IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

The role of MAS based CBRS using DM Techniques for the supplier selection

Jagjit Singh Dhatterwal[1*], Sarvottam Dixit[2], S. Srinivasan [3]

[1] Research Scholar, Department of Computer Science, Mewar University, Gangrar, Chittorgarh, Rajasthan, India
[2] Department of Computer Science, Mewar University, Gangrar, Chittorgarh, Rajasthan, India
[3] Department of Computer Science & Applications, PDM University, Bahadurgarh, Haryana, India

[*]*Corresponding Author: dss_dce@yahoo.com, Tel.: +919313998006*

DOI: https://doi.org/10.26438/ijcse/v7i5.16581665 | Available online at: www.ijcseonline.org

Accepted: 24/May/2019, Published: 31/May/2019

***Abstract*—**The supplier selection process is an important activity in any organization whether it is a manufacturing unit or educational institution to procure products or raw materials with the right price, strictly following the delivery schedule and the right quantities ordered. Anything that goes wrong in the form of delay in delivery or shortage of items ordered would lead to a chaotic situation in the organization. It often happens in the case of systems that are not properly integrated. The aim of this paper is to introduce a particular way of selecting a supplier by using Multi-agent system (MAS) based Case Based Reasoning system (CBRS) using Data mining (DM) techniques. The method that is suggested in this paper is to combines MAS, CBRS and DM for supplier selection. Finally this integrated model MAS-CBRS-DM is illustrated by an example in a seller selection process.

***Keywords*—** Multi-agent system, Case base reasoning system, Data mining

I. INTRODUCTION

A case-based reasoning (CBR) is one of the AI techniques in solving a problem by matching the problem description to a previously solved case, using the case base in which solved cases are already stored [1]. After each problem solving session you learn from case-based reasoning by retaining relevant information from a newly resolved problem for future problem solving. An important step in a CBR process involves finding a good match for a new problem and adapting an earlier solution in order to solve the new problem, index and store a new case to be retrieved later.

Given a new problem, a person often recollects a previous problem similar to the one at hand. For example: A Librarian after having examined several book sellers gets a reminding to a seller that he purchased a few months ago. If the reminding was caused by a cost of a book he intends to purchase, discount he offers and delivery, the librarian uses the list of book sellers of the purchases to determine the cost, discount and delivery of the purposed purchase. The example above illustrates how reasoning by re-using or revising past experiences is a powerful and frequently applied technique for human problem solving. This claim has supports from results of cognitive psychological research. Analogy based learning is closely related to case based reasoning, but the study issues differ in the two disciplines since a main research issue in analogy by leaning is the mapping of a current problem description to a known problem in a particular domain while case-based methods are used to index and match domain-specific strategy cases [14].

The main principle behind case-based methods in AI is based on the methodology for machine-learning and case-based reasoning (CBR) is a machine-learning subfield. CBR teaches how to solve problems. The CBR system comprises a number of methods used to organize, index, retrieve, revise and use past experience. The solution obtained from the past case can be matched directly with the current problem or revised based on differences between the two cases [6] . A comprehensiveness of the domain knowledge is required to produce sufficient arguments for why two cases match and how close this match is to find and match cases based on semantical similarities and other related features. The fig.1 depicts how a case is retrieved, matched, reused or revised and retained [12]. This paper is comprised of six section in which Section-I contain introduction of case based reasoning system (CBRS). Section- II contains elaborate discussion on CBRS with as an associated open source software jColibri. Section- III introduce multi-agent system MAS and the open source software JADE. Section- IV the integration of MAS and CBR alongwith its framework is discussed. Section- V

introduce Data mining and its associated software tool WEKA. Section- VI all the three software have been integrated and with a sample of data containing book. Sellers with cross ponding discount and delivery time, it is shown how a seller is selection who can give good discount and delivery time.

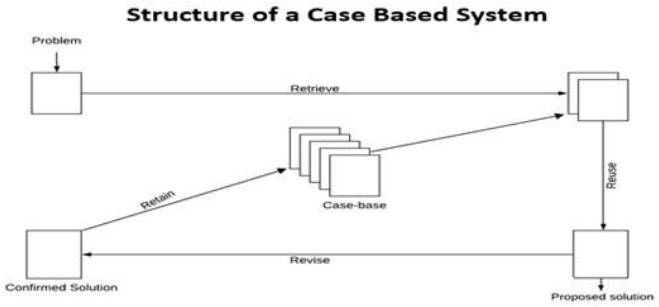


Figure. 1 Case base reasoning system architecture

II. CBR AND JCOLIBRI

The paper uses an open source jCOLIBRI, an object-oriented Java-based framework for building CBR systems and aims to determine how powerful it is for any given application. jCOLIBRI's design is based on a broad spectrum framework that supports several CBR systems, from a simple nearby method to a complex, knowledge-intensive structure. JCOLIBRI provides a process of development that is less complex and is dependent on reuse of previous designs. It offers a clear structure for CBR knowledge, a description of the CBR tasks at the knowledge level and a library of recurring problem solving methods (PSMs). This CBR knowledge is directed towards classes within the framework. [4]. The graphical interface supports the instantiation of the framework and guides the setup of an individual CBR system. Among others, we are using jCOLIBRI to build the Purchase Domain Buyer Seller System, a popular CBR-community example.

III. MULTI-AGENT SYSTEM

An agent is an entity often a computer system, able to act in an environment, interact/communicate with other agents, perceive its environment only in part, have specific skills and services. Agents are Autonomous that are able to perform its functionality independently and tries to achieve its goals autonomously. They are Intelligent; it means the agents are designed, trained, or fitted for one particular knowledge in one or more application fields [5]. The agents can collect information or are reactive on conditions of its environment. They are generally reactive so they react depending on the inputs from its environment. The agents are Pro-activity and goal orientated. The agents are bound to change its behavior based on its earlier experience. They are mobile, it means that agents to move themselves from one node to another in he given network. The important features are that agents are generally communicative and cooperative. Multiagent system (MAS) consists of agents in which several agents interact with each other to solve problems [10]. Agents in MAS are designed in such a way that they know when and how to interact with which agents. The main features of agents in multiagent systems are their inherent distribution and complexity. The nature of multiagent systems is that agents are distributed and flexible and so generally MAS performs in increased speed, robustness, scalability and reusability. Agents have only partial information and their Control and data are decentralized.

A JAVE AGENT FRAMEWORK (JADE)

The java-based MAS.JADE is a feature that supports technological development, a middleware for developing and run-time peer-to-peer applications based on a paradigm of agents that can work seamlessly and interact wirelessly. FIPA Application programmer's interface, FIPA Interaction Protocols Library, such as Contract Net and Graphical User Interface for managing several agents from the same Remote Management Agent are just some of the functionalities of the distributer platform that can be spaced to several machines. Fig. 3 depicts its internal architecture. This architecture includes agents which possesses unique name, residing in container. The collection of containers is called a platform. It has a unique container called main container. The remaining containers must register their names in the main container [9]. Jade includes the main packages available in JADE. Jade.proto, jade.wrapper, jade.gui, jade.lang.acl, jade.content, jade.domain, jade.gui.

IV. INTEGRATION OF CBRS AND MAS

The main conclusions of the study may be presented in a short Conclusion Section. In this section, the author(s) should also briefly discuss the limitations of the research and Future Scope for improvement. This research paper presents some of the critical points to be faced at different levels of integration of CBR techniques with Multi-agent system components and techniques for building up a full-fledged Buyer – Seller negotiation system. The integration requirement at the level of system components is identified from the integration requirements at the level of specific stages in the CBR problem solving cycle. In this paper it is described how these aspects have been dealt inside jCOLIBRI and JADE [2]. JADE for MAS and jCOLIBRI for CBR system for Buyer – Seller negotiation system are considered for the discussion on how MAS and CBR system are integrated .The framework shown in the fig. 4 illustrates how the integration takes place in this scenario [10].

In the framework of Buyer – Seller negotiation system, it is shown how the integration of JADE and jCOLBIRI works [2]. The buyer agent uses Retrieving agent to pick up information from casebase of jCOLBIRI . Afterwards the negotiation begins with seller agent for the purchase of a book [8].

V. Data mining

Data mining is the central process where a number of complex and intelligent methods are applied to extract patterns from data. Data mining process includes a number of tasks such a) Classification, b) Prediction,c) Time - Series Analysis, Association, e) Clustering,f) Summarization.

A Clustering

- It is a data mining technique which is used to place the data elements into their related groups.
- By this process the data in one class is more similar to each other than to those in other cluster.
- The process of grouping data into subclasses is called as cluster.
- A cluster consists of data object with high inter similarity and low intra similarity.
- The method used will provide quality clusters.
- It is also called as data segmentation, because it groups large data into groups according to their similarity

Clustering methods can be classified into the following categories:

1.Partitioning In this approach, data is partitioned and then evaluation is done based on given criteria.
2. Hierarchical method In this method, the data is decomposed (multilevel) hierarchically by using some criteria.
3. Density-based method This method is based on density (density reachability and density connectivity).

B Grid-based methods

This approach is based on multi-resolution grid data structure

C About WEKA

Weka is an open source software having many machine learning algorithms for various tasks of data mining . **Weka** contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization.

D. Integrating MAS and CBR

It is proposed integrating MAS and CBR for the selection of supplier of books for a University Library . The framework suggests how a retrieving agent picks of similar cases [3]. Suppose the librarian here after called a buyer wants to buy a book with a title t and the number of copies may be thousands . So he looks for a decent discount and at the same time the buyer wants them to be delivered as early as possible . Let the delivery days be d and the discount be e . The following is the set of similar cases : { (s_i,d_i,e_i) } , here s_i is a seller , d_i is the delivery days and e_i the discount . In this case 50 similar cases are fetched . Various tools with regard to clustering that are available in WEKA, are applied . The table 1 contains 50 such cases.

The following framework describes how the retrieving agent matching the best cases in huge number of cases in jColibri where: t refer time, d refer delivery, e refer days.

.

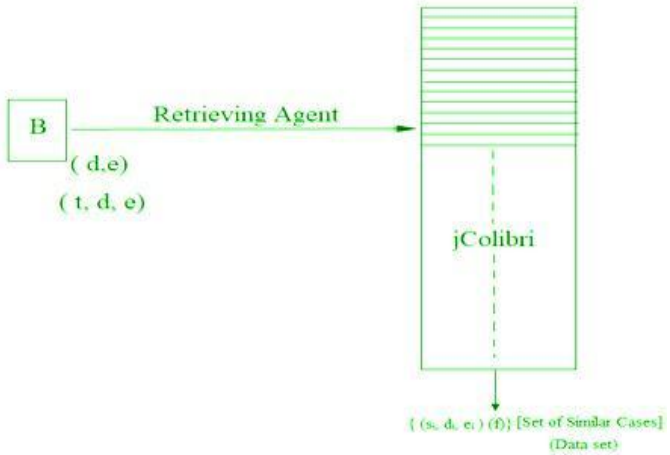


Figure. 2 Buyer and Seller Framework

The table1 contains 50 sellers (s1,s50) and every seller provides information on how many days a book can be delivered and corresponding discount .

Table. 1 Buyer and Seller

Relation: Buysel2				Relation: Buysel2			
No.	1: Seller Nominal	2: Days Numeric	3: **Discount** Numeric	No.	1: Seller Nominal	2: Days Numeric	3: **Discount** Numeric
1	s1	2.0	3.0	26	s26	16.0	8.7
2	s2	3.0	4.0	27	s27	16.0	8.8
3	s3	5.0	4.2	28	s28	16.0	8.9
4	s4	6.0	5.0	29	s29	16.0	9.0
5	s5	7.0	5.2	30	s30	16.0	10.0
6	s6	8.0	5.3	31	s31	16.0	11.1
7	s7	9.0	5.4	32	s32	17.0	11.2
8	s8	9.0	5.6	33	s33	17.0	11.3
9	s9	9.0	5.8	34	s34	17.0	11.4
10	s10	10.0	5.9	35	s35	17.0	11.5
11	s11	10.0	6.0	36	s36	17.0	11.7
12	s12	10.0	6.1	37	s37	17.0	11.8
13	s13	10.0	6.2	38	s38	17.0	12.0
14	s14	11.0	6.3	39	s39	20.0	13.0
15	s15	12.0	6.4	40	s40	21.0	14.0
16	s16	13.0	6.6	41	s41	22.0	15.0
17	s17	14.0	6.7	42	s42	23.0	16.0
18	s18	14.0	6.8	43	s43	24.0	17.0
19	s19	14.0	6.9	44	s44	26.0	17.2
20	s20	14.0	7.0	45	s45	27.0	17.5
21	s21	14.0	8.0	46	s46	27.0	17.8
22	s22	15.0	8.1	47	s47	27.0	17.9
23	s23	16.0	8.2	48	s48	28.0	18.0
24	s24	16.0	8.4	49	s49	29.0	19.0
25	s25	16.0	8.5	50	s50	29.0	20.0

E. EM algorithm

One of the tools is EM algorithm . An expectation–maximization (**EM**) **algorithm** is an iterative method to find maximum likelihood and it is shown in the flowchart1 . Expectation–Maximization (EM) flowchart search and select the best pattern after giving data. It depends on the strength of data to work accordingly. But this flowchart more reliable for select relevant data (suppliers) in this context.

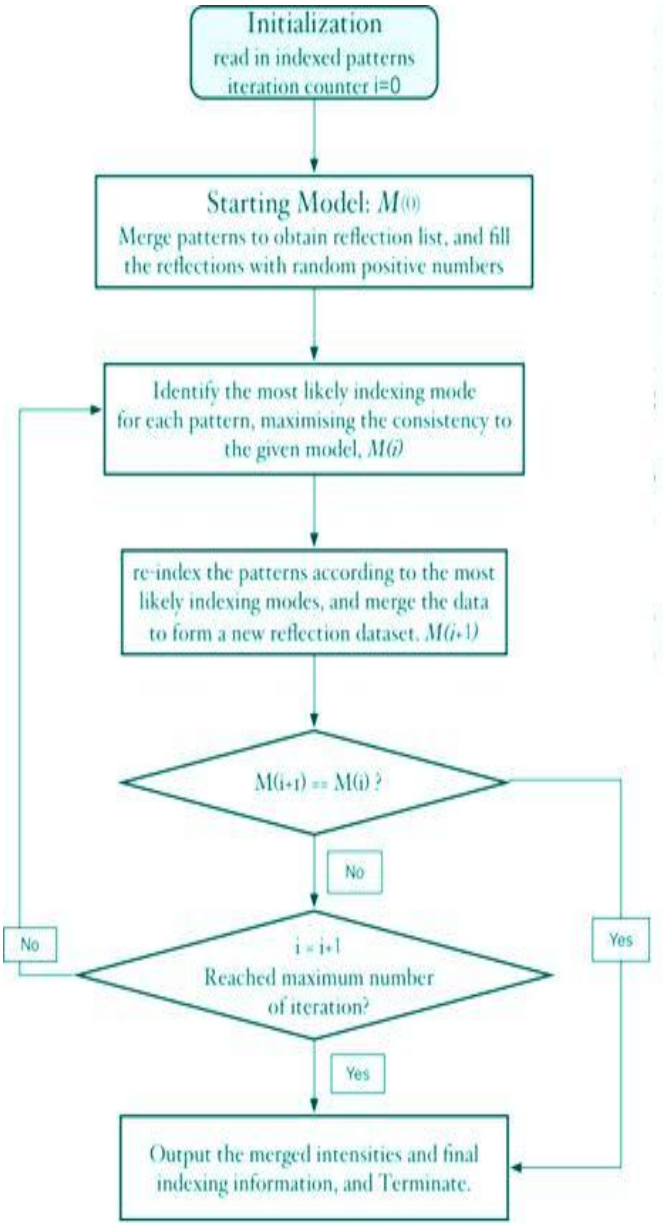


Figure. 3 Flowchart 1

EM algorithm use on 50 data find 6 best pair of sellers . Second time the EM algorithm is applied on 6 data to get the seller s_{31} with delivery days 16 and discount 11 ie. (s_{31},16,11). The Result1 shows how table reduced to table of 6 data and finally leading the selection of a seller .

```
  s41              2       1       1       1       1       1
  s42              2       1       1       1       1       1
  s43         1.9846       1       1  1.0154       1       1
  s44         1.0054       1       1  1.9946       1       1
  s45         1.0001       1       1  1.9999       1       1
  s46         1.0001       1       1  1.9999       1       1
  s47         1.0001       1       1  1.9999       1       1
  s48              1       1       1       2       1       1
  s49              1       1       1       2       1       1
  s50              1       1       1       2       1       1
  [total]    54.9902 62.0503 57.9852 57.0098 53.0632 64.9012
Days
  mean       21.9983  9.3046 16.8766 27.5648  3.3887 15.0815
  std. dev.   1.4178  1.6128  0.3289  1.0614  1.2919  1.0551

Discount
  mean       14.9963   5.779 11.5007 18.1981  3.7596  8.0537
  std. dev.   1.4132  0.4459  0.2915  0.8999    0.55  0.9897
```

Figure. 4 Result 1

Again 6 best seller data insert in EM algorithm then make table shown in the table 2.

Figure. 5 Table 2

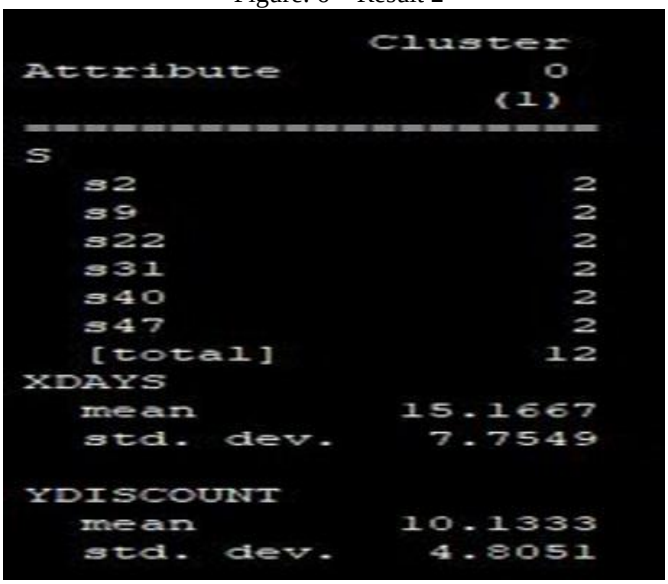
Viewer

Relation: Buysel1

No.	1: S	2: XDAYS	3: YDISCOUNT
	Nominal	Numeric	Numeric
1	s2	3.0	4.0
2	s9	9.0	5.8
3	s22	15.0	8.0
4	s31	16.0	11.0
5	s40	21.0	14.0
6	s47	27.0	18.0

Only one best seller comes out seller s31 result 2 shows how to select best seller.

Figure. 6 Result 2



Attribute	Cluster 0 (1)
S	
s2	2
s9	2
s22	2
s31	2
s40	2
s47	2
[total]	12
XDAYS	
mean	15.1667
std. dev.	7.7549
YDISCOUNT	
mean	10.1333
std. dev.	4.8051

F. Simple k-means Clusters

Simple k-means clustering is an algorithm to classify or to group the objects based on attributes features into k number of groups. Here K is positive integer. The grouping is done by minimizing the sum of squares of distance between data and the corresponding clusters centroid. Simple k-means flowchart mend for perspective goal and is given in Flowchart 2. It works on the distinct centroids its best suited for supplier.

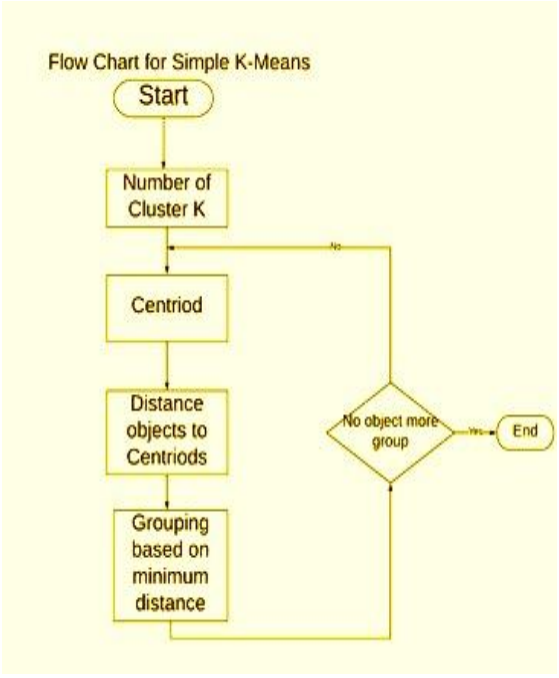


Figure. 7 Flowchart 2

Steps:

- Number of clusters must be known to be k
- k number of cluster centers such that they are farthest apart from each other
- Consider each data point and assign it to the cluster which is closest
- Recalculate cluster centers by finding mean of data points belonging to the same
- Repeat step 3 and 4 till shifting cluster centers are observed

It is observed that s38 with delivery days 17 and 12 % discount seem to be working solution and it is exhibited in Result 3.

```
kMeans
======

Number of iterations: 4
Within cluster sum of squared errors: 50.37236782005127

Initial starting points (random):

Cluster 0: s14,11,6.3
Cluster 1: s37,17,11.8

Missing values globally replaced with mean/mode

Final cluster centroids:
                          Cluster#
Attribute    Full Data          0          1
                (50.0)     (29.0)     (21.0)
=============================================
Seller             s1         s1        s23
Days            15.7    11.4138     21.619
Discount       9.908     6.6483    14.4095
```

Figure. 8 Result 3

G. Density-Based Clustering

Clustering based on density (local cluster criterion) such as density connected points or based on an an explicitly constructed density function.

1. *Density-based Algorithm*

- Density= number of points within a specified radius r
- core point= A point is a core point if it has more than a specified number of point (Min pts) within radius
  These are points that are at the interior of a cluster
- border point= a border point has fewer than Min pts with radius, but is in the neighbourhood Of a core point

- A noise point is any point that is not a core point or a border point

In this below flowchart 3 it is shown how Density based algorithm works and it is similar result to simple k-means but its working is totally different to simple k-means it works on a lot of separate types of points and analysis. The Result 4 provides 2 sellers s_{14} and s_{38}.

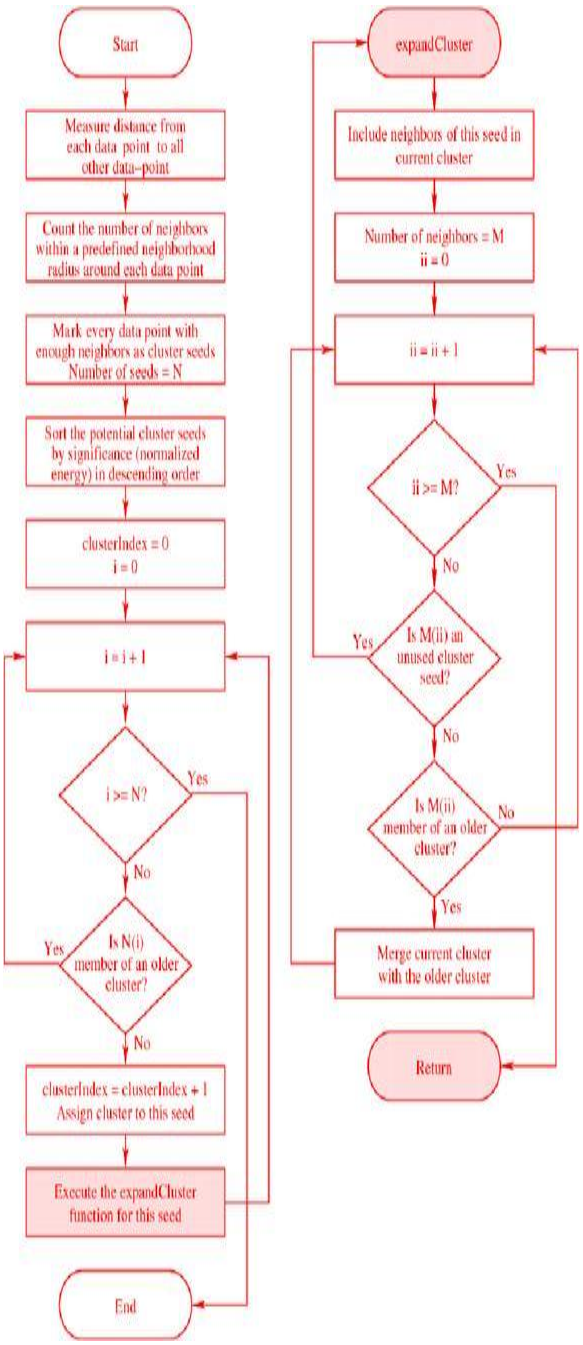


Figure. 9 Flowchart 3

```
MakeDensityBasedClusterer:

Wrapped clusterer:
kMeans
======

Number of iterations: 4
Within cluster sum of squared errors: 50.37236782005127

Initial starting points (random):

Cluster 0: s14,11,6.3
Cluster 1: s37,17,11.8

Missing values globally replaced with mean/mode

Final cluster centroids:
                        Cluster#
Attribute    Full Data          0          1
                (50.0)     (29.0)     (21.0)
=======================================
Seller              s1         s1        s23
Days              15.7    11.4138    21.619
Discount         9.908     6.6483   14.4095
```

Figure. 10 Result 4

H. FarthestFirst Algorithm

- Choose an arbitrary node I as a starting node.
- Find the farthest node from I, the node j for which distance Cij is longest. From a subtour T=i-j-i.
- Find a node k not in the subtour that is farthest from any of the subtour nodes.
- Find an edge [I,j] of the subtour to which the insertion of k gives the smallest increase of length, for which if =Cik+Ckj-Cij is smallest.
  Modify the subtour by inserting k between I and j.
- Go to 3 until a Hamiltonian cycle is formed.

It generate two clusters (0,1) in below table and comparing between two clusters come out best match of supplier s_{36}


```
FarthestFirst
==============

Cluster centroids:

Cluster 0
        s36 17.0 11.7
Cluster 1
        s1 2.0 3.0



Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       36 ( 72%)
1       14 ( 28%)
```

Figure.11 Result 5

The seller selection table 3 contains the sellers selected by the algorithms that are considered . The last column in the selection table provides the list of selected sellers . Here s31 , s36 and s38 are selected for further negotiation in order to get the desired delivery days and discount by the buyer .

Figure. 12 Table 3 Final Buyer and Seller Selection

Algorithm	Seller Selected (S_i, $Days_j$, Dis_m)	Repeated Algorithm (S_i, $Days_j$, Dis_m)	Seller Final Selection
EM	(S_2, 3, 4), (S_9, 9, 5.8), (S_{22}, 15, 8), (S_{31}, 16, 11), (S_{40}, 21, 14), (S_{47}, 27, 18)	(S_{31}, 16, 11)	(S_{31}, 16, 11)
KM	(S_{15}, 11.41, 6.64), (S_{40}, 21.6, 14.4)	(S_{38}, 17, 12)	(S_{38}, 17, 12)
DB	(S_{15}, 11.41, 6.64), (S_{40}, 21.6, 14.4)	(S_{38}, 17, 12)	(S_{38}, 17, 12)
FAR	(S_{36}, 17, 11.7), (S_1, 2, 3)	-	(S_{36}, 17, 11.7),

VI Conclusion

The CBR has a mechanism for rationalizing preceding situations or difficulties with new situations. It is seen in various problem-solving tasks and applications as a successful tool. However, with poor case-by-case and reasoning engine quality, the problems with reliability, inaccuracy and the extent of a limited application remain limited. A multi-agent system improves the CBR's ability to address complex problems with precision and accuracy. The first approach is to distribute agents through all phases of the CBR cycle and to apply the two different approaches. The agents work together to improve the four steps of the CBR cycle. Each agent develops in the second approach the ability to find similar cases to achieve its objective. This paper has studied varieties of techniques and methods by combining MAS and CBR to solve difficult issues in different domains. This paper contains an overview of several CBR integration methods with MAS in several areas. There are many functionalities of agents, like scalability, adaptation, reactivity, proactive, collaborative, reusability, distribution, and autonomy. The CBR, on the other hand, has several features, such as direct representation and extensive of knowledge. In order to integrate both CBR and MAS, some of the exceptions to the CBR and MAS can be overcome, which usually involve unreliability, inaccuracy, limited application and poor administration. The retrieval mechanism also contributes to improving the quality of the solutions obtained. For any specific application like supplier selection Data mining (DM) techniques has been employed. For future work, the MAS-CBR-DM model may be used in different domain construction applications.

REFERENCES

[1] Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI communications, 7(1), 39-59.R. Solanki, "*Principle of Data Mining*", McGraw-Hill Publication, **India**, pp. 386-398, 1998.

[2] Abutair, H. Y., & Belghith, A. (2017 a). A Multi-Agent Case-Based Reasoning Architecture for Phishing Detection. Procedia Computer Science, 110, 492-497.

[3] Ayzenshtadt, V., Langenhan, C., Bukhari, S. S., Althoff, K. D., Petzold, F., & Dengel, A. (2016, February). Thinking With Containers: A Multi-Agent Retrieval Approach for the Case-Based Semantic Search of Architectural Designs. In ICAART (1) (pp. 149-156).

[4] Berman, A. F., Maltugueva, G. S., & Yurin, A. Y. (2018). Application of case-based reasoning and multi-criteria decision-making methods for material selection in petrochemistry. Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications, 232(3), 204-212.

[5] Byrski, A., Dreżewski, R., Siwik, L., & Kisiel-Dorohinicki, M. (2015). Evolutionary multi-agent systems. The Knowledge Engineering Review, 30(2), 171-186.

[6] De Mantaras, R. L., McSherry, D., Bridge, D., Leake, D., Smyth, B., Craw, S., ... & Keane, M. (2005). Retrieval, reuse, revision and retention in case-based reasoning. The Knowledge Engineering Review, 20(3), 215-240.

[7] Eibe Frank, Mark A. Hall and Ian H. witten. The WEKA Workbench. Fourth Edition, 2016

[8] Haque, B. U., Belecheanu, R. A., Barson, R. J., & Pawar, K. S. (2000). Towards the application of case based reasoning to decision-making in concurrent product development (concurrent engineering). Knowledge- Based Systems, 13(2-3), 101-112.

[9] Jani, H. M., & Mostafa, S. A. (2011). Implementing case-based reasoning technique to software requirements specifications quality analysis. International Journal of Advancements in Computing Technology, 3(1).

[10] Kwon, O., Im, G. P., & Lee, K. C. (2007). MACE-SCM: A multi-agent and case-based reasoning collaboration mechanism for supply chain management under supply and demand uncertainties. Expert Systems with Applications, 33(3), 690-705.

[11] Jiawei Han, Micheline kamber, Jian pei. DATA MINING Concepts and Techniques, 3rd Edition, 2011

[12] Leake, D. B. (1996). Case-Based Reasoning: Experiences, lessons and future directions. MIT press

[13] .Mostafa, S. A., Ahmad, M. S., & Firdaus, M. (2012). A soft computing modeling to case-based reasoning implementation. International Journal of Computer Applications, 47(7), 14-21.

[14] Richter, M. M., & Weber, R. O. (2016). Case-based reasoning. Springer-Verlag Berlin An.

Authors Profile

Jagjit Singh Dhatterwal is a PhD student in the Department of Computer Science, Mewar University. He earned his MCA from Maharshi Dayanand University, Rohtak and M.Tech degrees from KSOU , Karnataka India. His research interests include Artificial Intellegence and Agent Technology.

Sarvottam Dixit earned his PhD from Agra University, India in 1990. Currently, he is serving as the Professor and Advisor to Chancellor at Mewar University, India. His general research interests includes artificial intelligence, software engineering, multiagent etc.

S. Srinivasan completed his PhD from the Madurai Kamraj University, Madurai in 1978. Presently, he is a Professor at the PDM University, Bahadurgarh, Haryana – 124507. His research interests include artificial intelligence, software engineering, software testing etc.

