A Data-Driven Transition Model from Stem to UHV-STEM: Predictive Analytics for Holistic Student Outcomes Using Clustering

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

[1] Department of Computer Science, Doon University, Dehradun, Uttarakhand, India.
(dheerajsingh.phd@gmail.com)

[2] Department of Computer Science, Doon University, Dehradun, Uttarakhand, India.

Corresponding Author:

Dheeraj Kumar Singh[1*]
Email: dheerajsingh.phd@gmail.com

Abstract: Sustainable development is always possible through imparting and acquiring quality education, which should be oriented towards value-based skill development. The Science, Technology, Engineering, and Mathematics (STEM) education model adopted by countries definitely enhances the skills across the generations from the last many years, but this model has become a failure in giving a sustainable, holistic life to the people. For the alignment towards the sustainable development goals (SDG), the education model ought to be Universal Human Values (UHV)-based STEM education (UHV-STEM), which will integrate the values with the skill acquisition.So, this study will provide a path to move from STEM education to UHV-STEM education. In this, a data-driven analysis through using clustering of Educational Data Mining (EDM) has been adopted and analyzed a model of education that is UHV-STEM. The study identifies patterns and correlations that demonstrate the UHV impact on STEM education. K-means, DBSCAN, and other clustering algorithms are used to group data, providing insights into the effectiveness of UHV-STEM. The findings validate the need for incorporating UHV into STEM education, leading to a more inclusive, value-driven learning environment..

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

Introduction

Educational Data Mining (EDM) enables the following: predicting student performance, selecting relevant programs and subjects, and designing courses that are tailored to each individual student (Baker & Yacef, 2009).

UHV content is holistic in terms of its nature of being universal, rational, natural and verifiable, all encompassing, and its being the hold for leading to harmony in all human being. UHV content also sustainable at all time, places, and environment due to its characterstic to focus on relation, harmony, and coexistence at levels of living such as individual, family, society, nature and existence, and all orders such as material order, bio order, animal order, and human order (Gaur et al., 2010). As per study, UHV is a language gives the understanding that it is hoistic and sustainable education approach. UHV is a concept which focuses on human being to live with doing participation at all level of living such as individual, family, society, nature and existence, and verifiable at all time and places with seeing all orders such as material, bio, animal, human, and gives the base in human being to live with understanding of relation, harmony, and coexistence.

One of the biggest issues facing education today is predicting student performance accurately enough to offer appropriate assistance and intervention. Students can be categorized using clustering techniques such as K-Means, DBSCAN, BIRCH, and others based on their academic performance, learning preferences, and other relevant criteria. In order to improve student performance, teachers might employ tailored interventions by examining these clusters to identify patterns leading to potential areas of academic strength or weakness.

In STEM / STE(A)M (Science, Technology, Engineering, (Arts,) Mathematics) course prediction analysis, clustering algorithms are used to identify demand and create curricula that align with students' interests (Choi et al., 2017; Liao, 2016;Shin & Shim, 2021) .

One significant area in which clustering is employed in education is course prediction analysis, curriculum design, and course planning (Regueras et al., 2019). Teachers can examine personalized learning routes by classifying students based on their interests, academic abilities, and constraints. Additionally, clustering algorithms can help teachers to identify students who might need special attention by accurately estimating future academic progress. Demographic data, behavior, psychology, family socioeconomic background, and school environment all have a big impact on academics and course choices (Batool et al., 2023;Issah et al., 2023;Kukkar et al., 2024).

This approach enriches students' quality of life by equipping them with the knowledge and abilities needed to thrive in the twenty-first century and secure physical facilities, in addition to preparing them for bright professions in STEM (Cetto et al., 2000). But it can't address the stress, worry, and anxiety developing in the students (Gaur et al., 2010).

In order to create individuals who are not only physically and intellectually skilled but also ethically and cognitively mature, schools need to teach Universal Human Values (UHV) (Singh & Kumar, 2024;Singh & Kumar, 2024).

UHV has the capacity to enact revolutionary transformation on the levels of the person, the family, society, and the natural world (Gaur et al., 2010).

In conclusion, the integration of Universal Human Values with STEM/STEAM education (UHVSTEM / UHVSTEAM) holds significant promise for transforming the educational landscape. This analysis is made feasible by the application of clustering techniques in Educational Data Mining. By using these technologies and approaches, educators can assess student performance, design customized programs and courses, and promote ideas essential to sustainable development. This comprehensive approach to education enhances students' general wellbeing and development in addition to their academic performance, giving them the compassion and self-assurance they need to tackle the challenges of the future.

Factors impacting students' performance

In addition to the aforementioned studies, there are several other factors that could influence students' academic performance, including the decision-making, subject, and program or courses, as well as their prediction and analysis (Table 1), which are explained below:

Table 1. Variables related to students' performance, choice of course, subject, and program

Interest	Motivation	Relevance	Ambition	Confidence
Attendance	Age	Gender	Challenge	Location
Satisfaction	Alignment	Engagement	Support	Marital Status
Hostel Factor	Passion	Curiosity	Inspiration	Recognition
District of Origin	Surrounding Environment	Self Employed	Acceptance Rate	Type of Institute
Past Academic Performance	Behaviour of society	Credit Hours Required	Commination skills	Grades / Marks
Financial Aid/ Scholarship	Location of Institute	Teaching Quality	Teachers Behaviour	Home Issues
Student-to-Faculty Ratio	Value based living of Teacher	Body centric living of teacher	Family Income / Background	Teachers Perception / Motivation
Personal Issues	Daily / weekly study	Group wise study	Make new Friends	Connection with Religion
Father / Mother Education	Bunk from lectures	Writing habits in class	Notes making ability	Engaged in social media / time spent
Course / Program Duration	tuition fees, overall expenses	Job Placement Rates	Average Salary of Graduates	Admission Cut-off Scores

These variables, which can be either qualitative or quantitative in nature, were taken from studies conducted by (Bansal et al., 2023;Crisp et al., 2009;Gasiewski et al., 2012;Gaur et al., 2010;Williams & Williams, 2011).

In addition to the previously mentioned considerations, it is possible to have a better understanding of students' choices about courses, subjects, and programs by considering all of the aforementioned information.

Method

Pattern recognition is data mining's main goal, often known as knowledge discovery in databases (Agrawal et al., 1993). Data mining technologies are specifically used for pattern identification in the educational domain, known as educational data mining, or EDM. Patterns discovered through the use of educational data mining techniques will revolutionize the research and prediction of student growth (Romero & Ventura, 2007).

Finally, EDM may forecast both offline and online education that helps society and the economy through colleges, universities, and other academic institutions. The most common applications of EDM include learning objectives, course completion rates, student identification, curriculum building, efficient test-taking, and performance prediction. It is also used in analysis, course selection support, and the prevention of student dropouts.

To create clusters for grouping analysis, there are several clustering techniques available (Asif et al., 2017;Durairaj & Vijitha, 2014;Peña-Ayala, 2014;Suciati et al., 2023). One of the effective unsupervised learning strategies for finding patterns for predictive analysis is clustering. Among the clustering algorithms that can be used for result formulation for applicability of course, subject, and program formation in educational environments are Affinity Propagation, Mean-Shift, Agglomerative Hierarchical Clustering, Mini-Batch K-Means, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies), OPTICS (Ordering Points To Identify the Clustering Structure), and K-Means Clustering (Ankerst et al., 1999; Khan et al., 2014;Peng et al., 2018;Sasirekha & Baby, 2013;Wang et al., 2018;Wu & Yang, 2007).

It is possible to calculate a confusion matrix or matching matrix that correlates to each method in order to evaluate details like accuracy and efficiency. The confusion matrix's findings provide insight into how well clustering techniques can be used to organize courses based on their characteristics. Teachers can use these data to improve course recommendations, curriculum design, and the overall learning experience for students.

Forecasting any new program, course, or subject is a challenging task, much more than just predicting a student's grade, according to an analysis of multiple studies. Clearly, the work will become even more difficult when figuring out what kind of education is needed worldwide to achieve the UN's fourth sustainable development goal.

The Universal Human Values framework of education has been adopted by the Indian Ministry of Education, AICTE (All India Council for Technical Education), and UGC (University Grant Commission). On the other hand, research indicates that UHV education can address attaining harmony at existential, natural, family, societal, and individual levels of being human more successfully (Gaur et al., 2010).

This study employed a qualitative research design. First, for approximately eighteen hours, teachers were expected to teach UHV material, and for additional eighteen hours, students took part in an efficient train-the-trainers program. Second, using Google Form Analytics, research inquiries with the following constructs—"Strongly Agree," "Agree," "Undecided," "Disagree," and "Strongly Disagree"—were created and based on the Likert five-point scale (Armstrong, 1987).

In the end, as previously stated in the paper, a confusion matrix was produced after a number of clustering techniques were applied. Up to a point of severity, courses that should be part of the current educational system have also been identified by professors and students in the arts and sciences. Here are the listed research questions:

RQ1. IS UHV incorporation with current education will lead students towards Happiness?

RQ2. IS only STEM (Science, Technology, Engineering, and Mathematics) education will lead Harmony in Human being?

RQ3. Is addition of UHV education with STEM education will lead Harmony in Human being?

RQ4. If Value Education cell open-up in your University, would you ready to work as Volunteer for it?

RQ5. Would you like to work as Volunteer for UHV?

RQ6. Would you like to share UHV with school children?

RQ7. Is UHV Based Education - Sanskar will reduce the crime level of society?

RQ8. Is UHV Based 'Education - Sanskar' delivered to students and society will reduce the corruption level?

RQ9. IS UHV based education delivered to students, will it reduces the stress level of students?

RQ10. Will UHV based education be holistic and sustainable education approach.

"Education" can be understood as knowing “what is correct”, whereas "Sanskar" refers to “living well” (Gaur et al., 2010). A research framework or educational framework for a course, subject, or program based on STEM (science, technology, engineering, and mathematics) or NON-STEAM (science + arts) disciplines based on Universal Human Values (UHV) may be predicted by considering several important factors.

Important facets of education include incorporating values into the curriculum, choosing teaching strategies, assessing value integration, professional development for teachers, facilitators, and co-explorers, engaging with stakeholders, and choosing strategies that improve understanding and application of UHV in STEM, STEAM, and non-STEAM fields. Examples of this include role-playing games, case studies involving human life, and reflective strategies that help students consider the moral, ethical, and value-based components of their work. As per the findings of (Gaur et al., 2010;Singh & Kumar, 2024;Singh & Kumar, 2024), the implementation of this value-based collaborative approach is expected to promote complementarity and ensure the relevance and efficacy of the framework.

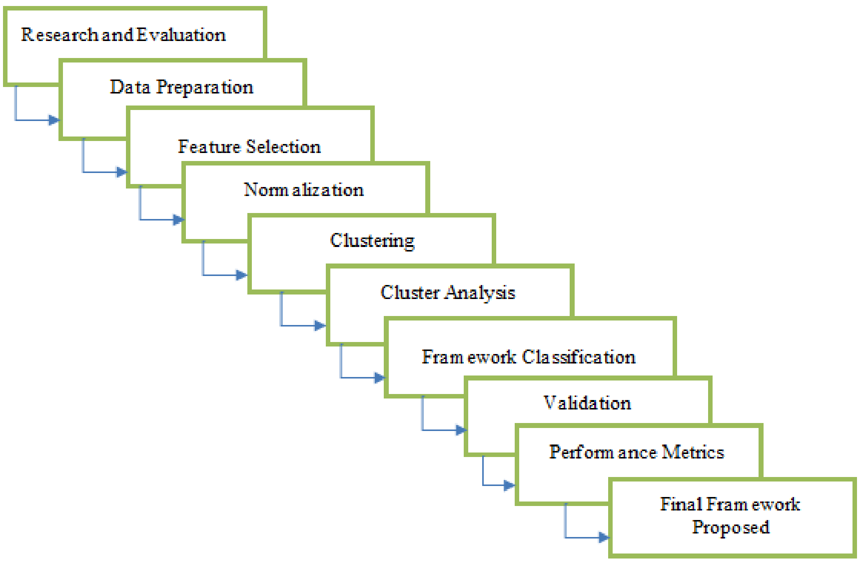

Figure 1. Methodology used for proposing educational framework through clustering

When examining and enhancing research frameworks and curricula based on Universal Human Values (UHV) and targeted at STEM (science, technology, engineering, and mathematics) or NON-STEAM (non-science, technology, and arts) disciplines, clustering algorithms are a helpful tool. Gather data about stakeholder involvement, research and evaluation, assessment strategies, professional growth, and core values. Every data unit needs to match a specific framework component (Feldman-Maggor et al., 2021). To make the clustering procedure easier, choose pertinent features from the dataset (Parhizkar et al., 2023). Apply clustering methods using the customized data. These algorithms will merge similar data points based on the feature values of each individual data point.

Using, analyzing, and executing cluster analysis algorithms, a framework for a UHV-based STEM / STEAM / NON-STEAM course, subject, program, or instruction is being produced. Using the clusters as a reference, ensure that the framework is designed with core principles and promoting UHV in all aspects of teaching and learning.

In order to confirm the proposed framework, share it with the relevant parties and create a comparison between it and the current one. Adapt the framework as needed to make sure it is acceptable, practical, and compliant with UHV standards.

3.1 Mathematical Modeling

Mapping the mathematical and predictive modeling framework to the 10-questions through Likert scale clustering model.

Problem Formaization (Mathematical Layer)

Objective

To identify latent learner orientation (Art vs Science oriented) and predict UHV-STEM course adoptability using responses to 10 Likert-scale questions.

Mathematical Definition

Let:

$Q = \{q_1, q_2, \ldots, q_{10}\}$

Each learner I is represented as:

$x_i = (x_{i1}, x_{i2}, \ldots, x_{i10}),\ xij \in \{1,2,3,4,5\}$

Thus, the dataset is:

$X = R^{N \times 10}$

Data Representation and Preprocessing (Modeling Layer)

Normalization

To ensure equal contribution of all questions (Min-Max Scaling):

$x'_{ij} = (x^{ij} - 1) / 4$

Feature Semantics (UHV Mapping)

Question Group	**Meaning**
$q_1 - q_4$	Human Values & Ethics
$Q_5 - q_7$	Scientific Reasoning
$q_8 - q_{10}$	Creativity & Social orientation

Similarity and Distance Modeling (Mathematical Core)

Distance Function

Learner similarity is measured using:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{10} \left(x'_{ik} - x'_{jk}\right)^2}$$

This distance defines:

ϵ – neighborhoods (DBSCAN / OPTICS)

Cluster compactness (K-Means / BIRCH)

Unsupervised Clustering (Learning Layer)

Step -1: Density-based Structure Discovery

DBSCAN / OPTICS

Detects natural groupings

Handles noise and outliers

Finds intrinsic count

Formally:

$C = \{C_1, C_2, \ldots, C_m\}$

Step 2:

For each cluster C_k :

$$\mu_k = \frac{1}{C_k} \sum_{x_i \epsilon C_k} x_i$$

These centroids summarize value-skill profiles. μ refers cluster centroid (mean vector).

Step 3: Partition Refinement

K-Means / BIRCH

Uses extraceted centroids as initialization:

$$\min \sum_{i=1}^{N} \left|\left|x_i - \mu_{c(i)}\right|\right|^2$$

Cluster Interpretation (Semantic Layer)

Cluster Labeling Rule

Let:

$$S = \sum_{j \epsilon Science} \mu_{kj},\ A = \sum_{j \epsilon Art} \mu_{kj}$$

If:

S > A => Science oriented Cluster

A > S => Art oriented Cluster

This converts numerical clusters -> interpretable learner archetypes.

Predictive Modeling (Inference Layer)

Adoptability Score

$UHV\_Adopt_i = w_1.Value_i + w_2.STEM_i + w_3.Creativity_i$

Where:

$$Value_i = \frac{1}{4}\sum_{j=1}^{4} x_{ij}$$

Cluster membership acts as a predictive prior:

P (Adopt | C_k)

Performance Metrics

The confusion matrix and its associated parameters—Accuracy, Precision, Recall, F1-Score, and Specificity—are needed to arrive at the final goal of the methodology (Xia, 2020). The following are these parameters:

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)

Accuracy measures overall correctness of a model. It also tells what proportion of total predictionas are correct.

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

Precision measures prediction correctness for a specific class. It also tells how many predicted positives are actually correct.

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

Recall measures coverage of actual positives. It also tells how many real positives are correctly identified.

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

F1 Score balances Precision and Recall. It also reflects overall effectiveness for a target class.

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

Specificity measures correct identification of negatives. It also tells how well non-target cases are excluded.

In short methodology include steps as Feature Selection, Data Preparation, Normalization, cluster analysis, framework classification, and finally validation through performance metrics.

Result Analysis

This section presents the findings from a few particular study phases of the provided topics. Several clustering techniques have been identified through educational data mining in order to evaluate the dataset for further study, and the results have been validated. To predict the course UHV-STEM / UHVSTEM, these clustering techniques include K-Means, DBSCAN, BIRCH, Affinity Propagation, Mean-Shift, OPTICS, Agglomerative Hierarchy, and Mini-Batch K-means.

For validating study, the association between different clustering algorithms and characteristics associated to confusion matrices, such as precision, recall, specificity, and sensitivity, has been studied.

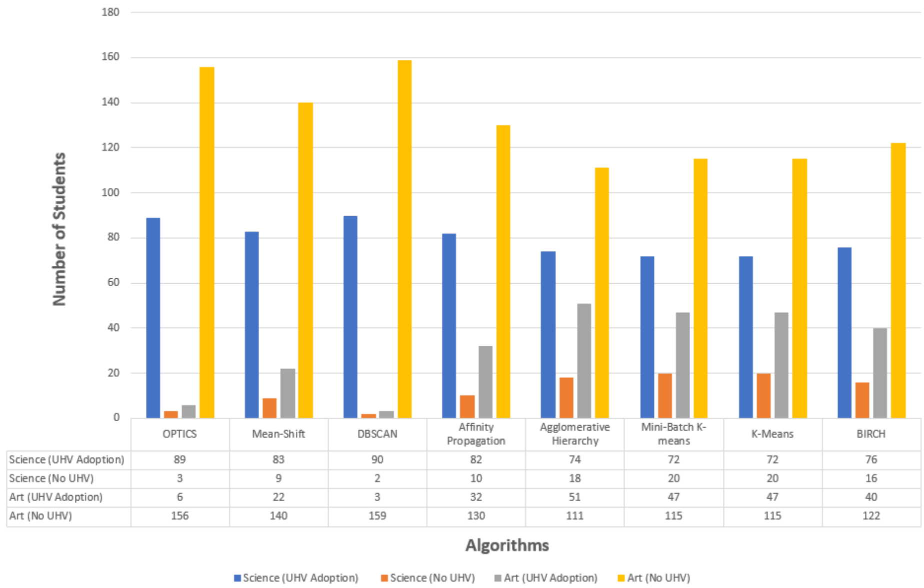

Figure 2. ML Model outcome using Various Clustering Feature Selection Algorithms

The rate of UHV adoption by teachers and students in the art stream is extremely low across all clustering algorithms, whereas it is highly high among those in the scientific stream, as indicated by the confusion matrix seen in Figure 2. The rates of UHV adoption among all science faculty members and students were as follows: The clustering algorithms that yielded the following results: 96.74% by OPTICS, 90.22% by Mean-Shift, 97.83% by DBSCAN, 89.13% by Affinity Propagation, 80.43% by Agglomerative Hierarchy, 78.26% by Mini-Batch K-Means, 78.26% by K-Means, and 82.61% by BIRCH.

As a result, the Mini-Batch K-Means algorithm and K-Means algorithm have decided the lowest predictive analysis rate for UHV, while the DBSCAN clustering algorithm has found the greatest. By creating the confusion matrix, more research has been done on the various classification analysis-related criteria, such as Precision, Recall, F1-Score, and Specificity, for the predictive analysis of the results pertaining to the produced clusters. It is feasible to ascertain the proportion of scientific and art students compared to teachers who take UHV courses because two clusters for these students were made specifically for this study.

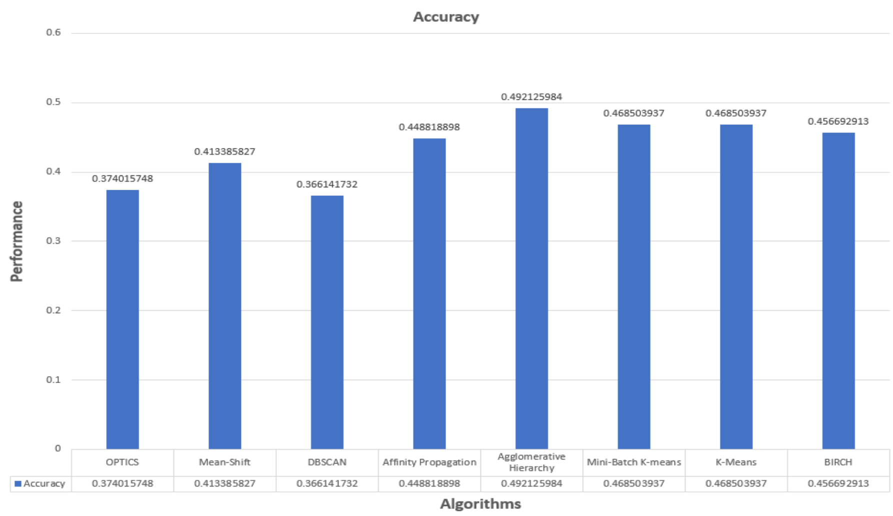

Figure 3.assessing the accuracy level associated with different clustering algorithms

Understanding the real accuracy computation is aided by the depiction in Figure 3. The results show that the Agglomerative Hierarchy method produces the best results, followed by the Mini-Batch K-Means algorithm. Based on this, it is possible that the K-Means algorithm is the most accurate predictive analytic technique, whereas DBSCAN produces the least accurate results.

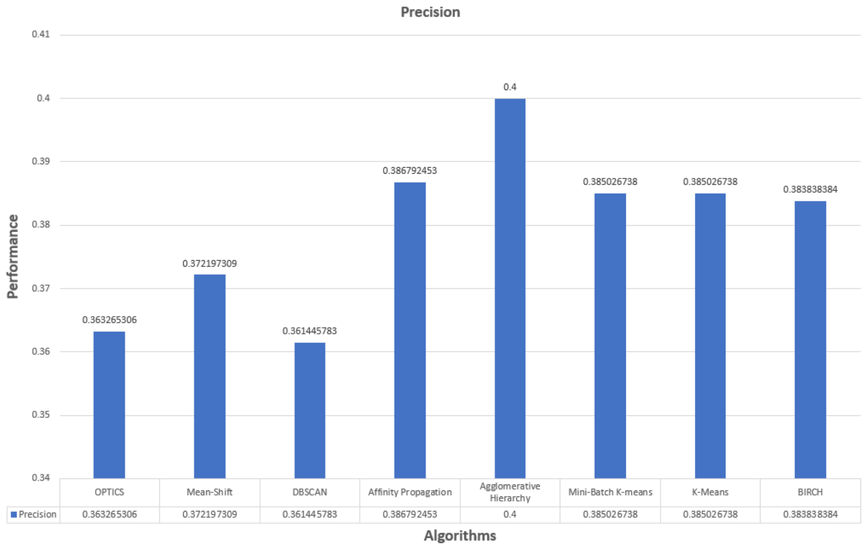

Figure 4.calculating precision level in relation to various clustering algorithms

While DBSCAN produces the lowest precision. Through the visual representation produced in Figure 4, the real computation of precision may be appreciated. Evidently, the Affinity Propagation algorithm comes in second for precision, with Agglomerative Hierarchy algorithm producing the best results.

By examining the representation shown in Figure 5, one can comprehend the recall computation process itself. It is clear that DBSCAN produces the highest recall value, followed by OPTICS. In terms of recall, Mean-Shift is the optimal method for predictive analysis. Ultimately, it is clear that every algorithm evaluated had a recall value more than 78%. It means sensitivity towards adoption of UHV among students vs teachers is very much high. So, UHV based education is necessary in today education system.

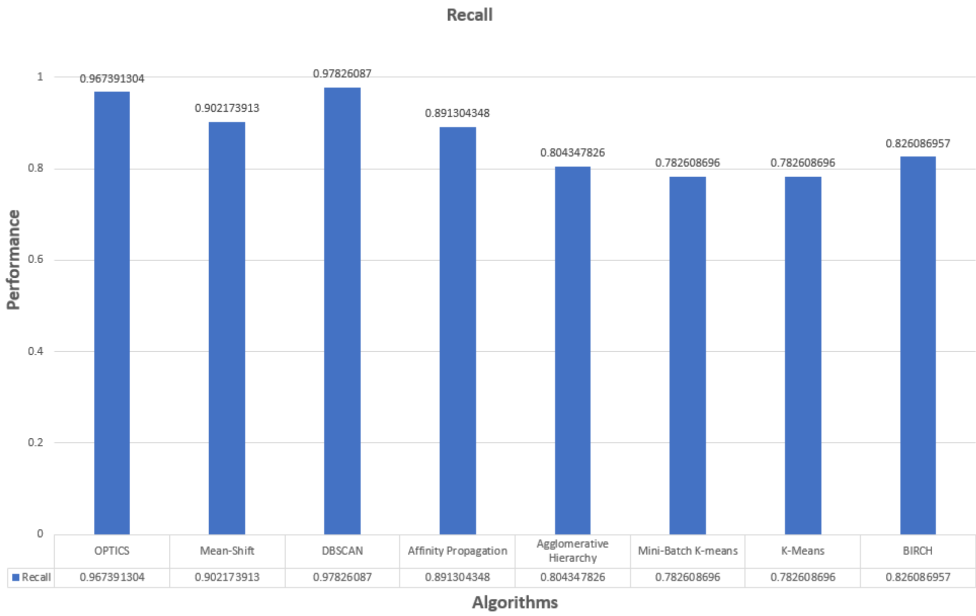

Figure 5.calculating the recall level for various clustering algorithms

The visualization in Figure 6 illustrates the actual computation of the F1-Score. It is clear that the Affinity Propagation algorithm produces the best results for F1-Score, followed by the Agglomerative

Hierarchy algorithm; OPTICS is then the best algorithm for predictive analysis in terms of F1-Score; and the Mini-Batch K-Means algorithm produces the lowest F1-Score.

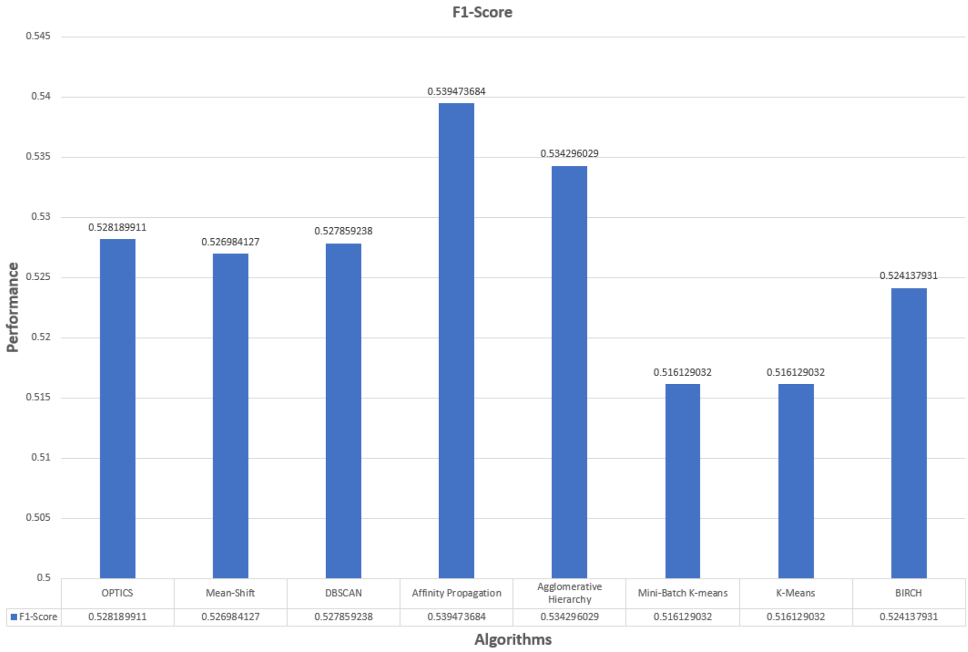

Figure 6.determining the F1-Score level for various clustering algorithms

The illustration in Figure 7 aids in the explanation of the real specificity computation. It is clear that the Agglomerative Hierarchy method produces the highest specificity, whereas the DBSCAN clustering technique produces the lowest.

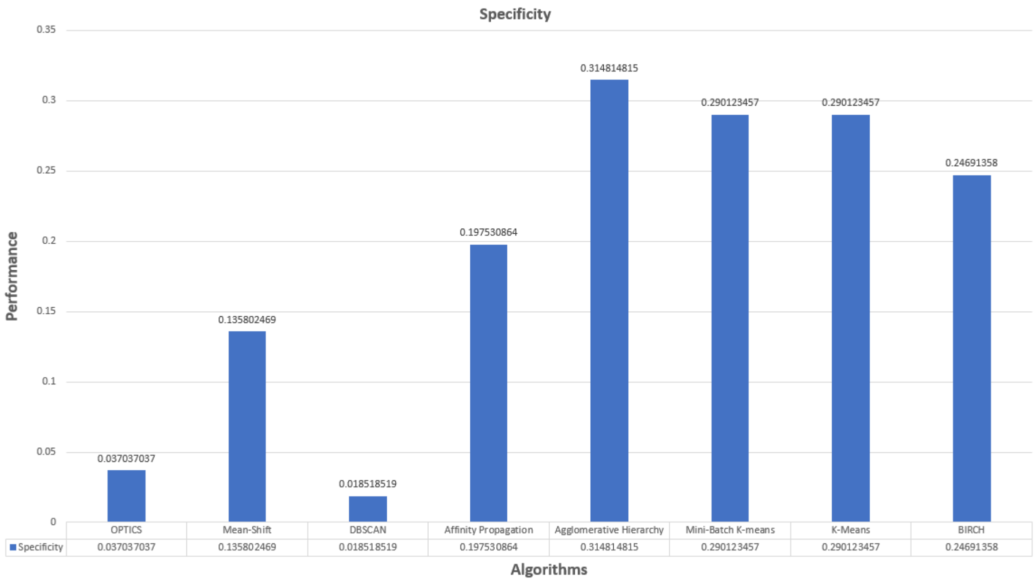

Figure 7. determining the specificity level associated with various clustering algorithms

As demonstrated in Table 3, recall value, accuracy, precision, F1-score, and specificity are only a few of the cluster-based categorization assessment metrics that can be evaluated at one site.

Suitability of clustering Algorithms

Suitability reflects whether an algorithm respects ordinal data, earner diversity, noise, and interpretability Table 2 .

Table 2: Recall-Oriented Educational Suitability of Clustering Algorithms

Algorithm	**Suitability**	**Reason**	**Suitability Meaning in Education**	**Recall Behavior**
OPTICS	Excellent	Handles varying densities	Identifies multi-level learner / value patterns	High recall across multiple learner types
Mean-Shift	Weak	Kernal Density Estimation (KDE) unsuitable for discrete ordinal data	Poor for discrete educational data	Low recall
DBSCAN	Good	Density based, noise-aware	Dtects atypical or disengaged learners	High recall for atypical learners
Affinity Propagation	Conditional	Sensitive to similarity matrix	Exemplar learner identification	Moderate recall
Agglomerative Hierarchy	Good	Works with ordinal distances	Reveals learning hierarchies or stages	Good reacall for structured groups
Mini-Batch K-means	Conditional	Same assumption as K-means	Fast grouping, weak pedagogical validity	Moderate low recall
K-Means	Conditional	Assumes Euclidean distance	Quick grouping, weak pedagogical validity	Low recall for minority learners
Birch	Conditional	Assumes spherical clusters	Institutional-level large data clustering	Moderate recall at institutional scale

Density-based and hierarchical methods better reveal meaningful learner or value patterns, while centroid-based methods offer speed but limited pedagogical insight. Table 2 explains how different clustering algorithms align with educational data characteristics. Clustering algorithms are sutables to evaluates algorithm effectiveness, interpretability, and learner recall.

Table 3. Comparisons showing evaluation measures that correspond to student performance using various clustering algorithms.

Algorithms	Accuracy	Precision	Recall	F1-Score	Specificity
OPTICS	0.374015748	0.363265306	0.967391304	0.528189911	0.037037037
Mean-Shift	0.413385827	0.372197309	0.902173913	0.526984127	0.135802469
DBSCAN	0.366141732	0.361445783	0.97826087	0.527859238	0.018518519
Affinity Propagation	0.448818898	0.386792453	0.891304348	0.539473684	0.197530864
Agglomerative Hierarchy	0.492125984	0.4	0.804347826	0.534296029	0.314814815
Mini-Batch K-means	0.468503937	0.385026738	0.782608696	0.516129032	0.290123457
K-Means	0.468503937	0.385026738	0.782608696	0.516129032	0.290123457
BIRCH	0.456692913	0.383838384	0.826086957	0.524137931	0.24691358

Significance of the study

This qualitative study shows that teachers and students have a very high level of sensitivity to UHV adoption. Therefore, the paradigm for education in the modern world need to be STEM (STEM) based on Universal Human Values (UHV) rather than just STEM, as STEM is already prevalent worldwide. By integrating UHV with STEM (UHV-STEM), students will be able to take innovative, personalized courses and learn values that are crucial to their long-term success.

The comparison between both model of education STEM and UHV-STEM can be better understood through Table 4, Figure 8.

Table 4. Comparison Analysis between STEM and UHV-STEM model of Education

Parameter	**STEM (Science, Technology, Engineering, and Mathematics)**	**UHV-STEM (Universal Human Values - Science, Technology, Engineering, and Mathematics)**
Definition:	STEM stands for Science, Technology, Engineering, and Mathematics. It encompasses a broad range of disciplines focused on these fields.	UHV-STEM combines Universal Human Values (UHV) with STEM education. It integrates ethical, moral, and value-based education into the traditional STEM curriculum.
Focus:	**Technical and Analytical Skills:** Emphasizes developing skills in scientific research, technological innovation, engineering design, and mathematical analysis. **Career Preparation:** Aims to prepare students for careers in industries that rely on these core	**Holistic Development:** Emphasizes not only technical skills but also the development of ethical, social, and emotional values. **Ethical Application:** Encourages students to apply STEM knowledge in ways that are socially responsible and aligned with universal human values.

	areas.	
Educational Objectives:	**Skill Development:** Focuses on problem-solving, critical thinking, and technical expertise. **Innovation:** Encourages advancements in technology and science.	**Value-Based Education:** Aims to develop not only technical expertise but also a strong sense of responsibility, empathy, and ethical judgment. **Sustainable Development:** Focuses on using STEM skills for the betterment of society and addressing global challenges in a responsible manner.
Curriculum:	**Courses:** Includes subjects like physics, chemistry, biology, computer science, engineering, and mathematics. **Methodologies:** Often involves hands-on experiments, technical projects, and quantitative analysis.	**Courses:** Includes traditional STEM subjects integrated with courses on ethics, human values, and societal impact. **Methodologies:** Combines technical projects with discussions on ethics, sustainability, and human-centered design.
Summary	**STEM** focuses on developing technical and analytical skills in science, technology, engineering, and mathematics, preparing students for careers in these fields.	**UHV-STEM** integrates Universal Human Values with STEM education to ensure that students not only excel technically but also understand the ethical and societal implications of their work, promoting holistic development and responsible innovation.

In this case, clustering approaches are used to independently analyze the significance of UHV for educators and learners in STEM and non-STEM domains. Predicting an educational model (UHV-STEM) for a social, economic, sustainable, co-existential, holistic, and all-encompassing approach to education is the ultimate goal of this research endeavor.

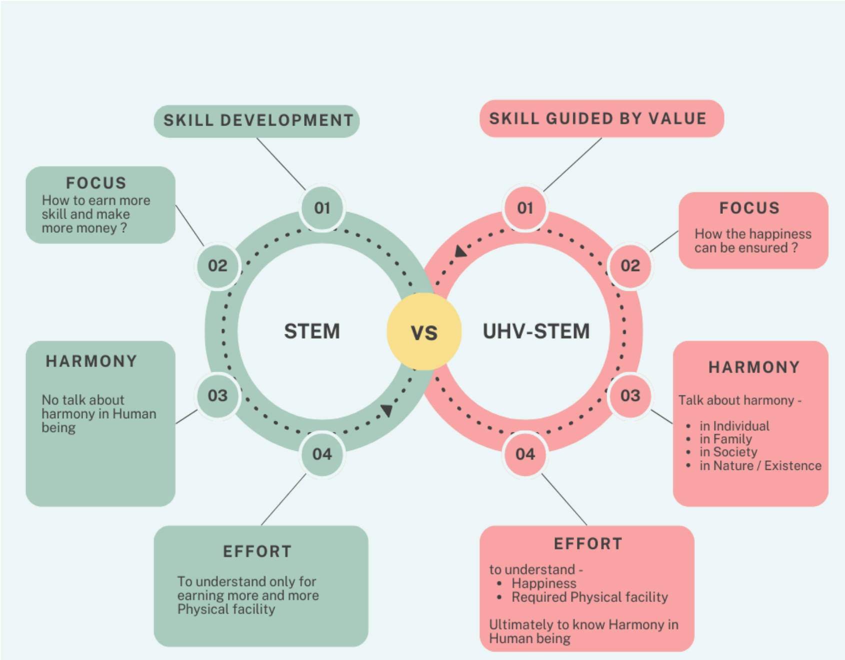

Figure 8: Learnability comparison between STEM and UHV-STEM Model of Education

Conclusion and Future Work

The study aimed to explore the incorporation of Universal Human Values (UHV) into science and art courses, focusing on the perspectives of both teachers and students. It was found that scientific instructors and students have higher feelings for UHV subjects, indicating the need for its mandatory study. STEM education does not foster human relationships, even though it may increase skill sets; this might lead to a separation in later life. EDM algorithms were used to cluster the dataset, and the outcomes showed how crucial UHV is for understanding human relationships and producing happiness.

The mathematical study done in this research through likert vectors, and distances computaion. Cluster based classification (Art vs Science) used for predictive analysis, and UHVSTEM / UHV-STEM framework proposed is a reusable model for UHV course adoptablity.

The relevancy of data is not drawn from the analysis charts or graphs that determine whether UHV should be implemented or not; rather, the conclusion is drawn from the nature of questionnaires as to how it can be analyzed or predicted that UHV should be implemented or not, or up to what level it should be implemented and how.

It is observed and analyzed through questionnaires and clustering algorithms employed that to promote a happy, prosperous, and peaceful society, the UHV-STEM / UHV-STEAM educational model ought to be employed. By evaluating things like comparison outcomes, impact assessments for students' holistic development, and self-assessment batteries based on UHV content exploration among students and teachers, future research may be able to predict the behavior of students.

References

Agrawal, R., Imieliński, T., & Swami, A. (1993, June). Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD international conference on Management of data (pp. 207-216).

Ankerst, M., Breunig, M. M., Kriegel, H. P., & Sander, J. (1999). OPTICS: Ordering pointsto identify the clustering structure. ACM Sigmod record, 28(2), 49-60.

Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. Computers & education, 113, 177-194.

Armstrong, R. L. (1987). The midpoint on a five-point Likert-type scale. Perceptual and motor skills, 64(2), 359-362.

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.

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.

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.

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

Choi, Y., Lim, Y., & Son, D. (2017). A semantic network analysis on the recognition of STEAM by middle school students in South Korea. EURASIA Journal of Mathematics, Science and Technology Education, 13(10), 6457-6469.

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. 924-942.

Durairaj, M., & Vijitha, C. (2014). Educational data mining for prediction of student performance using clustering algorithms. International Journal of Computer Science and Information Technologies, 5(4), 5987-5991.

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.

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.

Gaur, R.R.,Sangal, R., & Bagaria, G.P.(2010). A Foundation course in Human Values and Professional Ethics. 2010, Excel Books Private Limited.

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.

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) (pp. 232-238). IEEE.

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.

Liao, C. (2016). From interdisciplinary to transdisciplinary: An arts-integrated approach to STEAM education. ART education, 69(6), 44-49.

Parhizkar, A., Tejeddin, G., & Khatibi, T. (2023). Student performance prediction using datamining classification algorithms: Evaluating generalizability of models from geographical aspect. Education and Information Technologies, 28(11), 14167-14185.

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.

Peng, K., Leung, V. C., & Huang, Q. (2018). Clustering approach based on mini batch kmeans for intrusion detection system over big data. IEEE access, 6, 11897-11906.

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.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. Expert systems with applications, 33(1), 135-146.

Sasirekha, K., & Baby, P. (2013). Agglomerative hierarchical clustering algorithm-International Journal of Scientific and Research Publications, 83(3), 83.

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.

Singh, Dheeraj Kumar, Kumar Narender (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 (pp. 145-164). London, United Kingdom: HP HAMILTON LIMITED

Singh, Dheeraj Kumar & Kumar, Narender (2024). A Systematic Survey for Students Performance Prediction with Holistic and Sustainable Education approach using Educational Data Mining. Tuijin Jishu/Journal of Propulsion Technology, Vol. 45 No. 2(2024), 232-269. Retrieved from https://www.propulsiontechjournal.com/index.php/journal/article/view/5698

Suciati, N., Fabroyir, H., &Pardede, E. (2023). Educational Data Mining Clustering Approach: Case Study of Undergraduate Student Thesis Topic. IEEE Access, 11, 130072-130088.

Xia, X. (2020). Clustering Analysis of Interactive Learning Activities Based on Improved BIRCH Algorithm. arXiv preprint arXiv:2010.03821.

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.

Williams, K. C., & Williams, C. C. (2011). Five key ingredients for improving student motivation. Research in higher education journal, 12, 1.

Wu, K. L., & Yang, M. S. (2007). Mean shift-based clustering. Pattern Recognition, 40(11), 3035-3052..