

ABSTRACT

Through a variety of learning experiences, education develops knowledge, competencies, outlooks, convictions, and habits. It gives people a way to hone their critical thinking, problem-solving skills, efficient communication, and all-encompassing comprehension of their surroundings. Bloom's taxonomy has had a big impact on education, emphasizing how important it is to help pupils develop higher-order cognitive abilities. Bloom's idea of educational action holds that, a person's cognitive and affective domains have an impact on their capacity to learn. Education data mining is one of the newest applications in the area, assisting in gaining a greater knowledge of the way students acquire knowledge and communicate in the processes aimed at raising the academic system's standard and financial sustainability over time.

The focus of the study is how students' affective state affects their cognitive state. By tying together characteristics from the mental abilities, emotional abilities and educational accomplishment, the study suggests a recommendation a technique for forecasting student achievement in academia. The suggested investigation has four steps, which are as follows: The First Phase entails the use of polarity scores in Quest-SA: A Sentiment Analysis-Based Evaluating Methodology for Closed-Ended Questionnaires. Re-clust: The second stage of the procedure involves re-clustering and cluster validation that utilizes clustering technique on data set that contains mixed data. A third stage is called ARMpred: An Effective Prediction Employing Associative Classifier Rule Construction for grouped Data. A recommender system is used in Phase 4 to forecast student academic performance.

Sentiment analysis has a important role in the field of natural language processing, finding applications in diverse sectors such as

education, news, politics, product reviews, and various industries. In the realm of education, sentiment analysis serves as a valuable tool for educators to gain quick insights into students' feelings regarding a course, enabling them to adapt their teaching strategies and elevate the overall excellence of education and instruction. Moreover, sentiment analysis also empowers students by identifying their feelings, academic progress, and engagement in various activities, thus fostering a more personalized and effective learning experience.

This study examines students' affective state for instance emotional intelligence (EI), emotional happiness (EH), personality, positive affect negative affect (PANA), self-esteem (SE), self-determination (SD), and general self-efficacy (GSE) through the analysis of closed-ended questionnaires. The proposed approach, named Quest_SA, utilizes sentiment analysis techniques to discern students' emotions from their responses to closed-ended questionnaires. Each questionnaire scale is allocated a polarity value, and subsequently, the students' replies are gathered through closed-ended questionnaires. These responses undergo sentiment analysis, wherein their emotions are categorized based on polarity using a sentiment analysis method. Finally, the emotion variance and sentiment scores derived from this analysis are utilized to evaluate and interpret the results.

Data mining relies heavily on clustering, which is a crucial process for identifying data clusters. Practical applications frequently involve a mix of numeric and categorical features to describe the data. To handle such datasets effectively, various clustering algorithms have been proposed that accommodate both categorical and numerical data. In evaluating the quality of clustering outcomes, cluster validation is employed within these algorithms.

In this study, a proficient clustering algorithm is introduced to handle datasets containing mixed data types (numbers and category). The approach incorporates repeated clustering and validation of clusters for improving accuracy. Initially, the mixed dataset undergoes clustering using two distinct algorithms: Modified Probability and Similarity based K-Means (MPS-KM) clustering and Modified Self organizing Map (MSOM). The above algorithms are then subjected to validation, and the superior one is selected to perform the clustering again. The re-clustering procedure aims to enhance the overall quality of the cluster outcomes by identifying and rectifying improperly clustered data points, guided by cluster validation assessments.

Data mining utilizes an associative classification (AC) model, which is rule-based; to construct precise classifiers through supervised learning datasets. This approach combines classification and association rule mining to achieve effective classification. Frequent rules are generated based on support and confidence measures. However, AC has some drawbacks, such as producing numerous rules for lower support, and some of these rules may not be used in the classification process. Additionally, it fails to classify minority classes effectively.

This study suggests (ARMpred), an effective prediction algorithm relying on rules generated from clustered data using AC. The approach involves three phases: Preprocessing, Rule Generation, and Prediction. In the initial phase, the gathered data undergoes preprocessing through discretization, converting numerical attributes into categorical. The second phase generates the rules set based on the cluster label. This method chooses More Important Rules (MIR) for prediction from the rules generated based on the two support values. The final phase involves making predictions for the corresponding class using the created rules.

Recommender systems (RS) produces recommendations that are important in the field of education, commercial and many other sectors. RS has numerous uses in a variety of industries, including writing, travel, cuisine, music, and movies. In the field of education, RS has been used to suggest student learning preferences. The recommender system for predicting students' affective states and academic achievement is suggested by this study. Collaborative filtering, one of the often employed RS methodologies, provides the most relevant recommendations by detecting similar persons or items based on previous interactions. However, the low accuracy of collaborative filtering is a significant problem. The suggested method addresses this issue by employing closed-ended questionnaires and association rule mining to predict pupil academic achievement accordingly to the chosen affective state for research.

To investigate the affective state encompassing Emotional Intelligence, Positive and Negative Emotions, Personality, Self-Determination, Self-Esteem, Self-Efficacy, and Happiness using the suggested sentiment analysis, closed-ended inventory in the real world such as the Eysenck Personality Questionnaire (EPQ), Self-Determination Scale (SDS), General Self Efficacy Scale (GSE), Rosenberg's Self Esteem Scale, Positive and Negative Affect Schedule (PANAS), and Oxford Happiness Questionnaire (OHQ).