Exploring The Integration Of Universal Human Values In Arts And Science Education: Perceptions, Predictive Analysis, And Pathways To Uhv-Stem Using Clustering

Dheeraj Kumar Singh[1*] and Narender Kumar[2]

[1,2]Department of Computer Science, Doon University, Dehradun, Uttarakhand-248001, India
dheerajsingh.phd@gmail.com[1]
Corresponding Author: Dheeraj Kumar Singh

ABSTRACT
One of the key educational trends of the modern era is the emphasis on Science, Technology, Engineering, Mathematics (STEM) education. STEM concepts are used in almost every educational institution, creating the impression that STEM education is promoted globally. At the one end STEM education model enhancing across the world, another side agitation among students is also enhancing. Therefore, this seems that STEM movement, only cannot, lead harmony in Human being. In this paper UHV-STEM model has been proposed and research has been validated using real data set collected by teaching Universal Human Values (UHV) to students and teachers of different discipline, and outcome has been analyzed by applying various clustering algorithms used for identifying pattern for course, subject, program selection by art and science students as well teachers. Two aspects of students' performance have been focused upon course selection specifically related to adoption of education model as UHV-STEM or UHV-STEAM. Combinations of the progression and cluster-based classification results are formulated using various identified clustering Algorithms.

Keywords: Universal Human Values, STEM, UHV-STEM, Clustering Algorithm, Educational Data Mining.

1. INTRODUCTION

Educational Data Mining (EDM) (Baker & Yacef, 2009) plays a crucial role in predicting students' performance, selecting suitable programs, subjects, and designing courses tailored to individual needs. One of the key challenges in education is predicting students' performance accurately to provide timely support and intervention. Clustering techniques, such as K-Means, DBSCAN, BIRCH, and others, can help in grouping students based on their academic performance, learning styles, and other relevant factors. By analysing these clusters, educators can identify patterns that indicate potential academic challenges or strengths, allowing for personalized interventions to improve students' performance.

STEM / STE(A)M (Science, Technology, Engineering, (Arts,) Mathematics) (Choi et al., 2017; Liao, 2016; Shin & Shim, 2021) course predictive analysis utilizes clustering algorithms to forecast demand and design programs aligning with student interests. Techniques like K-Means and Mini-Batch K-means can analyze better for course demand, ensuring institutions offer relevant and engaging STEM/STEAM programs that prepare students for future careers.

Course classification analysis techniques, including machine learning algorithms like K-Means and Mini-Batch K-means, can help in forecasting the demand for courses and designing new programs that align with students' interests and future career prospects. These techniques analyze historical data on course enrolments, student preferences, and market trends to make informed predictive analysis about future course demand.

Program selection techniques involve using clustering algorithms to group students based on their interests, aptitudes, and career goals. By analysing these clusters, educators can recommend suitable programs that align with students' aspirations and abilities. This personalized approach to program selection ensures that students are placed in programs where they are likely to excel and achieve their goals, leading to higher satisfaction and retention rates.

Similarly, subject selection techniques utilize clustering algorithms to categorize students based on their preferences, strengths, and weaknesses in different subjects. By analysing these clusters, educators can recommend subjects that are well-suited to each student's abilities and interests, fostering a more engaging and fulfilling learning experience.

STEM/STEAM education prediction focuses on forecasting the demand for STEM/STEAM courses and programs to meet the growing need for skilled professionals in these fields. By using clustering algorithms to analyze data on student interest in STEM/STEAM subjects, career aspirations, and market trends, educators can design STEM programs that are relevant, engaging, and aligned with students' future career prospects. This approach not only prepares students for successful careers in STEM but also contributes to the overall welfare of students by equipping them with the skills and knowledge needed for success to earn physical facility in the 21st-century (Cetto et al., 2000), but unable to handle anxiety, agitation creating among students (Gaur et al., 2010).

The inclusion of Universal Human Values (UHV) in education is essential for nurturing well-rounded individuals who are not only academically proficient but also ethical and empathetic (Singh & Kumar, 2024). By integrating UHV into STEM/STEAM courses, educators can instill values such as trust, gratitude, reverence, compassion, integrity, and respect for diversity, which are crucial for building a sustainable and inclusive society. This holistic approach to education ensures that students are not only prepared for academic and professional success but also equipped with the values and principles needed to be responsible global citizens. (Gaur et al., 2010) said that UHV can impart transformational change at the level of individual, family, Society, and Nature / Existence.

In conclusion, the integration of Universal Human Values with STEM/STEAM education (UHVSTEM / UHVSTEAM), supported by the application of clustering techniques in Educational Data Mining, holds immense potential for transforming the educational landscape. By leveraging these technologies and approaches, educators can analyses students' performance, design personalized courses and programs, and promote values essential for sustainable development. This holistic approach to education not only enhances academic outcomes but also fosters the holistic development and welfare of students, preparing them to meet the challenges of the future with confidence and compassion.

2. LITERATURE REVIEW

Clustering algorithms have become increasingly popular in the field of education, particularly in pedagogy and analyzing student performance. Algorithms like K-means, DBSCAN, BIRCH, and others provide a robust framework for analyzing educational data and extracting valuable insights. By grouping students based on their similarities and performance trends, educators and administrators can customize teaching approaches, forecast future outcomes, and enhance the overall learning experience.

One of the key uses of clustering in education is in course planning, curriculum design, and course predictive analysis (Regueras et al., 2019). By clustering students based on their academic strengths, weaknesses, and interests, educators can analyses personalized learning paths. Moreover, clustering algorithms can accurately analyse future academic performance, helping educators identify students who may need additional support. Demographic data, behavior, psychology, family socio-economic background, and school environment also play a great role in academic as well course selection (Batool et al., 2023; Issah et al., 2023; Kukkar et al., 2024).

3. FACTORS IMPACTING STUDENTS' PERFORMANCE

In addition to the studies mentioned earlier, several other parameters can affect student's achievements, course, subject, and program selection / prediction / analysis (Table 1) as mentioned below:

Table 1: Factors associated with students' achievements, course, subject, and program selection

Career Goals	Skill Compatibility	Peer Influence	Parental Influence	Financial Considerations	Program Reputation
Previous Academic Performance	Programme Fees	Attendance and Engagement	Learning Style and Support System	Interest and Motivation	Study Habits & Flexible Scheduling
Institute Location	Networking Opportunities	Extracurricular Activities	Industry Trends	Personal Interests and Passions	Work-Life Balance
Availability of Resources	Language of Instruction	Cultural Fit	Accessibility	Feedback and Reviews	Technology and Innovation
Institutional Support	Personal Development	Research Opportunities	Ethical and Social Impact	Accreditation	Student Diversity
Practical Application	Industry Alignment / Partnership	Networking and Alumni Connections	Cost and Financial Aid	Graduate Outcomes	Institutional Brand Value
Interdisciplinary Opportunities	Global Perspective	Flexibility and Customization	Personal Circumstances	Technology and Resources	Location of Graduates
Ethical and Moral Development	Character Development	Pass Percentage Ratio	Placement Analytics	Inter and Intra Country Relationships	Teachers Quality
Private and Government Body	Learning Outcomes	Alignment with Values	Admission Requirements	Post-Graduation Support	Future Opportunities

These above factors are in quantitative or qualitative in nature, and have been taken from the study (Bansal et al., 2023; Crisp et al., 2009; Gasiewski et al., 2012; Gaur et al., 2010; Williams & Williams, 2011).

Considering these above parameters alongside the previously mentioned factors can provide a more comprehensive, holistic understanding of students' course, subject, and program selection decisions.

4. PROPOSED WORK: ALGORITHMS TO BE USED

Pattern recognition is the primary application of data mining, commonly referred to as knowledge discovery in databases (Agrawal et al., 1993). The application of data mining techniques for pattern recognition, specifically in the field of education, is known as educational data mining, or EDM.

Patterns that can be discovered with the use of educational data mining techniques will revolutionize for student's success analysis and prediction (Romero & Ventura, 2007). Using educational data mining tools, it is possible to look at Vocational Training Programs and predict student course, subjects, program, and subsequently job mapping. By creating psychological or self-assessment batteries, demographic data about students—such as age, gender, and socioeconomic status—as well as a variety of other parameters—can be used to develop educational data mining techniques that predict and sustain collaborative governance and collaboration within and between organizations. Finally, EDM can predict offline and online education that is beneficial to society and the economy through colleges, universities, and other academic organizations. EDM is most commonly used to identify students, develop curriculum efficient test, forecast student performance, assist in determining courses, help learners avoid dropping out, increase course completion rates, and target learning.

Clustering can also be one of the fruitful unsupervised learning approaches for identifying pattern of students in blended learning (Asif et al., 2017; Durairaj & Vijitha, 2014; Peña-Ayala, 2014; Suciati et al., 2023) identify various clustering algorithm for making clusters for performing classification analysis. The various clustering algorithms are used for result formulation has been used are K-Means Clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies), Affinity Propagation, Mean-Shift, OPTICS (Ordering Points To Identify the Clustering Structure), Agglomerative Hierarchical Clustering, Mini-Batch K-Means (Ankerst et al., 1999; Khan et al., 2014; Peng et al., 2018; Sasirekha & Baby, 2013; Wang et al., 2018; Wu & Yang, 2007). Corresponding to all algorithms confusion matrix can be calculated for measuring the accuracy, performance and other related features. The confusion matrix results provide insights into how well clustering algorithms can group courses based on their attributes. Educators can use these results to improve course recommendations, curriculum design, and overall student learning experience.

5. RESEARCH METHODOLOGY

After reviewing of various studies has been identified that prediction of any new course, subject, and programme is not an easy task, even it is difficult to predict than that of grade prediction of any of the students. Definitely, it will be much more difficult when it has to be identified that what education have to be provide to the world, so that it can achieve fourth sustainable development goal of the United Nations. Some of the Organizations such as All India Council for Technical Education (AICTE), and University Grant Commission (UGC) and Ministry of Education Government of India has adopted the model of education of Universal Human Values. Where it has been identified that UHV education can better deal with acquiring Harmony at level of living of Human being such as Individual, Family, Society, Nature and Existence identified by (Gaur et al., 2010).

In this research, methodology used is qualitative, and firstly students have to taught UHV content for approximately of 18 hours, as well-run train to trainers' program for teachers for approx. 18 Hours, and, then Research Questions have been formulated and has been asked through Google form analytics on the base of Likert five-point scale "Strongly Agree", "Agree", "Undecided", "Disagree", "Strongly Disagree" (Armstrong, 1987). Finally various clustering algorithms has been applied as mentioned above in the paper, and confusion matrix formulated and course selection has been identified by art and science students as well teachers up to the level of severity for the inclusion of UHV with current education system. The research questions asked were as:

RQ1. Is It be sustainable and holistic education when incorporation of UHV done with current education?

RQ2. Is UHV based education will satisfy the holistic and sustainable living?

RQ3. Is UHV based education will lead students towards definite human conduct?

RQ4. Is UHV based education will lead students towards sustainable and holistic development?

RQ5. Is UHV based education will lead Harmony in Human being?

RQ6. IS UHV incorporation with current education will lead students towards Happiness?

RQ7. IS UHV important for all students / teachers / society for holistic living?

RQ8. IS only STEM (Science, Technology, Engineering, and Mathematics) education will lead Harmony in Human being?

RQ9. Is addition of UHV education with STEM education will lead Harmony in Human being?

RQ10. Is delivery of UHV "Education- Sanskar" is necessary for all students?

Where 'Education' is about understanding the right thing and 'Sanskar' is to live accordingly (Gaur et al., 2010). Predicting the research framework or educational framework for a course, subject, or program focused on Universal Human Values (UHV) based STEM (Science, Technology, Engineering, and Mathematics), STEAM (STEM + Arts), or NON-STEAM disciplines involves considering several key factors.

Integration of Values into Curriculum, Selection of Teaching Methods, Assessment of Value Integration, Professional Development for Educators / Facilitator / Co-Explorer, Engagement with Stakeholders, and to choose methods that promote the understanding and application of UHV in STEM / STEAM / NON-STEAM disciplines. This could include case studies related to human living, role-playing exercises, and reflective practices that encourage students to consider the ethical, moral, value-based dimensions of their work. This value-based collaborative approach will help to ensure the framework's relevance, effectiveness, and enhance complementarity (Gaur et al., 2010) (Singh & Kumar, 2024).

5.1 Research and Evaluation:

Overall, Predicting the research framework or educational framework for a course, subject, or program focused on Universal Human Values (UHV) based STEM (Science, Technology, Engineering, and Mathematics), STEAM (STEM + Arts), or NON-STEAM disciplines can be enhanced, analyzed, and predict by applying clustering algorithms.

5.2 Research Assessment Method:

5.2.1 Data Preparation:
Gather data related to core values, curriculum integration, teaching methods, assessment tools, professional development, stakeholder engagement, and research and evaluation. Each data point should represent a specific aspect of the framework (Feldman-Maggor et al., 2021).

5.2.2 Feature Selection:
Select relevant features from the data that are suitable for clustering (Parhizkar et al., 2023). These features should capture the key components of the framework and how they relate to UHV integration in STEM / STEAM / NON-STEAM disciplines.

5.2.3 Normalization:
Normalize the selected features to ensure that they are on the same scale and have a comparable impact on the clustering process or modeling (Parhizkar et al., 2023).

5.2.4 Clustering:
Apply clustering algorithms such as K-Means, DBSCAN, BIRCH, Affinity Propagation, Mean-Shift, OPTICS, Agglomerative Hierarchy, and Mini-Batch K-means to the normalized data. These algorithms will group similar data points together based on their feature values.

5.2.4.1 Cluster Analysis:
Analyze the clusters generated by the algorithms to identify patterns and groupings within the data. Look for clusters that represent different aspects of the framework, such as teaching methods that promote UHV integration or assessment tools that measure UHV understanding.

5.2.5 Framework Classification:
Based on the cluster analysis algorithms applied, analyze and doing cluster-based classification analysis and preparing the framework for the UHV-based STEM / STEAM / NON-STEAM course, subject, program, or education. Use the clusters to guide the development of the framework, ensuring that it integrates core values and promotes UHV in all aspects of teaching and learning.

5.2.6 Validation:
Validate the predicted framework by comparing it to existing frameworks and consulting with stakeholders. Make adjustments as necessary to ensure that the framework is relevant, effective, and aligned with UHV principles.

5.3 Performance Metrics:
In the methodology, final goal become possible to hit through taken confusion matrix and parameter associated with the matrix named as Accuracy, Precision, Recall, F1-Score, and Specificity (Xia, 2020), which are given as:

True Positive (TP): The number of correct predictions that an instance is positive.

True Negative (TN): The number of correct predictions that an instance is negative.

False Positive (FP): Incorrectly predicted as positive while it is negative.

False Negative (FN): Incorrectly predicted as negative while it is positive.

Accuracy = (TP + TN) / (TP + TN + FP + FN) (Formula 1)

Precision = TP / (TP + FP) (Formula 2)

Recall (Sensitivity) = TP / (TP + FN) (Formula 3)

F1 Score = 2 * (Precision * Recall) / (Precision + Recall) (Formula 4)

Specificity = TN / (TN + FP) (Formula 5)

6. RESULT ANALYSIS

In this section, we describe the results obtained in some specific phases of the Research Questions asked. For the evaluation of the dataset for further analysis various clustering techniques has been identified from Educational Data Mining, and results has been tested further. These clustering techniques are including K-Means, DBSCAN, BIRCH, Affinity Propagation, Mean-Shift, OPTICS, Agglomerative Hierarchy, and Mini-Batch K-means, to predict course UHV-STEM / UHVSTEM.

On behalf of confusion matrix precision, recall, specificity, sensitivity etc., has been measured corresponding to various clustering algorithms.

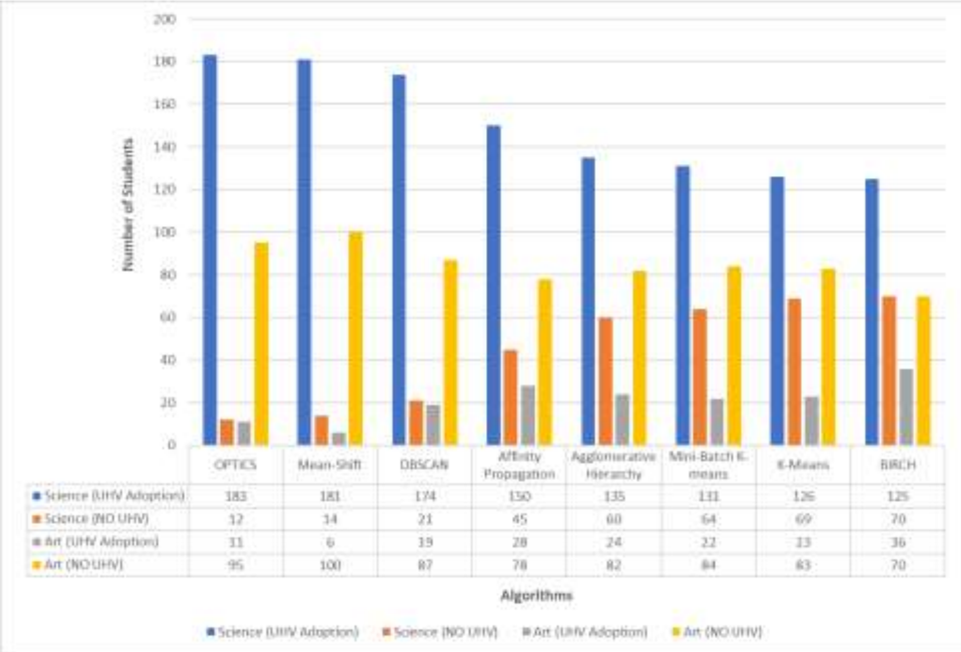

Figure 1: ML Model outcome using Various Clustering Feature Selection Algorithms

As per the confusion matrix identified (Figure 1) the rate of adoption of UHV is seen very less by students as well as faculties of art stream through all clustering algorithms, while the rate of adoption of UHV by the science students and faculties were very much high, and the rate of adoption of UHV by Science stream students and faculties were as 93.85% by OPTICS clustering algorithm, 92.82% by Mean-Shift Algorithm, 89.23% by DBSCAN, 76.92% by Affinity Propagation algorithm, 69.23% by Agglomerative Hierarchy algorithm, 67.18% by Mini-Batch K-Means algorithm, 64.62% by K-Means algorithm, and 64.10% by BIRCH clustering algorithm among all science students as well teachers. So, the highest predictive analysis rate for UHV has been identified by OPTICS clustering algorithm, and lowest predictive analysis rate has been identified by BIRCH clustering algorithm.

By deriving the confusion matrix generated the various classification analysis related parameter such as Precision, Recall, F1-Score, Specificity for the predictive analysis of the results corresponding to the clusters made has been further analyzed. Since in this research two cluster has been made for the art and science students on which it has been identified that what is the percentage ratio of arts as well science student’s vs teachers for occupying the UHV as a course.

Actual computation of accuracy can be understood through visualization done in Figure 2. It is clearly observable that OPTICS gives best result and then after DBSCAN, and further Mean-Shift algorithm can be the best for predictive analysis in terms of Accuracy, and K-Means gives lowest Accuracy.

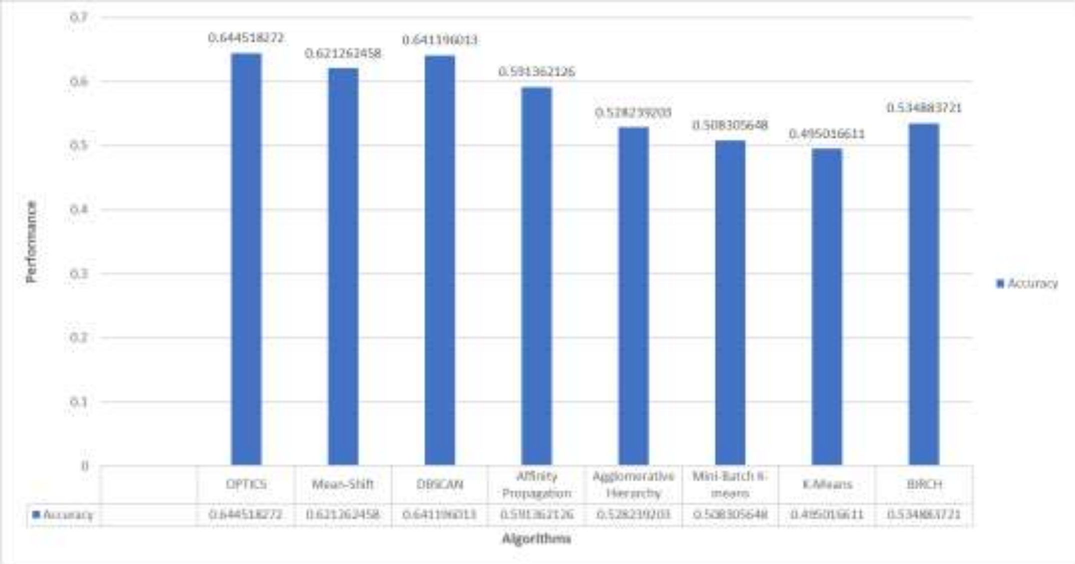

Figure 2: measuring level of Accuracy corresponding to various clustering Algorithms

Actual computation of precision can be understood through visualization done in Figure 3, and it is clearly observable that DBSCAN gives best result for precision and then after OPTICS, and further Affinity Propagation algorithm can be the best for predictive analysis in terms of Precision, and K-Means gives lowest precision.

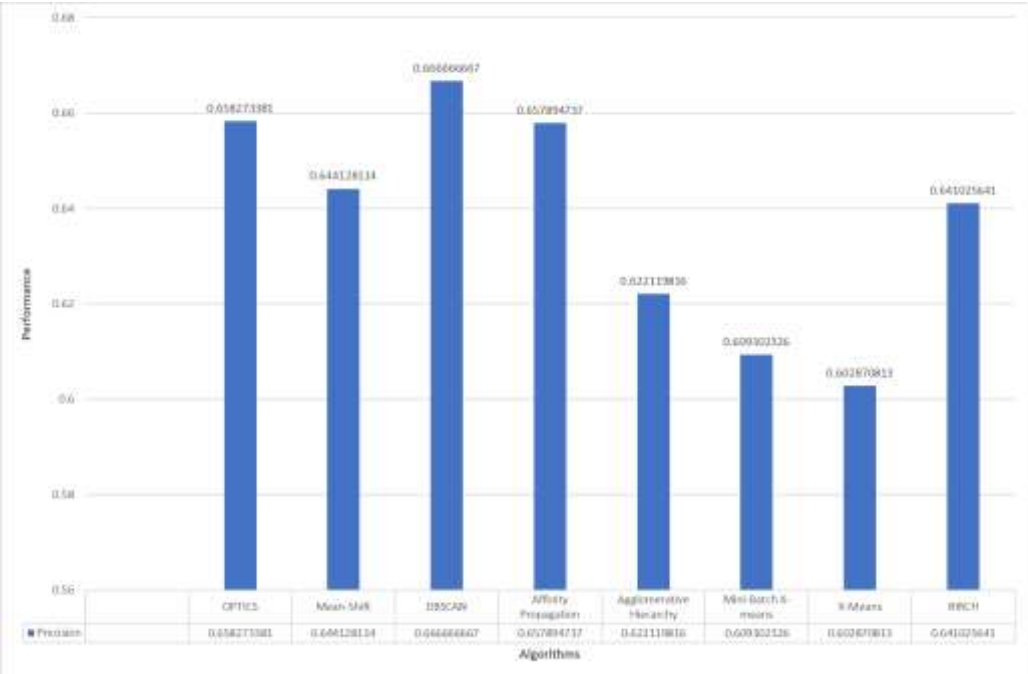

Figure 3: measuring level of precision corresponding to various clustering Algorithms

Actual computation of Recall can be understood through the visualization done in Figure 4, and it is clearly observable that OPTICS gives best result for recall and then after Mean-Shift, and further DBSCAN algorithm can be the best for predictive analysis in terms of Recall, and BIRCH gives lowest recall value.

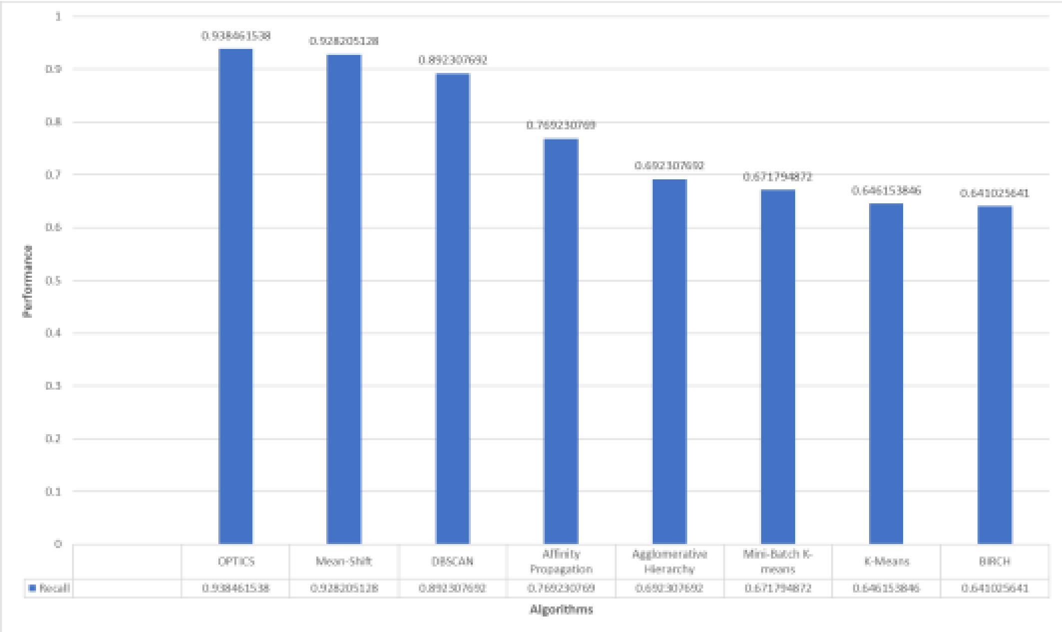

Figure 4: measuring level of Recall corresponding to various clustering Algorithms

Actual computation of F1-Score can be understood through the visualization done in Figure 5, and it is clearly observable that OPTICS gives best result for F1-Score and then after DBSCAN, and further Mean-Shift algorithm can be the best for predictive analysis in terms of F1-Score, and K-Means gives lowest F1-Score.

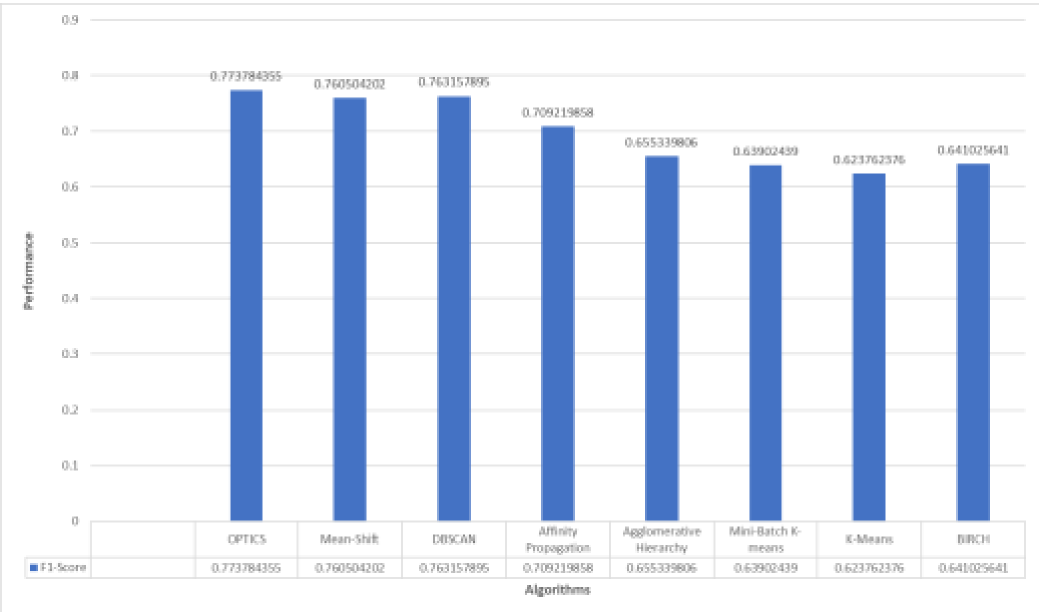

Figure 5: measuring level of F1-Score corresponding to various clustering Algorithms

Actual computation of Specificity can be understood through the visualization done in Figure 6, and it is clearly observable that BIRCH gives maximum specificity, and Mean-Shift clustering algorithm give minimum Specificity.

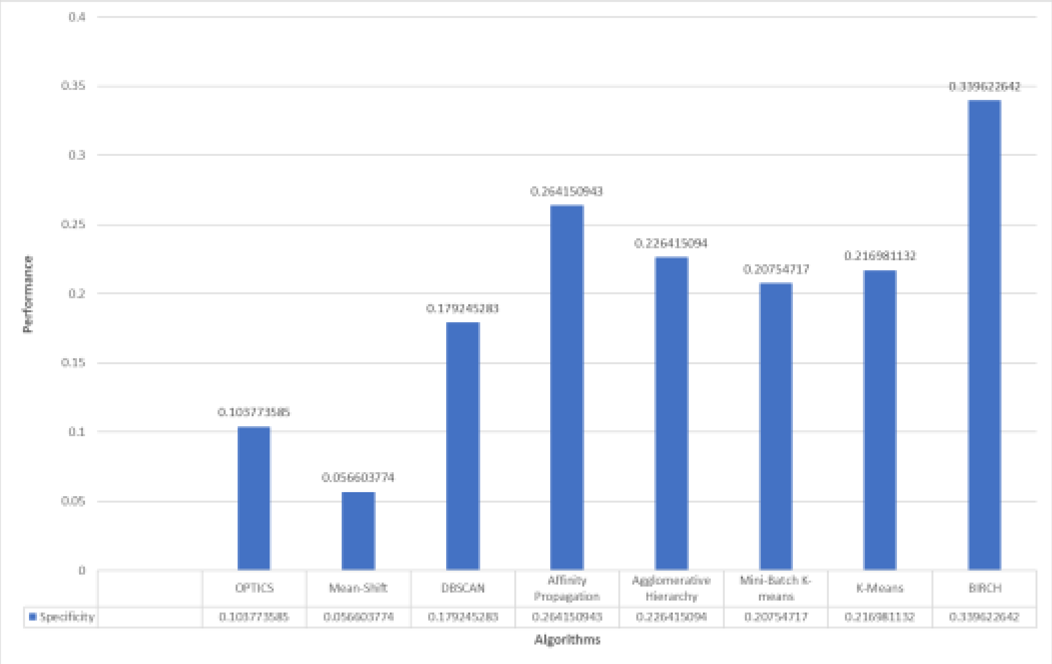

Figure 6: measuring level of Specificity corresponding to various clustering Algorithms

As shown in Table 2, all evaluation parameters, including recall value, accuracy, precision, and F1-score, and specificity related to cluster-based classification, may be evaluated at a single spot.

Table 2: Comparison table of evaluation metrics for various clustering algorithms based on student performance.

Algorithms	Accuracy	Precision	Recall	F1-Score	Specificity
OPTICS	0.644518272	0.658273381	0.938461538	0.773784355	0.103773585
Mean-Shift	0.621262458	0.644128114	0.928205128	0.760504202	0.056603774
DBSCAN	0.641196013	0.666666667	0.892307692	0.763157895	0.179245283
Affinity Propagation	0.591362126	0.657894737	0.769230769	0.709219858	0.264150943
Agglomerative Hierarchy	0.528239203	0.622119816	0.692307692	0.655339806	0.226415094
Mini-Batch K-means	0.508305648	0.609302326	0.671794872	0.63902439	0.20754717
K-Means	0.495016611	0.602870813	0.646153846	0.623762376	0.216981132
BIRCH	0.534883721	0.641025641	0.641025641	0.641025641	0.339622642

6.1 Significance of the study:

This study highlights the significance of using educational data mining approaches to include Universal Human Values (UHV) with STEM / STEAM education. This integration will ensure new personalized courses, and cultivating values important for sustainable growth. Ultimately UHV can be treated as a language for harmonious environment. Ultimately the results depict that education model should be value based STEM / STEAM (VbSTEM / VbSTEAM), and content of exploration should be UHV, so more precisely the model of Education should be UHVSTEM / UHVSTEAM.

7. CONCLUSION AND FUTURE WORK

The study aimed to analyze the inclusion of Universal Human Values (UHV) in arts and science education, focusing on students' and teachers' perceptions. It found that science students and teachers are more inclined towards UHV content, indicating a need for its compulsory study. While STEM education enhances skills, it lacks in fostering human relations, leading to a sense of disconnect later in life. Clustering algorithms applied to the dataset showed that UHV is crucial for understanding human relationships and promoting happiness. OPTICS, Mean-Shift, and DBSCAN found better for classification analysis of the course. To create societal harmony, the education model should be UHVSTEM / UHVSTEAM. Future research could predict student behavior using self-assessment batteries based on UHV content exploration among students and teachers, the impact assessment for students' holistic development, and comparative outcomes study can be possible to measure.

DECLARATIONS

Conflict of Interest: The researcher declares that they have no conflict of interest.

REFERENCES

1. Agrawal, R., Imieliński, T., & Swami, A. (1993, June). Mining association rules
2. between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD international conference on Management of data (pp. 207-216).
3. Ankerst, M., Breunig, M. M., Kriegel, H. P., & Sander, J. (1999). OPTICS: Ordering points to identify the clustering structure. ACM Sigmod record, 28(2), 49-60.
4. Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. Computers & education, 113, 177-194.
5. Armstrong, R. L. (1987). The midpoint on a five-point Likert-type scale. Perceptual and motor skills, 64(2), 359-362.
6. Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of educational data mining, 1(1), 3-17.
7. Bansal, A., Panchal, T., Jabeen, F., Mangla, S. K., & Singh, G. (2023). A study of human resource digital transformation (HRDT): A phenomenon of innovation capability led by digital and individual factors. Journal of Business Research, 157, 113611.
8. Batool, S., Rashid, J., Nisar, M. W., Kim, J., Kwon, H. Y., & Hussain, A. (2023). Educational data mining to predict students' academic performance: A survey study. Education and Information Technologies, 28(1), 905-971.
9. Cetto, A. M., Schneegans, S., & Moore, H. (2000). World Conference on Science: Science for the Twenty-first Century; a New Commitment, Budapest. p. 29-32
10. Choi, Y., Lim, Y., & Son, D. (2017). A semantic network analysis on the recognition of
11. STEAM by middle school students in South Korea. EURASIA Journal of Mathematics, Science and Technology Education, 13(10), 6457-6469.
12. Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. American Educational Research Journal, 46(4), 924-942.
13. Durairaj, M., & Vijitha, C. (2014). Educational data mining for prediction of student
14. performance using clustering algorithms. International Journal of Computer Science and Information Technologies, 5(4), 5987-5991.
15. Feldman-Maggor, Y., Barhoom, S., Blonder, R., & Tuvi-Arad, I. (2021). Behind the scenes of educational data mining. Education and Information Technologies, 26(2), 1455-1470.
16. Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. Research in higher education, 53, 229-261.
17. Gaur, R.R.,Sangal, R., & Bagaria, G.P.(2010). A Foundation course in Human Values and Professional Ethics. 2010, Excel Books Private Limited.
18. Issah, I., Appiah, O., Appiahene, P., & Inusah, F. (2023). A systematic review of the literature on machine learning application of determining the attributes influencing academic performance. Decision Analytics Journal, 100204.
19. Khan, K., Rehman, S. U., Aziz, K., Fong, S., & Sarasvady, S. (2014, February). DBSCAN: Past, present and future. In The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014) IEEE, 232-238.
20. Kukkar, A., Mohana, R., Sharma, A., & Nayyar, A. (2024). A novel methodology using RNN + LSTM + ML for predicting student's academic performance. Education and Information Technologies, 1-37.

21. Liao, C. (2016). From interdisciplinary to transdisciplinary: An arts-integrated approach to STEAM education. ART education, 69(6), 44-49.
22. Parhizkar, A., Tejeddin, G., & Khatibi, T. (2023). Student performance prediction using
23. datamining classification algorithms: Evaluating generalizability of models from geographical aspect. Education and Information Technologies, 28(11), 14167-14185.
24. Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. Expert systems with applications, 41(4), 1432-1462.
25. Peng, K., Leung, V. C., & Huang, Q. (2018). Clustering approach based on mini batch
26. kmeans for intrusion detection system over big data. IEEE access, 6, 11897-11906.
27. Regueras, L. M., Verdú, M. J., De Castro, J. P., & Verdu, E. (2019). Clustering analysis for automatic certification of LMS strategies in a university virtual campus. IEEE Access, 7, 137680-137690.
28. Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to
29. 2005. Expert systems with applications, 33(1), 135-146.
30. Sasirekha, K., & Baby, P. (2013). Agglomerative hierarchical clustering algorithm-a.
31. International Journal of Scientific and Research Publications, 83(3), 83.
32. Shin, D., & Shim, J. (2021). A systematic review on data mining for mathematics and science education. International Journal of Science and Mathematics Education, 19(4), 639-659.
33. Singh, D.K., & Kumar, N. (2024). Analyzing Students Performance through Holistic and Sustainable education approach in STEM through Educational Data Mining. In Tripathi, Durgesh, Sharma, Ramesh Kumar, & Tandon Surabhi (Eds.) Foundations of Change Analysing the Impact of National Education Policy, 2020 on India's Educational Landscape, London, United Kingdom: HP HAMILTON LIMITED, 145-164.
34. Singh, D. K., & Kumar, N. (2024). A systematic survey for students performance prediction with holistic and sustainable education approach using educational data mining. Tuijin Jishu/Journal of Propulsion Technology, 45(2), 232-269.
35. Suciati, N., Fabroyir, H., & Pardede, E. (2023). Educational Data Mining Clustering
36. Approach: Case Study of Undergraduate Student Thesis Topic. IEEE Access, 11, 130072-130088.
37. Xia, X. (2020). Clustering Analysis of Interactive Learning Activities Based on Improved BIRCH Algorithm. arXiv preprint arXiv:2010.03821.
38. Wang, L., Zheng, K., Tao, X., & Han, X. (2018). Affinity propagation clustering algorithm based on large-scale data-set. International Journal of Computers and Applications, 40(3), 1-6.
39. Williams, K. C., & Williams, C. C. (2011). Five key ingredients for improving student
40. motivation. Research in higher education journal, 12, 1.
41. Wu, K. L., & Yang, M. S. (2007). Mean shift-based clustering. Pattern Recognition, 40(11), 3035-3052.

BIOGRAPHIES OF AUTHORS

Dheeraj Kumar Singh is a Ph.D. Research scholar in the Department of Computer Science, School of Technology, Doon University, Dehradun. In addition to serving as an AICTE volunteer for UHV work, he co-facilitates, observes, and helps out at numerous UHV events in India. His MCA degree is from UPTU, Lucknow (U.P.), and his M.Tech. (Information technology) is from GGSIP University, Delhi. In addition to being a member of numerous academic and administrative committees, he has been in teaching from 2004. His expertise includes working with educational data mining, programming languages, machine learning, digital marketing, patents, trademarks, and copyrights. He can be contacted at email:

dheerajsingh.phd@gmail.com

Narender Kumar is Ph.D. from Indian Institute of Technology (I.I.T.), Roorkee. He is currently working as an Associate Professor in the Department of Computer Science, School of Technology, Doon University, Dehradun, since 2021. Before joining to Doon University, He has worked as Assistant Professor in the Department of Computer Science and Engineering, H.N.B Garhwal University, Srinagar Garhwal from 2012 to 2021. His main research work focuses on Optimization Algorithms, Blockchain Technology, Data Mining, Machine Learning, Genetic Algorithm, PSO and Digital Image Processing. He has 14 years of teaching experience and 10 years of Research Experience. He can be contacted at email: narenrawal@gmail.com