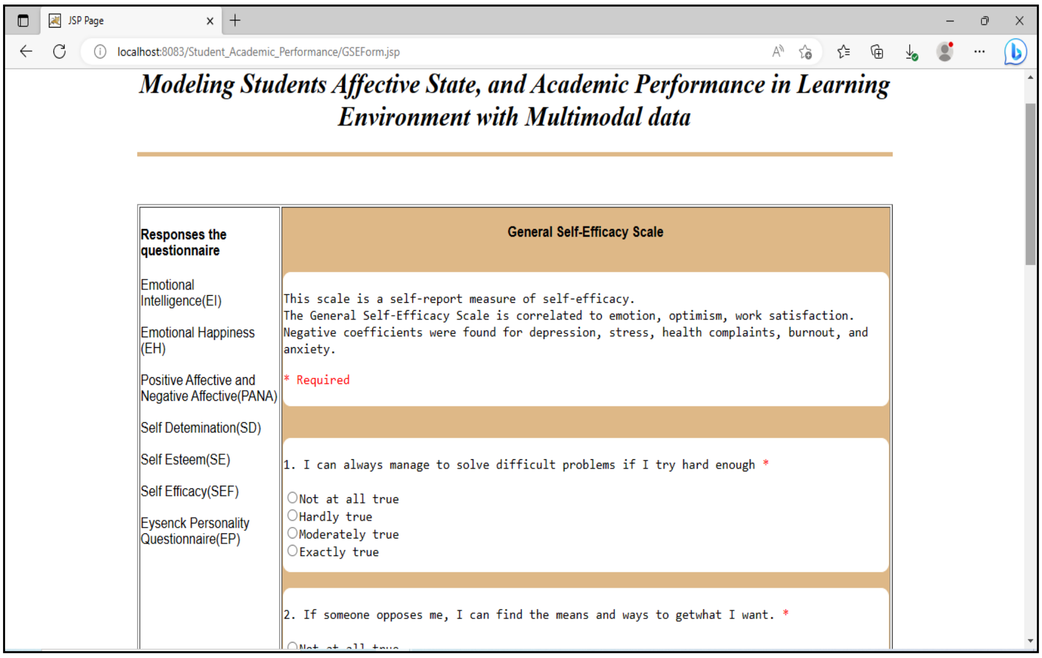


Figure 6. 10 GSE Questionnaire Page

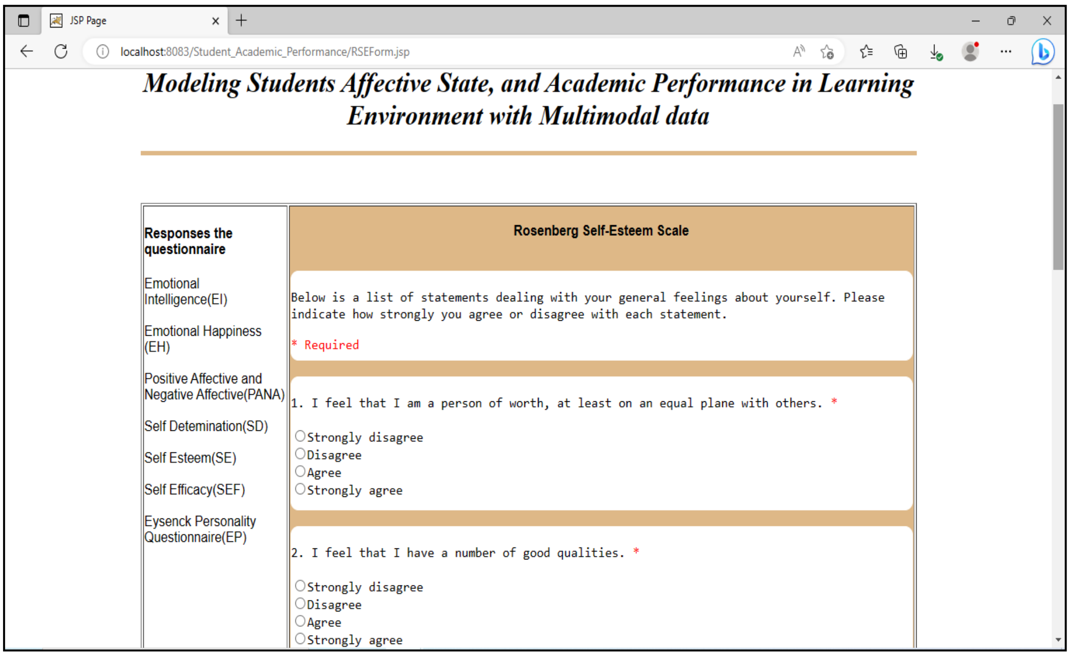


Figure 6. 11 RSE Questionnaire Page

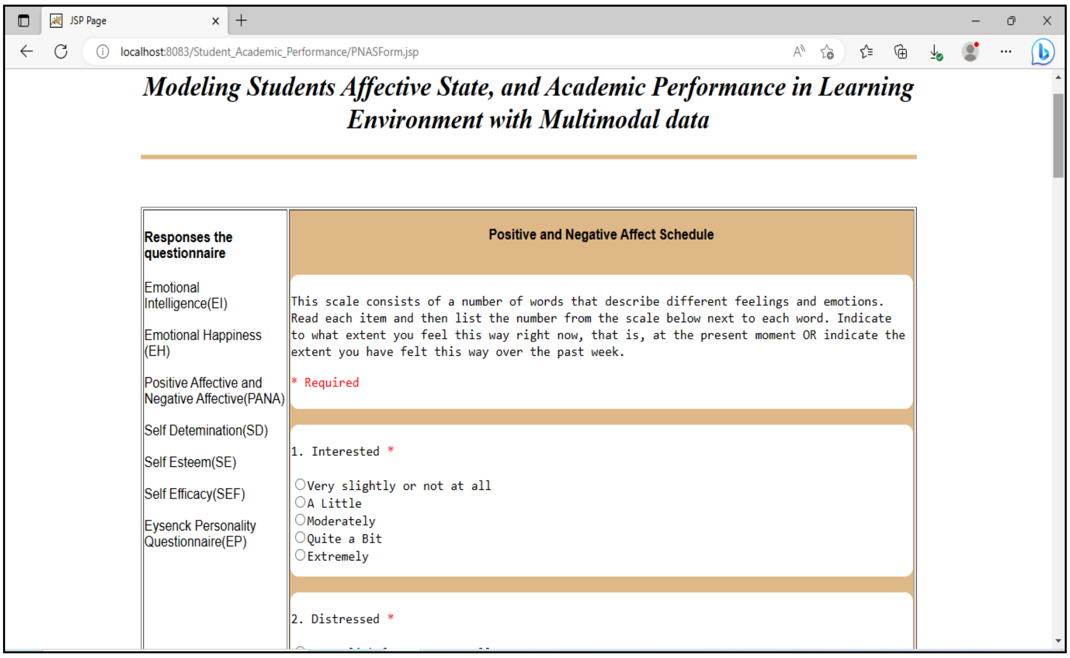


Figure 6. 12 PANA Questionnaire Page

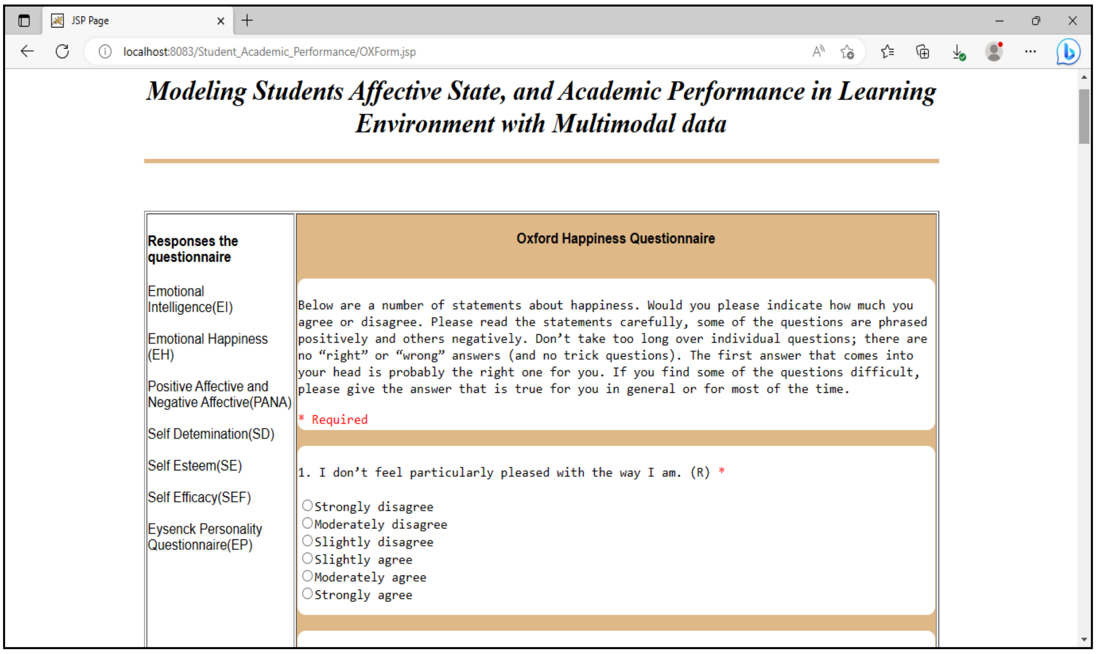


Figure 6. 13 EH Questionnaire Page

Figure 6.14 shows the clustering page. The system admin select the file (.csv format) for data clustering.

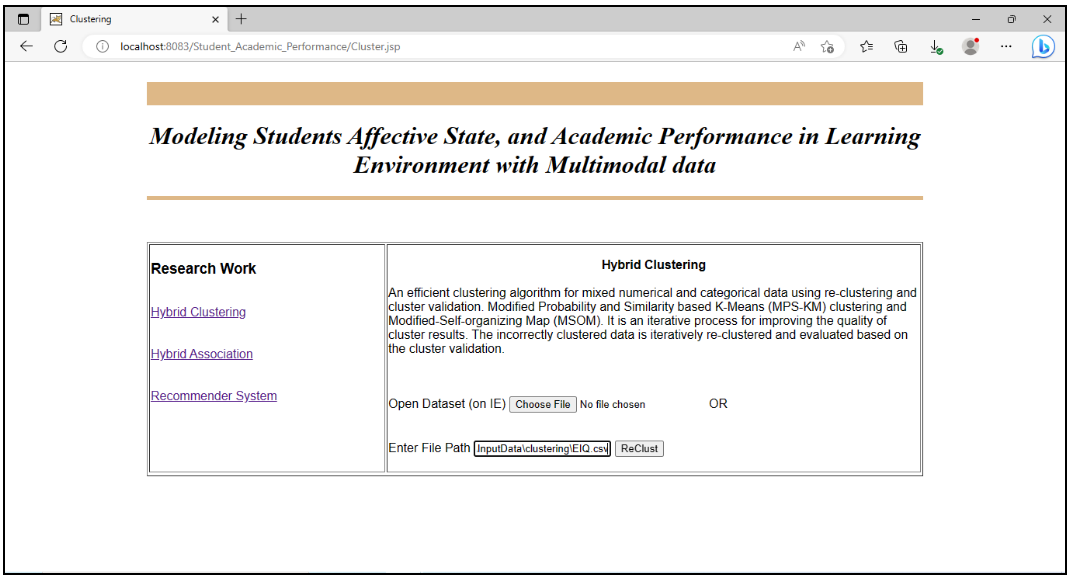


Figure 6. 14 Clustering Page

Figure 6.15 shows the cluster results. This page shows the classes to cluster assignment table with purity, rand index, NMI, Precision and recall metrics.

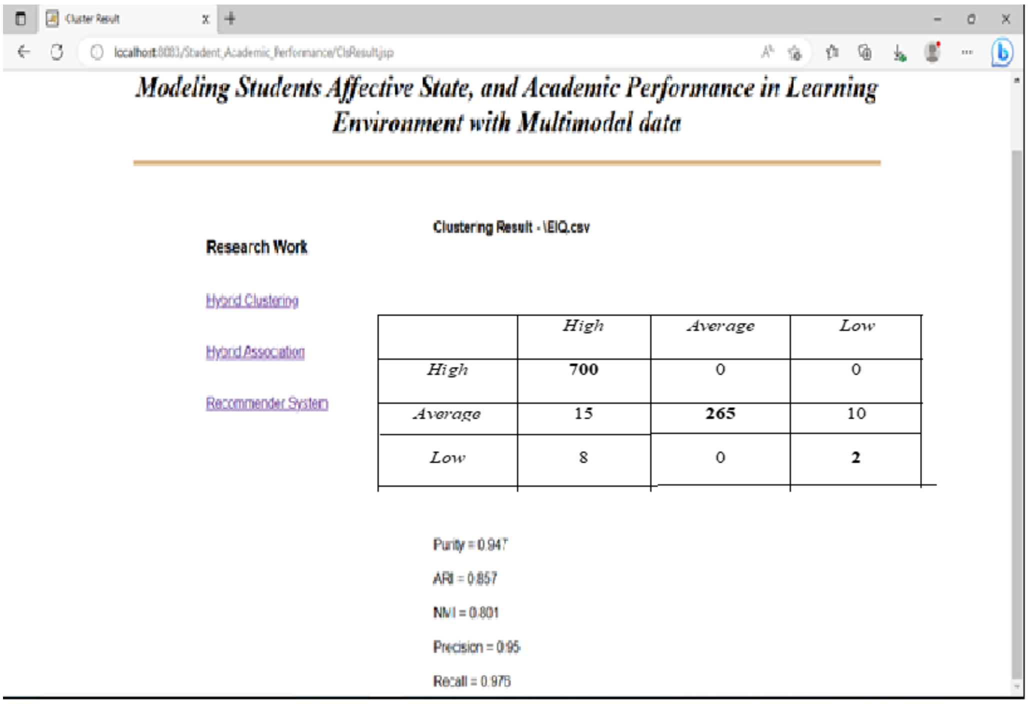


Figure 6. 15 Clustering Result Page

Figure 6.16 shows the association page.

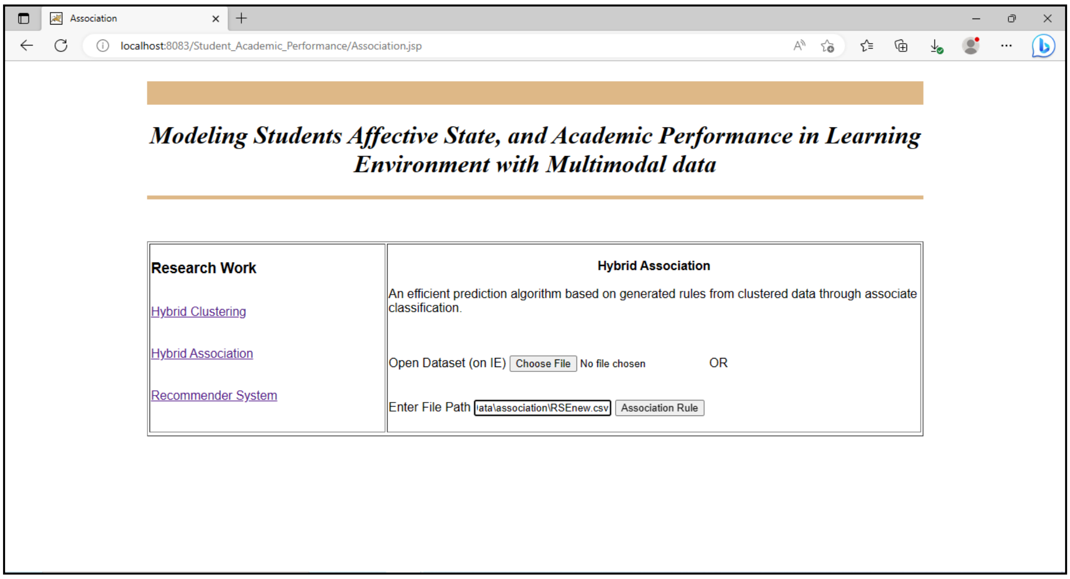


Figure 6. 16 Association Rule Mining Page

Figure 6.17 shows the association rule mining result. It display running time, number of generated rules and accuracy.

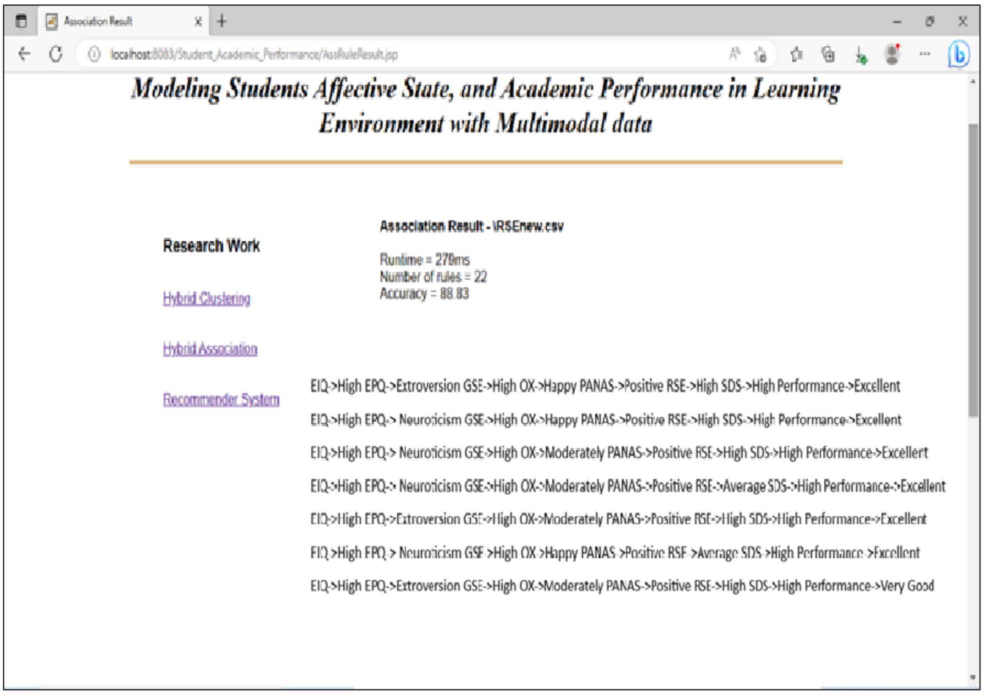
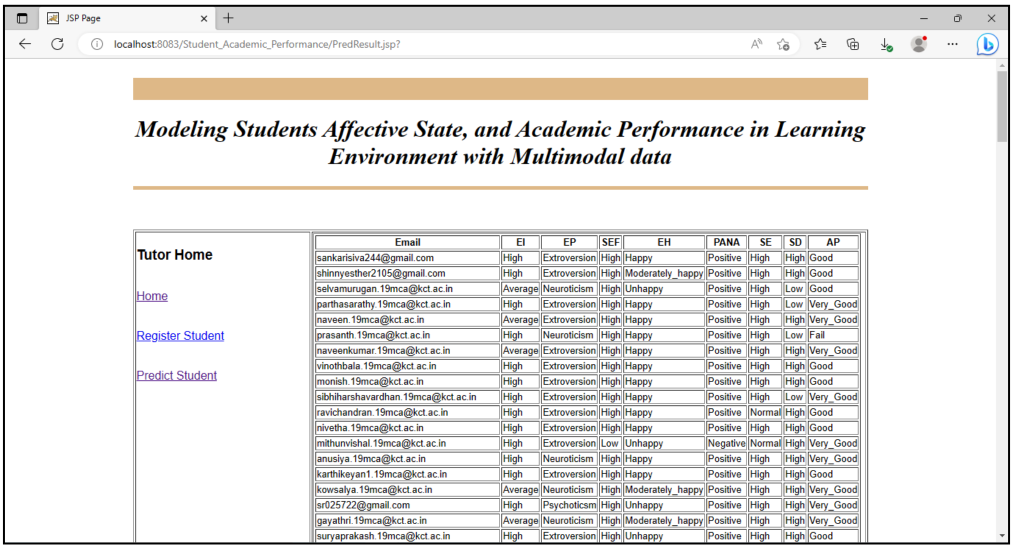


Figure 6. 17 ARM Result Page

Figure 6.18 shows the student affective state and academic performance. For each questionnaires (EIQ, OHQ, SDS etc), the system analyzer finds the corresponding sentiments through polarity assignment and the academic performance is predicted based on the association rules.



Email	EI	EP	SEF	EH	PANA	SE	SD	AP
sankarisiva244@gmail.com	High	Extroversion	High	Happy	Positive	High	High	Good
shinnyesther2105@gmail.com	High	Extroversion	High	Moderately_happy	Positive	High	High	Good
selvamurugan.19mca@kct.ac.in	Average	Neuroticism	High	Unhappy	Positive	High	Low	Good
parthasarathy.19mca@kct.ac.in	High	Extroversion	High	Happy	Positive	High	Low	Very_Good
naveen.19mca@kct.ac.in	Average	Extroversion	High	Happy	Positive	High	High	Very_Good
prasanth.19mca@kct.ac.in	High	Neuroticism	High	Happy	Positive	High	Low	Fail
naveenkumar.19mca@kct.ac.in	Average	Extroversion	High	Happy	Positive	High	High	Very_Good
vinothbala.19mca@kct.ac.in	High	Extroversion	High	Happy	Positive	High	High	Good
monish.19mca@kct.ac.in	High	Extroversion	High	Happy	Positive	High	High	Good
sibhiharshavardhan.19mca@kct.ac.in	High	Extroversion	High	Happy	Positive	High	Low	Very_Good
ravichandran.19mca@kct.ac.in	High	Extroversion	High	Happy	Positive	Normal	High	Good
nivetha.19mca@kct.ac.in	High	Extroversion	High	Happy	Positive	High	High	Good
mithunvishal.19mca@kct.ac.in	High	Extroversion	Low	Unhappy	Negative	Normal	High	Very_Good
anusiya.19mca@kct.ac.in	High	Neuroticism	High	Happy	Positive	High	High	Very_Good
karthikeyan1.19mca@kct.ac.in	High	Extroversion	High	Happy	Positive	High	High	Good
kowsalya.19mca@kct.ac.in	Average	Neuroticism	High	Moderately_happy	Positive	High	High	Very_Good
sr025722@gmail.com	High	Psychoticsm	High	Unhappy	Positive	High	High	Very_Good
gayathri.19mca@kct.ac.in	Average	Neuroticism	High	Moderately_happy	Positive	High	High	Very_Good
suryaprakash.19mca@kct.ac.in	High	Extroversion	High	Unhappy	Positive	High	High	Good

Figure 6. 18 Student Affective State and Academic Performance

6.6 Summary

In this chapter, a recommender system is introduced with the purpose of evaluating students' academic performance based on affective traits. The study collects student responses through a closed-ended questionnaire and employs SA based on polarity to identify their feelings. Subsequently, clustering and ARM techniques are employed to the student data. To assess the respective affective states, the research conducts experiments utilizing seven real-time questionnaires: PANAS,SDS,GSE,RSE,OHQ,EIQ,EPI. By mining the rules, the proposed research work predicts students' academic performance.