



Category: STEM (Science, Technology, Engineering and Mathematics)

ORIGINAL

Improved Cuckoo Search Optimization and Transductive Support Vector Machine Algorithm for E-Learning Recommendation System

Algoritmo de Máquina Vectorial de Soporte Transductivo y Optimización de Búsqueda de Cuco Mejorado para el Sistema de Recomendación de Aprendizaje Electrónico

D. Poornima¹  , D. Karthika²  

¹Research Scholar, P.K.R Arts College for Women. Gobichettipalayam, India.

²Associate Professor in Computer Science, VET Institute of Arts and Science College. Erode, India.

Cite as: D. P, D. K. Improved Cuckoo Search Optimization and Transductive Support Vector Machine Algorithm For E-Learning Recommendation System. Salud, Ciencia y Tecnología - Serie de Conferencias. 2024; 3:1118. <https://doi.org/10.56294/sctconf2024.1118>

Submitted: 19-02-2024

Revised: 12-05-2024

Accepted: 02-09-2024

Published: 03-09-2024

Editor: Dr. William Castillo-González 

Corresponding Author: D. Poornima 

ABSTRACT

Introduction: growing numbers of students opt for self-learning via the Internet, an established e-learning approach, as a result of the popularity and advancement of data search technology. A challenge for e-learning has constantly been the ability to learn different knowledge items methodically and effectively in a certain topic because the majority of the learning material on the network is dispersed. Still, the existing system has issue with higher error rate and computational complexity.

Method: to overcome this problem, Improved Cuckoo Search Optimization (ICSO) and Transductive Support Vector Machine (TSVM) algorithm were introduced. The main steps of this research are such as pre-processing, clustering, optimization and e-learning recommendation.

Results: initially, the pre-processing is performed utilizing K-Means Clustering (KMC) which is focused to deal with noise rates effectively. Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is used to cluster data where data space's dense object regions are examined to divide low-density areas. In the improved DBSCAN method, density reachability and density connectedness are used. Then, ICSO algorithm is applied to fine tune the parameters using best fitness values.

Conclusions: finally, the classification of recommendation system is done by using TSVM algorithm which more precise outcomes for the specified datasets. According to the findings, the recommended ICSO-TSVM approach excels the existing ones regards to higher accuracy, recall, precision, mean absolute error (MAE), and also time difficulty.

Keywords: E-Learning; Recommendation System; Clustering; Improved Cuckoo Search Optimization (ICSO); Transductive Support Vector Machine (TSVM).

RESUMEN

Introducción: un número creciente de estudiantes opta por el autoaprendizaje a través de Internet, un enfoque de aprendizaje electrónico establecido, como resultado de la popularidad y el avance de la tecnología de búsqueda de datos. Un desafío para el e-learning ha sido constantemente la capacidad de aprender diferentes conocimientos de manera metódica y efectiva sobre un tema determinado porque la mayoría del material de aprendizaje en la red está dispersa.

Método: aun así, el sistema existente tiene problemas con una mayor tasa de error y complejidad computacional.

Para superar este problema, se introdujeron la optimización mejorada de búsqueda de cuco (ICSO) y el algoritmo de máquina de vectores de soporte transductivo (TSVM). Los pasos principales de esta investigación son el preprocesamiento, la agrupación, la optimización y la recomendación de aprendizaje electrónico.

Resultados: inicialmente, el preprocesamiento se realiza utilizando K-Means Clustering (KMC), que está enfocado a manejar las tasas de ruido de manera efectiva. La agrupación espacial de aplicaciones con ruido basada en densidad (DBSCAN) se utiliza para agrupar datos donde se examinan las regiones de objetos densas del espacio de datos para dividir áreas de baja densidad. En el método DBSCAN mejorado, se utilizan la alcanzabilidad de la densidad y la conectividad de la densidad. Luego, se aplica el algoritmo ICSO para ajustar los parámetros utilizando los mejores valores de aptitud.

Conclusiones: finalmente, la clasificación del sistema de recomendación se realiza mediante el algoritmo TSVM, que proporciona resultados más precisos para los conjuntos de datos especificados. Según los hallazgos, el enfoque recomendado por ICSO-TSVM supera a los existentes en cuanto a mayor exactitud, recuperación, precisión, error absoluto medio (MAE) y también dificultad de tiempo.

Palabras clave: Aprendizaje Electrónico; Sistema de Recomendación; Agrupación en Clústeres; Optimización Mejorada de Búsqueda de Cuco (ICSO); Máquina de Vectores de Soporte Transductivo (TSVM).

INTRODUCTION

E-learning is currently regarded as one of the cutting-edge approaches to online learning in the long run compared to the conventional direct technique of teaching with society. An ever-growing number of students have benefited from various courses due to e-learning. The conservative one-size-fits-all system of education, in which a single layout of learning resources is prescribed to each student, are challenged by the enormously diverse range of pupils on the internet. Prominent recommenders have limitations and problems, such as significant differences in predicted absolute mistakes, slower query processing, and worse recommendations. E-learning has become more popular in recent years as an alternative to traditional education for achieving universal educational targets. E-learning offers a variety of explanations, some of which can be confusing.^(1,2) E-learning uses innovations of web for providing efficient education to students irrespective of locations or timings.

Online education has drawn much interest. Developments in Information Communication Technologies have resulted in scholars creating several e-learning platforms with individualized learning algorithms to help learners learn more effectively. A path of study is a collection of learning resources designed to help students advance their skills in particular topics or degree programs.⁽³⁾ Learning design can be a difficult task, particularly for students. Self-learners could construct effective learning paths and identify the right learning resources with the aid of a learning design recommendation system. Learning object recommendation systems have benefited from an astonishing number of innovations brought forth by educational data mining (EDM).

A growing kind of e-learning is self-learning using the Internet, which is becoming widespread as information search technology advances.^(4,5) The method that knowledge is transmitted to pupils has been dramatically revolutionized by e-learning. The learning materials are frequently deemed to lack relevance to the subject matter being presented. While students could use a variety of materials and videos for studying, retaining track of which ones are pertinent to them is a laborious task.

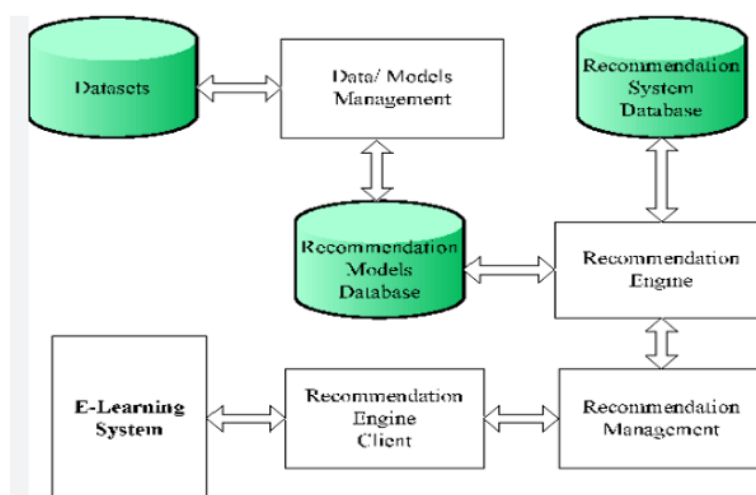


Figure 1. E-learning recommender systems

The use of recommendations systems is crucial in these circumstances. E-learning has benefits over the conventional teacher-centered learning method since it allows students to search and examine the learning target anytime, not just in class. It is constantly difficult for e-learning to study many knowledge items in a focused topic methodically and effectively because most learning content on the internet is scattered.⁽⁶⁾ Figure 1 depicts recommender systems for e-learning.

In academia, e-learning is now widely utilized. Since there are no in-person interactions in virtual classes, it is challenging to monitor student engagement and identify performance declines early through direct interaction, as is normally done in a classroom context. Therefore, it is imperative to use cutting-edge, non-traditional methods to drastically enhance the learning results of online courses. Learners' online footprints, comprising quiz results, logged entries, and login frequency, are frequently collected by web learning management systems used to provide university courses. This data's patterns can substantially aid instructors in understanding students' learning habits.⁽⁷⁾

Potential solutions include data mining and machine learning (ML) tools to examine pupil's online footprints to solve these issues. Software tools might be developed to profile students, recognize individuals who are performing poorly, and provide instructors with advice for improvement.⁽⁸⁾ The student learning experience and educational result would both be enhanced by the timely and targeted instructor involvement. DM and ML are frequently employed in education to assist educators in comprehending their pupils and improve learning outcomes. Online programs have evolved into a crucial component of higher learning due to its adaptable and typically focused on learner's design of syllabuses. As a result, there is an urgent need to greatly enhance the educational benefits of E-learning utilizing atypical methodologies. Automatic evaluation of pupil behaviour and performance along with intelligent and specific solutions using cutting-edge DM and ML techniques have a great deal of potential to accomplish the goal.

This study's primary goal is to establish an online learning recommendation system. Despite the numerous Studies and methods that have been created, the speed and precision of query processing are not considerably guaranteed. The current methods suffer from high error rates and erroneous findings. ICSO-TSVM technique address these issues while enhancing general efficacy of recommendations. Pre-processing, clustering, optimization, and e-learning recommendation are the primary contributions of this research. The recommended approach employs effective algorithms for the given datasets to produce better outcomes.

The research's remaining sections are organized as follows: Section 2 provides a concise summary of a few of the existing study in the recommendation system for e-learning. Section 3 provides specifics on the suggested technique for the ICSO-TSVM system. Section 4 provides the findings and a discussion of the performance analysis. Also, Section 5 summarizes the findings.

Related work

Wang et al.⁽⁹⁾ suggested A blended model-based method for recommending films is utilized, which partitions altered user space using Genetic Algorithms (GAs) and upgraded K-means clustering. It uses a data reduction approach called Principal Component Analysis (PCA) to compact the movie population space, which could lower the cost of intelligent movie guidance. When comparing to current techniques, the findings using Movie lens dataset show that the methodology can deliver high performance with respect to of accuracy and produce more dependable and customized movie suggestions.

Guo et al.⁽¹⁰⁾ developed a technique for Multiview clustering in which individuals are iteratively grouped based on both perspectives on rating trends and social trust relationships. Generating a forecast for a specific item utilizing a support vector regression technique based on attributes connected to the user, the item, and the prediction, allowing users who concurrently show in two different clusters. It used a probabilistic technique to provide an estimate that might accommodate people who lack sufficient knowledge to form clusters, such as chilly persons. Studies shows this method may successfully increase recommendations' accuracy and coverages in cold start scenarios, bringing clustering-based recommender systems closer to practical application.

Li et al.⁽¹¹⁾ suggested integrating recommenders in e-learning based on cross-platform HTML5 with hybrid programming for personalized digital education systems using Collaborative Filters (CF). The mobile terminal uses HTML5 and the structure to provide users with individualized lessons through collaborative filtering and recommendation methods. Server-side development makes advantage of well-known B/S architectures and development methodologies. The accuracy and efficacy of the suggestion are significantly improved by augmenting traditional recommendations CF resulting in course recommendations that were more in accordance with user preferences. This platform offers two forms of online instruction: interactive online live teaching and recorded videos. Each course has its own online classroom; the lecturer will post the timetable,

Intayoad et al.⁽¹²⁾ introduced strategy that is efficient in a dynamic setting utilizing contextual bandits and reinforcement learning issues. Additionally, establish a policy that will allow the reinforcement agent to make the best choice based on present pupil condition and previous pupil behaviours. To assess the approach, use actual data from an online learning system. The approach is contrasted with well-known approaches for

solving problems involving reinforcement learning, such as upper bound confidence, greedy optimistic starting value, and ϵ -greedy. The findings show that technique performs noticeably better in the case test than those benchmarking techniques.

Aziz et al.⁽¹³⁾ presented revised cuckoo search algorithm (CSA) for feature selection to handle high dimensionality data using a using rough sets. The altered CSA mimics the Lévy flight behaviour of birds along with cuckoo species. The fitness function for the altered cuckoo search incorporates multiple characteristics and classification using the rough sets theory. The method is evaluated and benchmarked using a variety of assessment criteria in multiple benchmark datasets, and an additional analysis is conducted using the Analysis of Variance test. On discrete datasets, the approach additionally contrasted with the current techniques. Also, the efficiency of the technique is assessed using K-nearest Neighbours and Support Vector Machines algorithm. The findings demonstrate that the method can substantially enhance classification performance.

Bhaskaran et al.⁽¹⁴⁾ developed a special custom hybrid recommender that applies the transductive support vector method to machine learning on open data sets. The learning process has been enhanced by the students' habits. This research proposes and investigates numerous new approaches to enhance a hybrid recommender's efficacy. The learner dataset is meant to be prepared using the modified one-source denoising approach. To improve performance metrics, an optimization strategy for a modified anarchic society has been devised. Learners' sequential patterns are found using the expanded and updated sequential pattern approach. The recovered behaviours and interests are assessed by means of the enhanced transductive support vector machine. Children receive the best support when these creative methods of measuring their unique rates are used. Publicly accessible machine learning datasets, including those from dating, scholarly publications, and open university courses, are used to assess the expanded model. The results demonstrate that different sizes of groups may be found using the enhanced clustering technique. For accuracy, ranking score, expected absolute error, and recall, the suggested tactics performed better than the previous approaches.

METHOD

Here, ICSO-TSVM technique is suggested to enhance the recommendation system for e-learning more effectively. It contains main steps are such as pre-processing, clustering, optimization and e-learning recommendation. Figure 2 displays the suggested method's general block diagram.

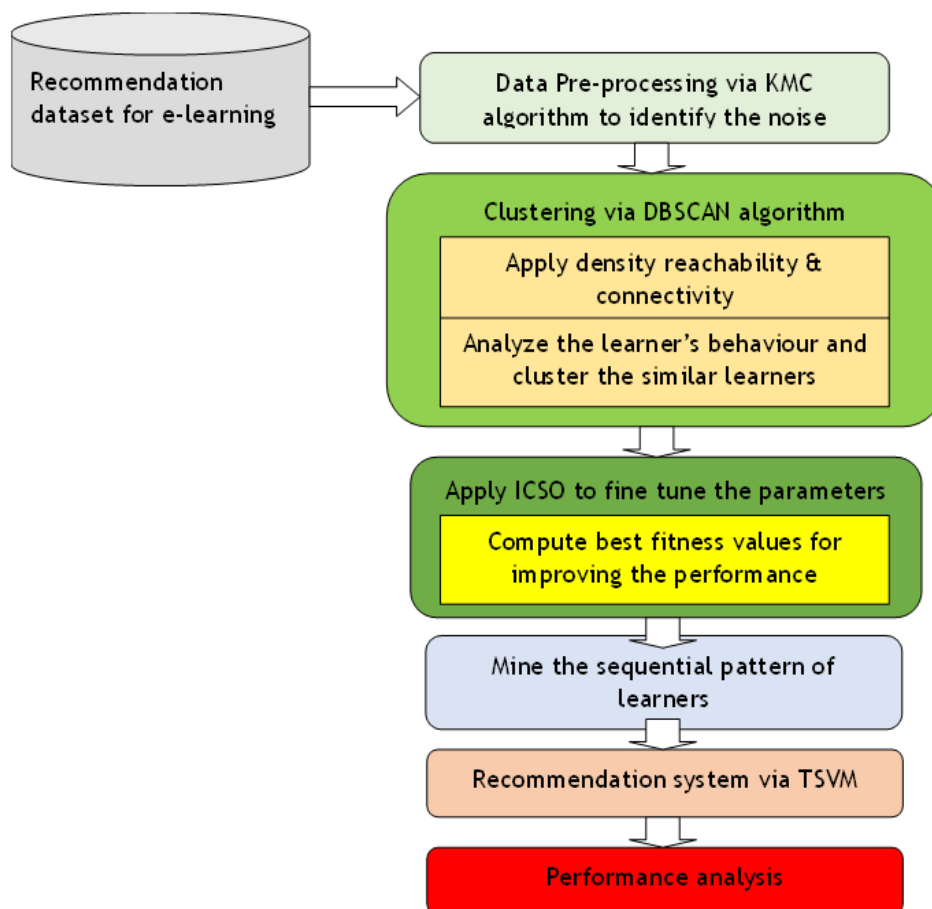


Figure 2. Overall block diagram of the proposed system

Pre-processing via K-Means Clustering (KMC) algorithm

This work uses KMC algorithm in pre-processing to enhance recommendations of accuracy in datasets. Input datasets consist of books, articles and open university learning recommendations where KMC clustering groups comparable data from initial centroids of clusters⁽¹⁵⁾ identified by euclidean distances. Beginning with random partitioning iterative computations are executed for (i) centers of current clusters (i.e., average vectors in data spaces for clusters), and (ii) sends information to clusters whose centers are closest. When there are no more assignments, the course is ended. This method reduces intra-cluster variances locally, which are total squares of variances between data attributes and corresponding cluster centers. An example of the KMC algorithm is shown in figure 3.

The primary benefits of KMC implementation are its efficiency and linear runtimes. One cluster per class is the maximum number that can exist. To get the cluster centroids, use the following formula to calculate the Euclidean distance.

$$d(i,j) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Here x_i and y_i are two points in Euclidean n-space.

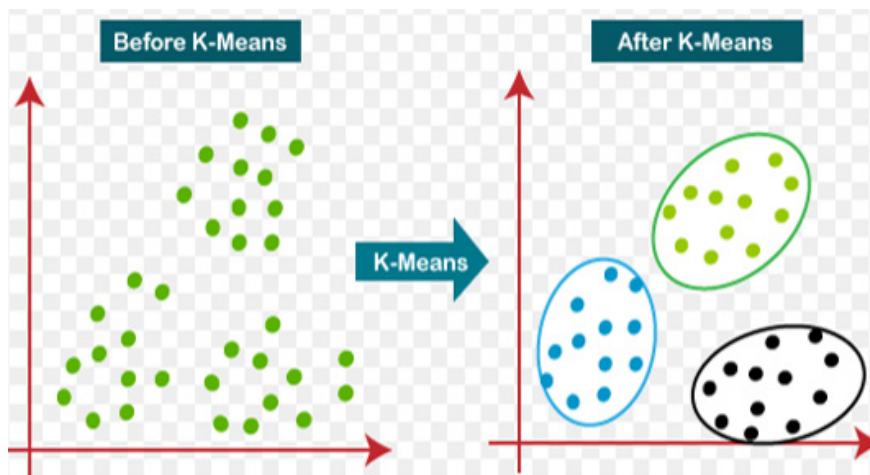


Figure 3. Example of KMC algorithm

Algorithm 1: KMC algorithm

Input: Recommendation dataset for e-learners

Output: Pre-processed data

1. From the dataset(D), choose k clusters
2. Create cluster centers $\mu_1 \dots \mu_k$
3. select k data points for alignments with cluster centers.
4. Compute cluster means and arbitrarily allocate cluster points.
5. Each data point's closest cluster center should be identified, and the noise values should be calculated using that distance.
6. Add clusters to data points.
7. Recalculate the cluster centers (cluster mean of data points).
8. Find and eliminate errors and missing values
9. On lack of fresh reassignments, stop.

The dataset does not include the cases from the original dataset that had missing characteristics. The dataset has been divided into two groups: fully complete examples with no missing values and partially complete examples with missing values from the other group. KMC is used to collect complete instances in order to form clusters of complete instances. As a result, each instance is examined separately, and any missing attributes are substituted in with potential values. The newly included instance is verified to determine whether it has been clustered in the right class after KMC has been used to the dataset from the resulting clusters. The allocated value becomes permanent in the appropriate cluster, and the procedure continues for the next instance. The next value, which puts the instance in the improper cluster, will be analyzed until the right cluster, is located. Consequently, the KMC technique is applied in the preprocessing strategy may improve the recommendation classification accuracy.

Clustering by using DBSCAN algorithm

Only for entire classes of students are inappropriate suggestions generated, as students' abilities to use connected learning notices and their task management skills can fluctuate appropriately in response to changes in their data stage. This work used the clustering approach for elements in information space to partition low-density object areas.

The idea of density reach ability and density connectivity is used in the upgraded DBSCAN strategy. $\{x_1, x_2, x_3 \dots x_n\}$ reflect the quantity of learner data positions that render data pointsets. This method uses two parameters: minimum-points, or the fewest points are required to build clusters, and distances.⁽¹⁶⁾ To locate the learners in gap, all that is needed is to identify every maximal solidity linked gap. Groups are the names for these areas. Data from each learner that is not part of a group is considered noise and is disregarded by the original learners. The idea of solidities like reachability and connection are used. In algorithm 2, the increased clustering is displayed.

Algorithm 2: DBSCAN algorithm for clustering the learners

1. Initialize the data positions
2. Extract the neighbourhood region of the learners data positions
3. If there is related learners then clustering progression initiates & notice learners data position
4. Else learners data position are labeled as noise
5. If learners data positions are clustering element then repeat it
6. Check the newly visited learner's data position and add additional cluster
7. Continue this process of learner's data position until one time visit
8. End

Density reach ability

When learners' data at locations q have distances of p , the optimal values of students' data in places where there are neighbors in space ϵ , the positions p are considered as solidly available from position q 's learner's data.

Density connectivity

When students at positions r have sufficient data at surrounding points, data at both positions p and q are supposedly related in the paths of their strengths. This is a progression of events. If q is the nearby data for the learner at point r , r is the nearby data for the learner at points, is the nearby data for the learner at point t , is the nearby data for the learner at point p , it follows that q is the nearby data for the learner at point p .

Computing ϵ and minimum-points

Two learner's data positions can be entered via the dynamic steps, and the automatically generated variables vary depending on the goal of each pupils data points. Distance (a, b) is used to calculate the separation between the two learners' data places and is described as follows:

$$\text{distance} = q \sqrt{(\mu_1 |X_{a1} - X_{b1}|^q + \mu_2 |X_{a2} - X_{b2}|^q + \dots \mu_p |X_{ap} - X_{bp}|^q)} \quad (2)$$

Where $a = (x_{a1}, x_{a2}, \dots, x_{ap})$ and $b = (x_{b1}, x_{b2}, \dots, x_{bp})$ are data positions for two p -dimensional learners, and $q > 0$.

Apply ICSO algorithm for fine tune parameters

Here, The DBSCAN method's output is subjected to the cuckoo search optimization method in order to improve the outcomes. A fitness function is built into the cluster to help users improve their centroid distances while changing prior centroids for a finite iteration. Next, categorise the users once more by figuring out the minimum centroid variances or by once more using DBSCAN. Three idealized rules that can be used to characterize the Cuckoo Search algorithm are as follows:

Assume a correlation for the framework in which an individual is compared to an egg. Every nest may be thought of as a cluster. Setup involves inserting a random population of n host nests to begin the process. The optimum answer is then obtained by calculating the behavior of levy flights and using the fitness function to estimate fitness.^(17,18) A random walk known as a "Levy flight" has step lengths that are distributed randomly. Laplace and Fourier transforms can be used to determine the step length and Levy stable distribution. The flight length is calculated using the cuckoo optimization method, where $x(N) \sim N(1/a)$ where $0 < a < 2$, x is random parameter and N is step size. The calculation mentioned earlier can be used to determine the cuckoo's travel distance iteratively. When selecting an arbitrary nest, State J compared the cuckoo egg fitness with host eggs fitness. If the quantity of the cuckoo egg's fitness value is less than or equal to the quantity of the arbitrarily chosen nest's fitness function, the quantity of the arbitrarily identified nest, j , is replaced by the initial resolution as stated in equation (5).

Finding new solutions and Levy flight

Levy flight is a technique used in the ICSO inspired feature selection technique to discover novel solutions to equation (3). To generate novel alternatives, a quick tour of the current best solution must be employed; this will expedite the local search. The novel solution for cuckoo i is obtained via Levy flying and is shown below as $x_i^{(t+1)}$.

$$x_i^{(t+1)} = x_i^{(t)} + C \oplus \text{Levy}(s, \lambda) \quad (3)$$

When step sizes are represented by t , step lengths follow Levy distributions.

$$\text{Levy}(S, \lambda) \sim s^{-\lambda}, \quad 1 < \lambda \leq 3 \quad (4)$$

Feature selections of training data are clustered where test accuracy of testing data is (f) . Equation (5) provides the accuracy of minimum distance clustering.

$$\text{Fitness function} = \text{Current best solution} - \text{Previous best solution} \quad (5)$$

The outcome of utilizing ICSO, the given dataset's vital and crucial properties are correctly selected.

Parameter Pa

Employing equation (6), the P_a value is modified dynamically in improved cuckoo search.

$$P_a = P_a \max - \frac{P_a \max - P_a \min}{\text{iter_max}} * \text{iter} \quad (6)$$

Algorithm 3: ICSO

Input: Recommendation dataset

Output: Optimal features

Generate populations of n host nests, starting with m eggs (features); for each nest, choose a cuckoo type (say i) at random based on $(t < \text{MaxGeneration})$ or (stop criteria).

Look at the species of cuckoo.

If cuckoo types are common cuckoos

Crossovers produce two best eggs in nests, and subsequently best eggs are selected.

If cuckoo_type = European cuckoo, then otherwise

By interacting with the uniform mutation operator, any two eggs in the nest can be utilized to create another two eggs.

The best egg in the nest should be selected.

Else

Produce an egg using an arbitrary method (cryptic egg)

End if

Examine its fitness f

Choose eggs in nests with worst solutions(j)

If $(f_i > f_j)$

Change j to the updated solution i .

End if

Sort the egg according to the solution.

Get eggs with greatest solutions from m eggs in nests

Remove some of the eggs that have the worse solutions and replace them with fresh eggs utilizing levy flight

(3) and (4)

Compute accuracies with (5)

Employing (6), update parameter

Maintain the greatest solutions

End for

To determine the present top egg, rank all the eggs according to fitness value.

Choose the elements that are essential and crucial.

End while

To enhance the number of informative characteristics for the provided dataset, the ICSO optimization technique is applied. With more iterations, the divergence among solutions reduces as the fitness function's value moves toward zero, thus it is concluded that if a conventional egg and a cuckoo egg are similar, the host bird will have trouble telling them apart. The fitness is determined by the variances and the substitution of the alternative with the arbitrarily selected solution. When cuckoo eggs are better matches than randomly selected nests, host birds can tell the differences between the two.

The host bird nest and the clusters used to categorize users by similarity of interests are both metaphors for the cuckoo optimization method, where each egg represents a user. The users who were initially categorised are regarded as host bird eggs. The clusters (nests) are initialized by choosing a few random individuals. The clusters are set up, and DBSCAN is used to categorize the users who were initially chosen, as stated in the technique. Users who are not chosen are picked at random. A randomized cluster and the fitness function are chosen for each arbitrarily chosen user. Depending on the user's fitness function, an egg may remain in the nest or be detected by the host bird. After hatching, the cuckoo egg tries tossing other eggs out of the nest at arbitrarily. The strategy succeeds if fitnesses of cuckoo eggs exceed predetermined percentages of users in the cluster.

Mine the sequential pattern of learners

Distinct learning styles result in distinct successions of repetitions for students. Learners are thus classified based on their behaviours, and using the EGSP, personal conduct standards are then identified for each student. For understanding continuous example mining concerns, the EGSP is suggested. The next sections illustrate how one approach to using the stage-wise model is to begin by discovering all the current via techniques for the EGSP strategy.

Candidate creation

The candidates utilized for the second pass are formed by merging $F_n(i-1)$, utilizing itself, given the range of recurrent $(i-1)$ and their recurrent cycles $F_n(i-1)$. By using the trimming step, at least a few of whose sub-cycles are irregular are eliminated.

Pruning phase

A pruning step eliminates at least one cycle whose series is not recurring.

TSVM for recommendation system

TSVM is often a continuous classifier that use a Transductive technique to comprise unidentified data in the training stage and routinely determine best hyperplanes in learning spaces.⁽¹⁹⁾ The improved TSVM is described here so that the learner's data may be understood. The learner's rating in the modified TSVM, that is completed in processes of dividing clustered individuals that depend on ratings from regular sequences, may be described based on optimal response percentages. The ETSVM's learning phase is defined as an issue of optimization and developed:

The user-specified training penalty values are represented by the constants C and C^* . d stands for the number of transductive samples, and the transductive samples and E_j are random variables. Improved TSVM training relates to how to approach solving the problems. The function of the improved TSVM's result with the Lagrange multipliers and α_j^* is outlined as follows.

$$f(x) = \text{sgn}[\sum_{i=1}^l y_i \alpha_i(x, x_i) + \sum_{j=1}^d y_j^* \alpha_j^*(x, x_j^* + b)] \quad (7)$$

Trust-based weighted mean item

A recommendation technique is used for estimating the standard rating for an objective item, i , by considering the rating from complete schemes' customer u , who are familiar with using i . This is the outcome of having a recommender system but not having a reliable system. The trust-based weighted mean approach is assessed since it provides the opposite level in the way that the raters' trust is intended. The trust-based weighted mean strategy is investigated by calculating the faith rates $t_{a,u}$. With the help of this technique, the sources can be divided up such that more weight is given to user reviews from more reliable sources. Where RT represent the collection of clients with faith rate $t_{a,u}$. Following that, the ranking of item i for customer a is assessed below:

$$P_{a,i} = \frac{\sum_{u \in RT} t_{a,u} r_{u,i}}{\sum_{u \in RT} t_{a,u}} \quad (8)$$

Recommendation process

The following categories describe the learner behaviour and choices:

- Clicks: describes a short set of chemicals.
- Selection: defines the material selected with additional attention to the haul.
- Learning: the material is defined as learning in this step.

RESULTS

The suggested algorithm is run on three public datasets: books, open university, and academic paper learning suggestions. The book dataset is taken from the following link <https://www.kaggle.com/datasets/dilaaslan/biggest-ecommerce-bestselling-books-dataset>. The open university dataset is taken from the following link: <https://www.kaggle.com/datasets/rocki37/open-university-learning-analytics-dataset?select=studentAssessment.csv>. Scholarly paper is taken from the following link: <https://www.kaggle.com/datasets/tayorm/arxiv-papers-metadata>.

In this section, the performance of the proposed ICSO-TSVM scheme is evaluated and compared with previous methods such as CF and TSVM algorithms. The experiments are conducted using Matlab. The existing and proposed methods are compared in terms of accuracy, precision, recall and lower MAE, time complexity.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	ISBN	Book-Title	Book-Author	Year-Of-Publication	Publisher	Image-URL-S	Image-URL-M	Image-URL-L						
2	0195153448	Classical Mythology	Mark P. O. Morford	2002	Oxford University Press	http://images.amazon.com/images/P/0195153448.								
3	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/images/P/0002005018.01.TH								
4	0060973129	Decision in Normandy	Carlo D'Este	1991	HarperPerennial	http://images.amazon.com/images/P/0060973129.01.THUMBZ								
5	0374157065	Flu: The Story of the Great Influenza Pandemic of 1918 and the Search for the Virus That Caused It	Gina Bari Kolata	1999	Far									
6	0393045218	The Mummies of Urumchi	E. J. W. Barber	1999	W. W. Norton & Company	http://images.amazon.com/images/P/0393045218.01.THUMBZ								
7	0399135782	The Kitchen God's Wife	Amy Tan	1991	Putnam Pub Group	http://images.amazon.com/images/P/0399135782.01.THUMBZ								
8	0425176428	What If?: The World's Foremost Military Historians Imagine What Might Have Been	Robert Cowley	2000	Berkley Publishing									
9	0671870432	PLEADING GUILTY	Scott Turow	1993	Audioworks	http://images.amazon.com/images/P/0671870432.01.THUMBZZZ.jpg								
10	0679425608	Under the Black Flag: The Romance and the Reality of Life Among the Pirates	David Cordingly	1996	Random House	http://								
11	074322678X	Where You'll Find Me: And Other Stories	Ann Beattie	2002	Scribner	http://images.amazon.com/images/P/074322678X.01								
12	0771074670	Nights Below Station Street	David Adams Richards	1988	Emblem Editions	http://images.amazon.com/images/P/07710746								
13	080652121X	Hitler's Secret Bankers: The Myth of Swiss Neutrality During the Holocaust	Adam Lebor	2000	Citadel Press	http://images.a								
14	0887841740	The Middle Stories	Sheila Heti	2004	House of Anansi Press	http://images.amazon.com/images/P/0887841740.01.THUMBZ								
15	1552041778	Jane Doe	R. J. Kaiser	1999	Mira Books	http://images.amazon.com/images/P/1552041778.01.THUMBZZZ.jpg								
16	1558746218	A Second Chicken Soup for the Woman's Soul (Chicken Soup for the Soul Series)	Jack Canfield	1998	Health Communications									
17	1567407781	The Witchfinder (Amos Walker Mystery Series)	Loren D. Estleman	1998	Brilliance Audio - Trade	http://images.amazon.co								

Figure 4. Dataset model (BXbooks)

Users are included as BX-Users. Notice that client IDs have been converted to whole numbers to make them anonymous. If available, statistical data (such as "Location" and "Age") is provided. Additionally, these meadows contain NULL values. BX-Books can be identified by their unique ISBN. Just now, illogical ISBNs were removed from the dataset. Additionally, a small amount of substance-based data is taken via Amazon Web Services. Data on book ratings can be found in BX-Book-Ratings. Assessments ('Book-Ratings') are either certain, conveyed through a 0, or unambiguous, communicated through a scale from 1 to 10 (increased attributes signifying elevated gratitude).

The efficiency of the suggested ICSO-TSVM system is evaluated and compared with previous approaches such as CF and TSVM algorithms. Utilizing Matlab, the tests are executed out. The accuracy, recall, precision, decreased MAE, and temporal complexity of the present and suggested approaches are evaluated.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
3	2;"stockton	california	usa";"18"											
4	3;"moscow	yukon territory	ruusia";NULL											
5	4;"porto	v.n.gaia	portugal";"17"											
6	5;"farnborough	hants	united kingdom";NULL											
7	6;"santa monica	california	usa";"61"											
8	7;"washington	dc	usa";NULL											
9	8;"timmins	ontario	canada";NULL											
10	9;"germantown	tennessee	usa";NULL											
11	10;"albacete	wisconsin	spain";"26"											
12	11;"melbourne	victoria	australia";"14"											
13	12;"fort bragg	california	usa";NULL											
14	13;"barcelona	barcelona	spain";"26"											
15	14;"mediapolis	iowa	usa";NULL											
16	15;"calgary	alberta	canada";NULL											
17	16;"albuquerque	new mexico	usa";NULL											
18	17;"chesapeake	virginia	usa";NULL											
19	18;"rio de janeiro	rio de janeiro	brazil";"25"											
20	19;"weston		;14"											
21	20;"langhorne	pennsylvania	usa";"19"											
22	21;"ferrol / spain	alabama	spain";"46"											
23	22;"erfurt	thueringen	germany";NULL											

Figure 5. Dataset model (BXUsers)

Accuracy

The total actual classification parameters $Tp+Tn$ is determined by calculating the model's accuracy and dividing it by the total of the categorization variables ($Tp+Tn+Tp+Fn$). The accuracy is determined as follows:

$$\text{Accuracy} = \frac{Tp+Tn}{(Tp+Tn+Tp+Fn)} \quad (9)$$

Here Tp is true positive, Tn is true negative, Fp is false positive and Fn is false negative.

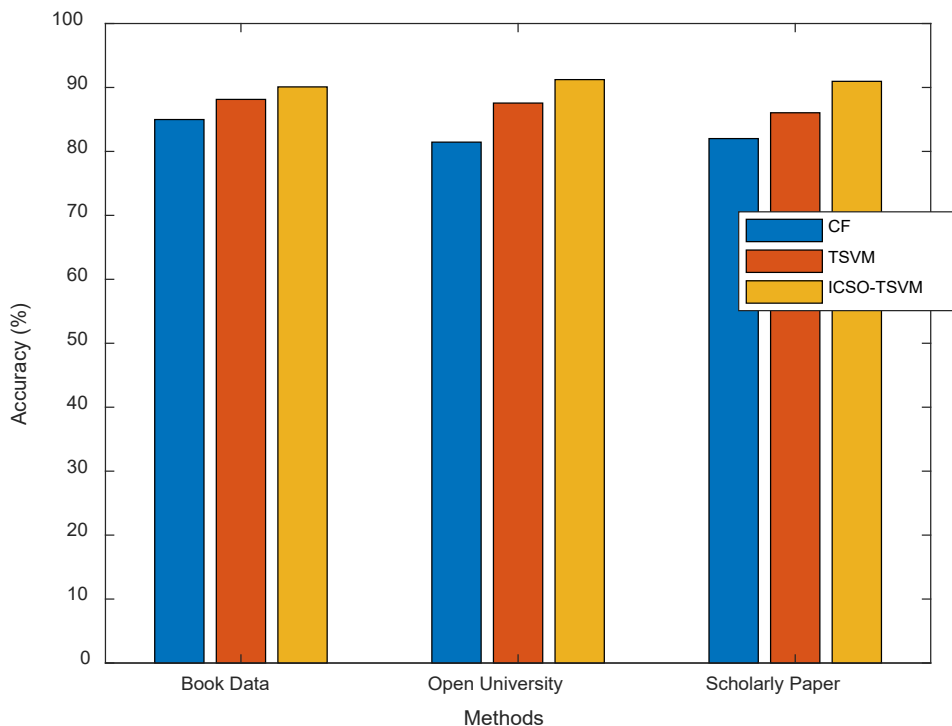


Figure 6. Accuracy

As seen in the preceding figure 6, a variety of cutting-edge and established techniques are used to assess the comparison metric's correctness. Methods form the x-axis while accuracy values are plotted on the y-axis. In comparison to existing techniques like CF and TSVM techniques, the ICSO-TSVM algorithm delivers enhanced accuracy for the presented recommendation datasets. Using the KMC algorithm, the pre-processing is used to enhance the accuracy of the recommendations. The ICSO algorithm improves the important features for the better results. The TSVM is utilized to compute the learners' data. As an outcome, the findings indicate that the stated ICSO-TSVM method improves the efficiency of the recommendation datasets using an optimized method.

Precision

The calculation for the precision is as followed:

$$\text{Precision} = \frac{T_p}{T_p + F_p} \quad (10)$$

Precisions are relevant documents found by searches divided by total counts of documents recovered by searches, whereas recall an ere relevant document counts found by searches divided by existing relevant document counts.

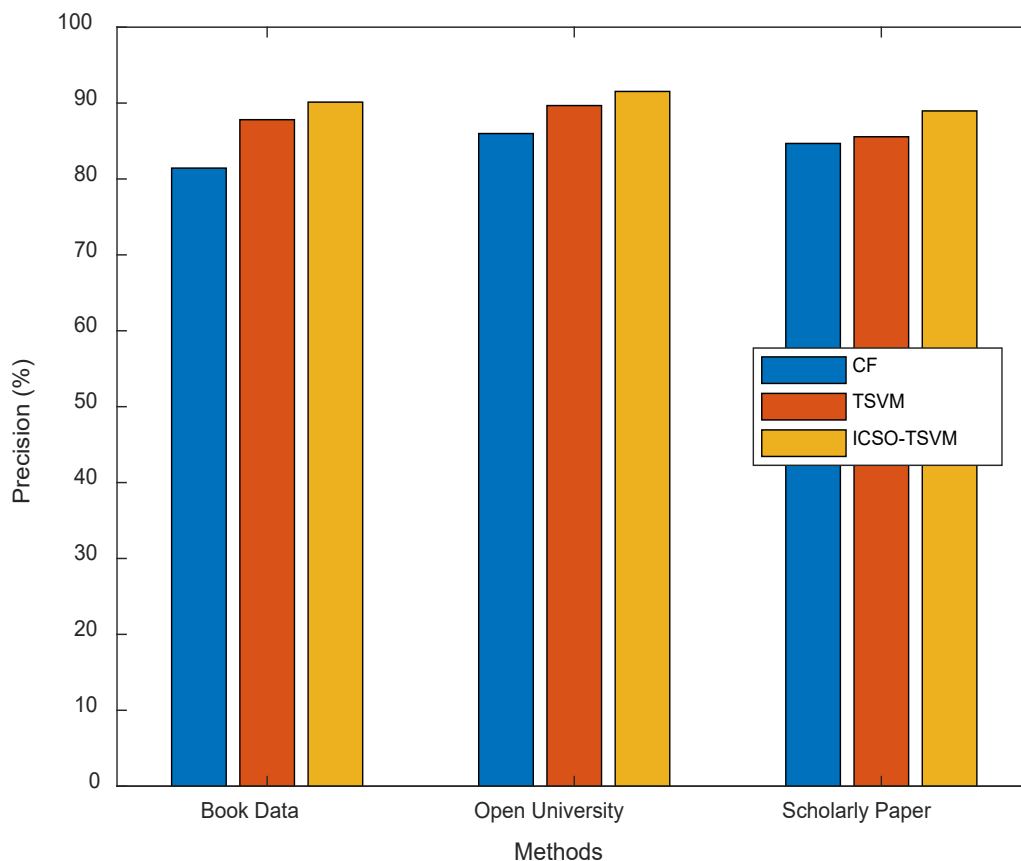


Figure 7. Precision

As seen in figure 7, Utilizing currently available methods, the precision of the opposing metric is evaluated. Methods form the x-axis while precision values are plotted on the y-axis. The proposed ICSO-TSVM algorithm provides higher precision whereas existing CF and TSVM methods provide lower precision. The learners' grade may clarify according to the gain from the excellent answer, and from there, relying on the repeated rating cycle, the clustered customers may be divided. Thus, the result concludes that the suggested ICSO-TSVM approach improves the informative features accurately for recommendation process.

Recall

The recall value is considered:

$$\text{Recall} = \frac{T_p}{T_p + F_n} \quad (11)$$

The comparison graph indicates as follows:

Recall measures completeness or quantity, while precision can be defined as a calculation of accuracy. High precision indicates that a procedure produced more relevant outcomes than unrelated ones. The ratio of true positives to all items labeled as belonging to the positive class determines the precision for a class in a clustering activity.

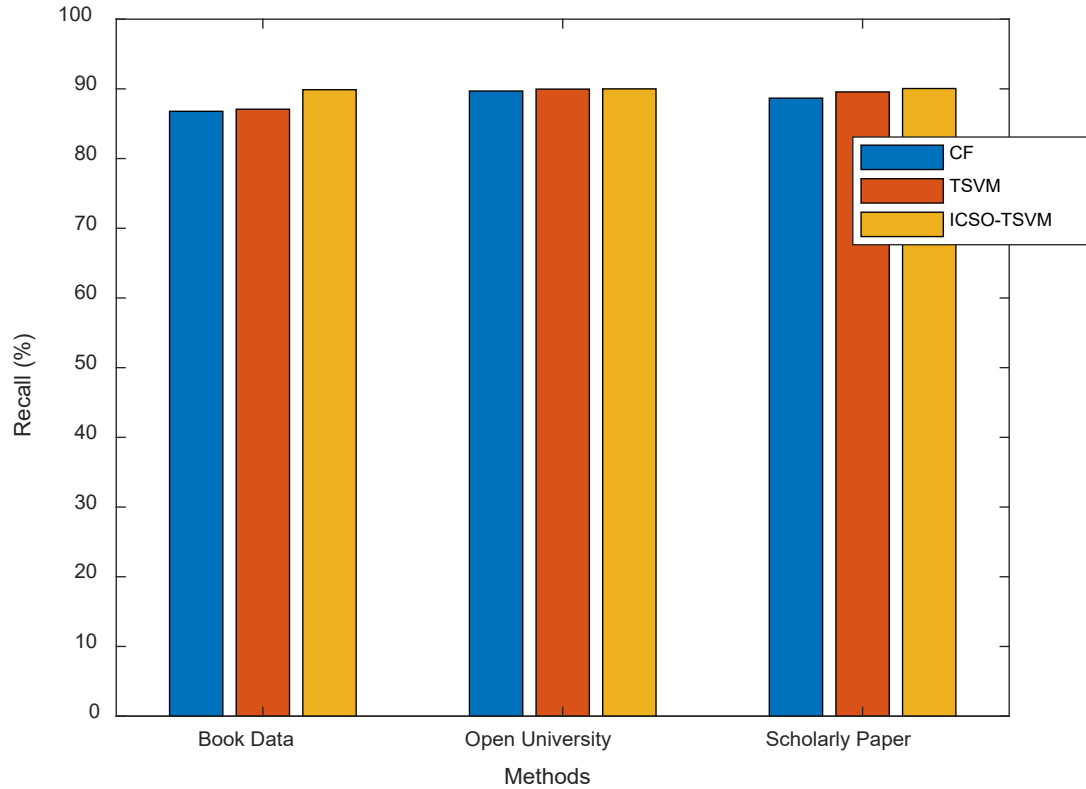


Figure 8. Recall

Present methods are used to evaluate the recall of the contrast metric, as shown in figure 8. Methods form the x-axis while recall values are plotted on the y-axis. The proposed ICSO-TSVM algorithm provides higher recall whereas existing CF and TSVM methods provide lower recall. The more pupils who use the suggested approach, the more accurate it develops. The conclusion is that the ICSO-TSVM method boosts the accuracy of the recommendation findings.

MAE Analysis

The MAE is calculated by summing together all the N's errors, which are proportional to evaluating forecast matches and then determining normal value. The MAE measure is calculated as follows for each combination p_i and q_i of anticipated assessments p_i and actual evaluations q_i :

$$MAE = \frac{\sum_{i=1}^N |p_i - q_i|}{N} \quad (12)$$

The contrast metric is evaluated using the most recent methods, as shown in figure 9. Methods form the x-axis while MAE values are plotted on the y-axis. While the current CF and TSVM approach offers greater MAE, the suggested ICSO-TSVM approach offers a lower MAE. The precision will be greater with a low MAE. The planned technique outperformed the current approaches in terms of output. It has been discovered that the mean absolute error lowers as the number of students rises. The findings indicate that the provided ICSO-TSVM technique greatly boosts the effectiveness of the recommendation system.

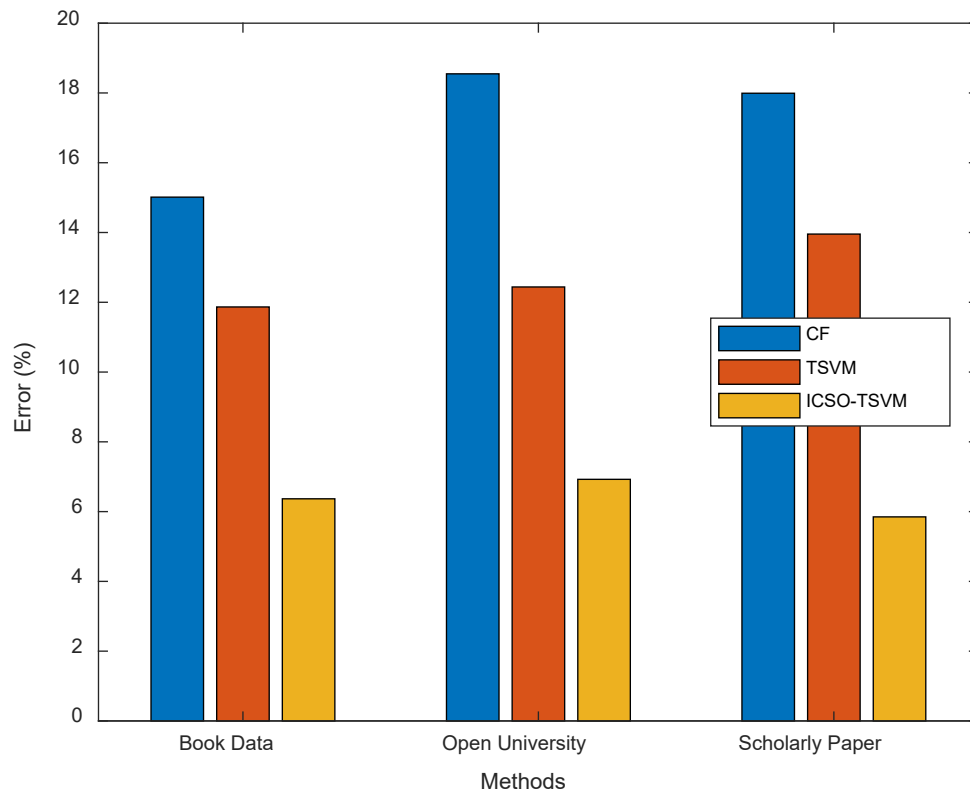


Figure 9. MAE

Computation time

The algorithm is better when it provides lower time complexity.

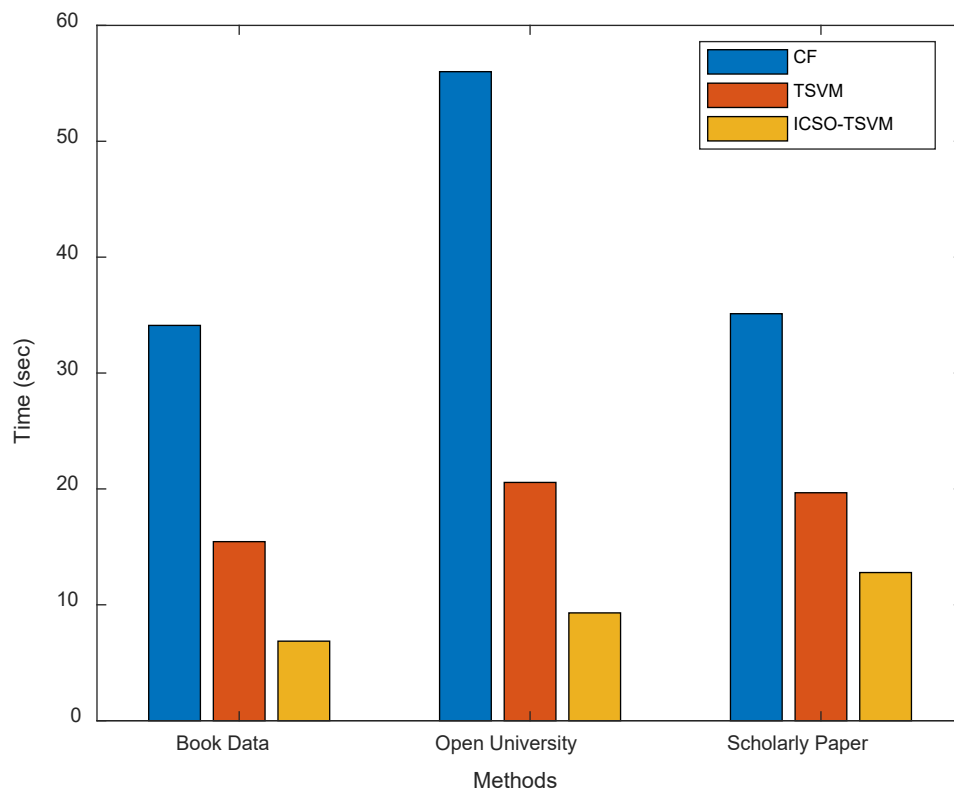


Figure 10. Execution time

As shown in figure 10, the comparative measure is evaluated based on execution time, using both the existing and suggested approaches. The methods are displayed on the x-axis, while the execution time is displayed on the y-axis. For the provided recommendation dataset, the suggested ICSO-TSVM algorithm executes more quickly than the current methods, such as CF and TSVM algorithms. The findings suggest that the ICSO-TSVM increases recommendation accuracy by clustering learner behaviours.

CONCLUSIONS

In this study, ICSO-TSVM algorithm is suggested to enhance the recommendation system performance for e-learning. There are four key modules in this research, including pre-processing, clustering, optimization and e-learning recommendation. Pre-processing is focused to improve the dataset quality by handling the missing values. The clustering is done by using DBSCAN algorithm which provides similar learner's into one group efficiently. Then the ICSO algorithm provides the useful and relevant feature for the given datasets. It is used to fine tune the parameters. After that, mine the sequential pattern of learners by pruning phase. Finally, the classification is done by using TSVM algorithm which provides more accurate recommendation performance. The proposed ICSO-TSVM model helps to improve the recommendation performance for e-learners in better way. According to the learner's habits and concerns, TSVM determines their evaluation. Experimental findings indicate that the recommended ICSO-TSVM method exceeds current methods according to accuracy, precision, recall, and time complexity. In future, deep learning-based algorithms will be implemented on the given datasets. Hybrid optimization algorithms also can be proposed to progress the recommendation performance using their local and global optimal features

BIBLIOGRAPHIC REFERENCES

1. Shi D, Wang T, Xing H, Xu H. A learning path recommendation model based on a multidimensional knowledge graph framework for e-learning. *Knowledge-Based Systems*. 195, pp. 105618. <https://doi.org/10.1016/j.knosys.2020.105618>
2. Chen W, Niu Z, Zhao X, Li Y. A hybrid recommendation algorithm adapted in e-learning environments. *World Wide Web*. 17, pp. 271-84. <https://doi.org/10.1007/s11280-012-0187-z>
3. Halawa MS, Hamed EM, Shehab ME. Personalized E-learning recommendation model based on psychological type and learning style models. In 2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS), pp. 578-584. <https://doi.org/10.1109/IntelCIS.2015.7397281>
4. Furtado F, Singh A. Movie recommendation system using machine learning. *International journal of research in industrial engineering*. 9(1), pp. 84-98. <https://doi.org/10.22105/riej.2020.226178.1128>
5. Choudhury SS, Mohanty SN, Jagadev AK. Multimodal trust based recommender system with machine learning approaches for movie recommendation. *International Journal of Information Technology*. 13. Pp. 475-82. <https://doi.org/10.1007/s41870-020-00553-2>
6. Kiran R, Kumar P, Bhasker B. DNNRec: A novel deep learning based hybrid recommender system. *Expert Systems with Applications*. 144, pp. 113054. <https://doi.org/10.1016/j.eswa.2019.113054>
7. Hussain M, Zhu W, Zhang W, Abidi SM. Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores. *Computational intelligence and neuroscience*. 2018(1), pp. 6347186. <https://doi.org/10.1155/2018/6347186>
8. Moubayed A, Injadat M, Nassif AB, Lutfiyya H, Shami A. E-learning: Challenges and research opportunities using machine learning & data analytics. *IEEE Access*, 6, pp. 39117-38. <https://doi.org/10.1109/ACCESS.2018.2851790>
9. Wang Z, Yu X, Feng N, Wang Z. An improved collaborative movie recommendation system using computational intelligence. *Journal of Visual Languages & Computing*. 25(6), pp. 667-75. <https://doi.org/10.1016/j.jvlc.2014.09.011>
10. Guo G, Zhang J, Yorke-Smith N. Leveraging multiviews of trust and similarity to enhance clustering-based recommender systems. *Knowledge-Based Systems*. 74, pp. 14-27. <https://doi.org/10.1016/j.knosys.2014.10.016>
11. Li J, Ye Z. Course recommendations in online education based on collaborative filtering recommendation algorithm. *Complexity*. 2020(1), pp. 6619249. <https://doi.org/10.1155/2020/6619249>

12. Intayoad W, Kamyod C, Temdee P. Reinforcement learning based on contextual bandits for personalized online learning recommendation systems. *Wireless Personal Communications*. 115(4), pp. 2917-32. <https://doi.org/10.1007/s11277-020-07199-0>
13. Aziz MA, Hassanien AE. Modified cuckoo search algorithm with rough sets for feature selection. *Neural Computing and Applications*. 29, pp. 925-34. <https://doi.org/10.1007/s00521-016-2473-7>
14. Bhaskaran S, Marappan R. Design and analysis of an efficient machine learning based hybrid recommendation system with enhanced density-based spatial clustering for digital e-learning applications. *Complex & Intelligent Systems*. 9(4), pp. 3517-33. <https://doi.org/10.1007/s40747-021-00509-4>
15. Mohamad IB, Usman D. Research article standardization and its effects on k-means clustering algorithm. *Res J Appl Sci Eng Technol*. 6(17), pp. 3299-303. <https://doi.org/10.19026/rjaset.6.3638>
16. Bagunaid W, Chilamkurti N, Veeraraghavan P. Aisar: Artificial intelligence-based student assessment and recommendation system for e-learning in big data. *Sustainability*. 14(17), pp. 10551. <https://doi.org/10.3390/su141710551>
17. Huang L, Ding S, Yu S, Wang J, Lu K. Chaos-enhanced Cuckoo search optimization algorithms for global optimization. *Applied Mathematical Modelling*. 40(5-6), pp. 3860-75. <https://doi.org/10.1016/j.apm.2015.10.052>
18. Kamoon AM, Patra JC, Stojcevski A. An enhanced cuckoo search algorithm for solving optimization problems. In 2018 IEEE congress on evolutionary computation (CEC), pp. 1-6. <https://doi.org/10.1109/CEC.2018.8477784>
19. Chen H, Yu Y, Jia Y, Gu B. Incremental learning for transductive support vector machine. *Pattern Recognition*. 133, pp. 108982. <https://doi.org/10.1016/j.patcog.2022.108982>

FINANCING

The authors did not receive financing for the development of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: D Poornima.

Data curation: D Poornima.

Formal analysis: D Poornima.

Research: D Karthika.

Methodology: D Karthika.

Drafting - original draft: D Poornima.

Writing - proofreading and editing: D Karthika.