ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 8, August 2016

A Novel Dynamic Rule Activation Method for Continuous Ant Colony Optimization Based Extended Belief Rule-Based Systems

Ms. J.Banupriya[1], Ms. S. Kiruthika[2]

1) Research Scholar, Department of Computer Science, P.K.R. Arts College for Women, Gobichettipalayam

2) Assistant Professor, Department of Computer Science, P.K.R. Arts College for Women, Gobichettipalayam.

Abstract: Generally, decision Making and Modeling in high pressure, fast paced, complicated environments is commonly confounded by the lack of the choice model to capture the requisite style of true in a very ungenerous manner. This analysis focuses on a sophisticated belief rule-based decision model and proposes a dynamic rule activation (DRA) technique to deal with each problem at the same time. DRA relies on "smart" rule activation, wherever the active optimum rules are selected in a very dynamic way to seek for a balance between the wholeness and inconsistency within the rule-base generated from sample knowledge to attain a stronger performance. A series of case studies demonstrate however the utilization of DRA improves the accuracy of this advanced rule-based call model, while not compromising its potency, particularly once handling multi-class classification datasets. DRA has been tested to be beneficial to pick out the foremost appropriate rules or knowledge instances rather than aggregating a complete rule-base. Beside the work performed in rule-based systems, DRA alone are often considered a generic dynamic similarity method that may be applied in several domains. During this analysis Continuous ant Colony optimization (CACO) is applied within the context of Extended Belief Rule-Bases (E-brbs). The experimental results illustrated during this analysis demonstrate however the utilization of DRA will improve the accuracy of E-BRB primarily based decision support models.

Keywords: Rule-based processing, incompleteness, inconsistency, decision support, Continuous Ant Colony Optimization

1. INTRODUCTION

Research applied within the space of decision Support Systems (dsss) since the Seventies has incontestable however the utilization of this technology will improve each potency and effectiveness of users once approaching call issues [1]. Dsss usually include two main parts [1, 2]: (1) a cognitive content, delineated by a information illustration scheme, that is employed to store, manage and visualize the knowledge accessible for a given downside and (2) an reasoning engine, that analyses inputs against information from the cognitive content to provide helpful info for call support. Various information illustration schemes are often found in literature. Among them, rule-based systems are widely known joined of the foremost utilized in dsss, since components of human reasoning sometimes ought to be combined with a point of intelligence provided by the computer-based system [3].

Although alternative models like neural networks or support vector machines are powerful and probably correct approaches, they're unable to satisfy some case-specific necessities in terms of knowledge illustration and method transparency and interpretability [4]. Notably in real-case situations, interpretability becomes a vital issue, and it's essential to incorporate mechanisms to handle all the obscure, incomplete and unsure information sometimes found in such things, alongside precise information. The information illustration model elite for this analysis, therefore referred to as Extended Belief Rule-Base (E-BRB) [5], offer a solid framework to incorporate these problems as an integral a part of the cognitive content for the projected call model.

Once information and uncertainty are depicted within the cognitive content, the DSS reasoning procedure can ought to extract info that's relevant for a given set of inputs and learn from the cognitive content to provide a choice outcome. Within the case of the strategy given during this analysis, a given set of inputs are assessed so as to extract relevant info from the E-BRB. To satisfy this purpose, a two-step reasoning is projected. To higher perceive and represent the ideas behind the recommended two-step model, decision issues are located within the context of a version house (i.e. As a group of all hypotheses according to a given E-BRB) [6]. Here, extended belief rules are often thought-about as information tuples compound of their antecedents and resultant values. These tuples are often directly premeditated within the version house by mistreatment their values. Tuples with similar antecedents are closely placed within the space and thus are often classified, making hypertuples.

The reasoning procedures of the projected model are often delineated as Associate in Nursing rules that: (1) searches for an appropriate hypertuple (i.e. Set of extended belief rules) and (2) aggregates it to provide a consequence. Therefore, the projected rule consists of 2 steps: (i) a pre-processing procedure, named Dynamic Rule Activation (DRA), designed to seek out Associate in Nursing applicable Associate in Nursingd optimum hyper tuple for a given set of inputs and (ii) an reasoning method that aggregates the knowledge delimited among the hyper tuple to produce a result. The most stress of this analysis is to focus on the advantages, in terms of accuracy and potency, of DRA-CACO as a knowledge pre-processing methodology. Therefore, the second step of the strategy is changed throughout the various case studies to demonstrate however DRA-CACO improves the accuracy of the system for many reasoning processes. Note that DRA can't be outlined as an improvement rule, since none of the parameters enclosed within the E-BRB is tuned to boost the performance of the system. DRA-CACO would be higher delineated as a pre-processing, accessory methodology, which can be run in each system execution to facilitate any steps. During this regard, it's been designed to attain high levels of potency, compared to the procedure price that what would go for commonest improvement ways referred to as Continuous Ant Colony improvement to tune the quantity of belief degree distributions, relative weights and attributes of a whole rule-base. Here with Continuous Ant Colony improvement models is applied to the E-BRB while not moving the performance of DRA and presumably enhancing the accuracy.

2. METHODOLOGY

2.1. Extended Belief Rule Bases

Extended Belief Rule Bases (E-brbs) [5] are designed in recognition of the necessity to handle uncertainty and heterogeneous info in human deciding. E-brbs are designed with belief degrees embedded altogether the antecedents and resulting terms of a rule. They're wont to capture nonlinear causative relationships additionally as uncertainty. Due to these belief degree distributions, the direct input transformation is achieved by mistreatment the linguistic terms of the antecedent or resulting attribute, which may be outlined in concert of the subsequent cases: (1) Matching perform strategies using fuzzy membership functions, (2) Rule-based or Utility transformation strategiesor (3) Subjective valuation methods, for attributes of qualitative nature. View as example the subsequent extended belief rule: IF "Crime" is AND "Education" is (1) THEN "Deprivation" is In (1), the "Education" antecedent contains a belief degree of fifty for the "Satisfactory" term and two hundredth for the "Poor" one, feat undefined the remaining half-hour. Therefore, the missing half-hour represents the wholeness degree of the antecedent. To produce a wider application scope to the current explicit analysis, the E-brbs are often wont to integrate within the call method some key parts of real-world call issues, like unclearness (with linguistic terms), uncertainty (with beliefs), info wholeness (partially better-known beliefs in antecedents and/or consequents) and nonlinear relationships between indicators (with IF-THEN rules; between its antecedents and also the consequent). They additionally give a versatile thanks to incorporate hybrid input info (both quantitative and qualitative) additionally as a lot of economical rule generation scheme. The choice model that integrates these new E-brbs with the evidential Reasoning (ER) inference scheme has been named as RIMER+ [5].

2.2. Extended Belief Rules as Hypertuples

In order to produce a standardized base for the planned decision model and to modify the illustration of the underlying ideas behind the logical thinking strategies for E-brbs, it's price to find the choice drawback to be approached within the context of a version area [6].In

Machine Learning, a version space is that the set of all hypotheses that are in line with a group of coaching samples (tuples), and it's delimited by general and specific boundaries [6]. The set of all the hypotheses (classes) can|that may} be inferred from the E-BRB will outline the precise boundary of the version area. Situating the conception of E-BRB during this framework could be a easy procedure: Since all the data of the first dataset is transmitted to the E-BRB throughout the rule generation method by mistreatment belief degree distributions, every extended belief rule are often directly painted as a tuple for the version area. The antecedent values of every extended belief rule are often painted as a tuple. Each of those tuples are often directly allotted inside the version area by mistreatment its values. Therefore, tuples with similar values are closely situated within the version area and will then be classified to make hypertuples. Consequently, a hypertuple are often delineate as a group of tuples (i.e. Extended belief rules) with similar characteristics. Figure lillustrates a dataset with three attainable categories, represented by crosses, circles and triangles. For simplicity problems, the dataset shown has simply two attributes (a_l And a_2), so as to be aforethought during a two-dimensional area. Some hypertuples also are illustrated in Figure 1 mistreatment dotted rectangles. All of them are consistent hypertuples, as a result of they simply contain tuples of constant category - that's, all the principles inside every hypertuple have constant resulting part: either cross, circle or triangle. Since they're all consistent, the set of all hypertuples illustrated in Figure 1 could be a hypothesis for the underlying conception silent by the dataset and it defines the precise boundary of the version area of the dataset.

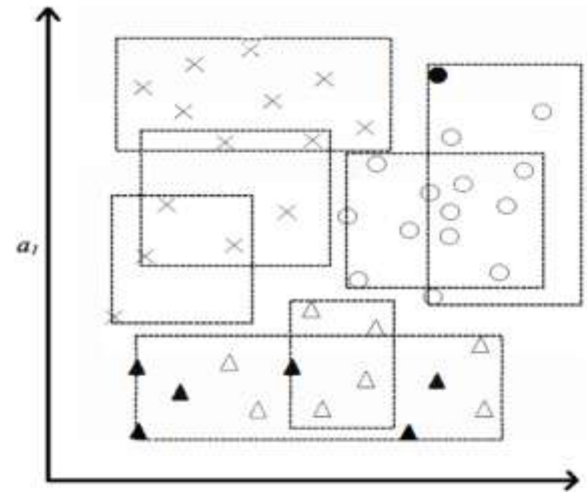

Figure l.A two-attribute (a_l And a_2) dataset with its specific boundary. Hypertuples are represented mistreatment dotted rectangles and every cross, circle or triangle represents a straightforward tuple. During this context, the method of assessing the inputs and rule activation in an E-BRB are often delineate as a procedure to seek out a standardized hypertuple containing appropriate relevant info (i.e. An acceptable set inside the E-BRB) to be inferred in additional steps to deliver helpful call support outcomes. Most call support models found in literature have confidence static rule activation procedures, wherever the primary hypertuple retrieved is taken without any consideration and is directly processed to retrieve a result.

2.3. Continuous Ant Colony Optimization

ACO has been applied to a broad range of hard combinatorial problems. Problems are defined in terms of components and states, which are sequences of components. Ant Colony Optimization incrementally generates solutions paths in the space of such components, adding new components to a state. The ACO system contains two rules:

1. Local pheromone update rule, which applied whilst constructing solutions.
2. Global pheromone updating rule, which applied after all ants construct a solution.

Furthermore, ACO algorithm includes two more mechanisms: trail evaporation and, optionally, daemon actions. Trail evaporation decreases all trail values over time, in order to avoid unlimited accumulation of trails

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 8, August 2016

over some component. Daemon actions can be used to implement centralized actions which cannot be performed by single ants, such as the invocation of a local optimization procedure, or the update of global information to be used to decide whether to bias the search process from a non-local perspective [7]. At each step, each ant computes a set of feasible expansions to its current state, and moves to one of these in probability. The probability distribution is specified as follows. For ant k, the probability of moving from state t to state n depends on the combination of two values [8]:

- The attractiveness of the move, as computed by some heuristic indicating the priori desirability of that move;
- The trail level of the move, indicating how proficient it has been in the past to make that particular move: it represents therefore an a posteriori indication of the desirability of that move.

Let consider a model $Q = (S, \Omega, f)$ of a continuous and discrete optimization problem consists of a search space S defined over a finite set of continuous decision variables and a set Ω of constraints among the variables and an objective function $f : S \to \mathbb{C}_0^+$ To be minimized. The search space S is defined as follows: Given is a set of continuous variables X_i , $i = 1, \ldots, n$, with values $v_i \in \mathcal{D}_i \subseteq \mathbb{C}$. A variable instantiation, that is, the assignment of a value v_i To a variable X_i, is denoted by $X_i \leftarrow v_i$. A solution $s \in S$—i.e., a complete assignment, in which each decision variable has a value assigned—that satisfies all the constraints in the set Ω is a feasible solution of the given cnop. If the set Ω is empty, Q is called an unconstrained problem model, otherwise is called a constrained one. A solution $s^* \subseteq S$ is called a global optimum if and only if: $f(s^*) \leq f(s)\, \forall s \in S$. The set of all globally optimal solutions is denoted by $s^* \subseteq S$. Solving a cnop requires finding at least one $s^* \subseteq S^*$. We use the Probability Density Function of a **Weibull** random variable is defined by: $f(x) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} & x \geq 0 \\ 0 & x < 0 \end{cases}$

The Weibull distribution PDF W_i Is parameterized by three vectors x, λ, and k. The solutions in the archive are used to calculate the values of these parameters, and hence shape the Weibull distribution PDF used to guide the ants in their search process. As mentioned earlier, in case of ACO, the pheromone information is stored as a solution archive. This implies that the pheromone update procedure has to perform some form of update on this archive. The size k of the archive T is a parameter of the algorithm. However, k may not be smaller than the number of dimensions of the problem being solved. At the start of the algorithm, the solution archive T is initialized generating k solutions by uniform random sampling. Pheromone update is accomplished by adding the set of newly generated solutions to the solution archive T and then removing the same number of worst solutions, so that the total size of the archive does not change. This process ensures that only the best solutions are kept in the archive, so that they efficiently guide the ants in the search process. So the first modified transition probability can be written as:

$$p_{ij}^k(t) = \begin{cases} \dfrac{\frac{k}{\lambda}\left(\frac{\tau_{ij}^k(t)}{\lambda}\right)^{k-1} e^{-\left(\frac{\tau_{ij}^k(t)}{\lambda}\right)^k} * \frac{k}{\lambda}\left(\frac{\eta_{ij}^k(t)}{\lambda}\right)^{k-1} e^{-\left(\frac{\eta_{ij}^k(t)}{\lambda}\right)^k}}{1 - \sum \frac{k}{\lambda}\left(\frac{\tau_{iu}^k(t)}{\lambda}\right)^{k-1} e^{-\left(\frac{\tau_{ij}^k(t)}{\lambda}\right)^k} * \frac{k}{\lambda}\left(\frac{\eta_{iu}^k(t)}{\lambda}\right)^{k-1} e^{-\left(\frac{\eta_{ij}^k(t)}{\lambda}\right)^k}} & u \in allowed \\ 0 & otherwise \end{cases}$$

The exponential function is characteristic by its large rate increase in values. That characteristic can be used to obtain the best result rapidly. In this work the refined Continuous Ant Colony Optimization algorithm is applied over the rules fetched from E-BRB. Using the confidence value as the pheromone (p) value and compute the path updating value $(P) = P + \Delta t$ Where$\Delta t = (2^{\,d+1} - 1)/d$, (d-Number of transaction set), then optimize association rule set is generated. The proposed method for generating better quality of rules by CACO is as follows:

1. Start
2. Load a patient's records from the database that fits in the memory.
3. Apply E-BRB algorithm to find the better rules.
4. Apply the CACO Algorithm.
5. Provide optimal rule generation results

The Dynamic Rule Activation (DRA) algorithm described in next Section has been designed to evaluate the hypertuple retrieved after the rule activation process and adjust it depending on each particular situation, in order to provide a more relevant and consistent set of rules to be processed afterwards.

2.4. Dynamic Rule Activation

The rule activation method is one in all the foremost vital steps in rule-based systems. Modifying the formulation during this element of the system could considerably have an effect on the retrieved outcomes. Therefore, the similarity live accustomed valuate the matching degree between inputs and rule antecedents becomes a vital issue, relevant to the performance of the whole system. Analysis has highlighted the importance of similarity measures and plenty of areas of analysis are centered on finding new strategies to enhance their performance [9]. Currently, most call support models (like RIMER+), are supported static rule activation, trusting in mere one set of activated rules retrieved from the similarity live for a given set of inputs. This reality not solely sometimes affects the output accuracy of the system, however typically results in things of integrity and/or inconsistency.

The projected Dynamic Rule Activation (DRA) technique enhances the performance of the whole DSS by evaluating the hypertuple of activated rules (Δ) for a group of inputs, when the similarity lives is applied. Supported the analysis retrieved, things of integrity (when |Δ| = 0) or inconsistency (when Δ contains rules with totally different consequents) is detected and stuck by enlarging or reducing the scale of Δ, severally. The scale of Δ is decided by its activation intervals (See Figure 2), that are controlled by powering the similarity operate to 1 parameter, named λ, exploitation the subsequent equation: $S_\lambda\ (i,a) = \left(S(i,a)\right)^\lambda$Where S is that the similarity operate applied to the input i and its corresponding antecedent attribute a, and λ is that the parameter dominant the scale of the activation intervals for Δ.

Any similarity operate may be used as S, as long because it meets the subsequent requirements: (1) it is normalized at intervals a [0, 1] interval and (2) the end result values of 1 and zero would mean excellent matching and total difference, severally. During this analysis, Milkowski's (Euclidean) distance was chosen as a operate that meets the necessities expected for S, though as aforesaid, the other operate might are chosen. Powering such S operates to λ could have an effect on the end result of SA in two totally different ways: (i) higher λ worths can punish rules with low activation weights (decreasing their activation weight to nearly zero) and (ii) lower λ values can increase their value. The previous can facilitate once endeavour inconsistency things, by checking out smaller, a lot of specific and consistent Δ hypertuples.

Meanwhile, the latter can facilitate avoiding integrity things, by opening the activation intervals checking out discourse data in larger Δ hypertuples. Figure a pair of shows however opening and shutting the activation intervals could facilitate in resolution these two things. Though this analysis is applied for datasets with any range of attributes, for illustrative functions the dataset pictured in Figure a pair of contains simply 2 attributes, premeditated in an exceedingly two- dimensional chart. Note that, during this study it's supposed that everyone the attributes of every rule of the E-BRB area unit connected with associate degree "and" relationship. Therefore, the eleven hypertuple is outlined because the intersection of the activation intervals of all the attributes of the dataset, pictured in Figure 2 exploitation gray rectangles.

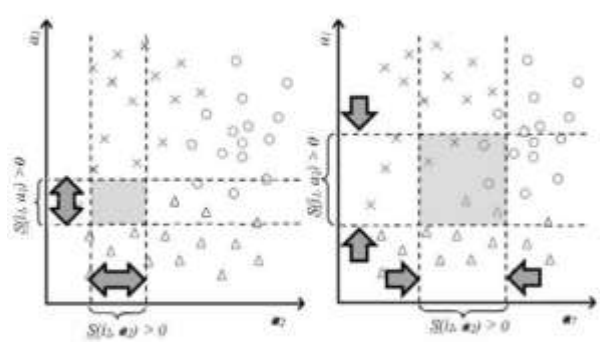

Figure 2. The left chart depicts a scenario of integrity, wherever the similarity operate describes the activation intervals (dashed lines) that outline associate degree empty Δhypertuple (depicted as a gray rectangle). The proper chart illustrates a scenario of inconsistency, during which the Δ hypertuple contains tuples from totally different categories. Daring arrows represent the chances offered to enhance these two things.

Let Φ be the hypertuple containing all the data at intervals the precise boundary of the version area and L the quantity of rules that completely match with the given set of inputs. Some general definitions is then explicit for the DRA method:

L. If λ=0, $\Delta_\lambda = \Phi$

2. If λ=∞, $|\Delta_\lambda|$=L

3. $\forall\lambda^* > \lambda$, $|\Delta^*_\lambda| \leq |\Delta\_\lambda|$, and what is more $\Delta^*_\lambda \in \Delta_\lambda$

Finally, note that if λ = one, there's no penalty or prize, simply the essential similarity live is applied. This last scenario will thus be thought-about because the start line of the algorithm. From this time, the strategy can frequently opt for either to lift the worth of A or decrease it, dominant during this means the scale of every Δ_λ counting on the case. The choice of either increasing or decreasing λ depends on three basic policies:

1. If there's a scenario of integrity (i.e. $|\Delta_\lambda|$=0), the worth of λ is small so as to open the activation intervals and thus increase the scale of Δ_λ (See Figure 2, left). Ultimately, this policy can cause capture discourse data [10] to be mass.

2. If there's a scenario of inconsistency (i.e. There's no agreement within the $\Delta_\lambda$hypertuple, as in Figure a pair of, right) increase the worth of λ so as to punish rules with low activation weights, shut the activation intervals and thus decrease the scale of Δ_λ, checking out higher levels of agreement.

3. It's fascinating to possess an oversized quantity of knowledge to be mass within the $\Delta_\lambda$hypertuple. It's thought-about that the larger quantity of relevant data accustomed generate the output, the a lot of reliable it'll be. To place this example in otherwise, take into account the worst case state of affairs, having a hypertuple of size $|\Delta_\lambda|$=1.

In this case, nearly each classifier would behave equally to the nearest Neighbor (INN) classifier, since it'd solely have one rule to combination and thus having a choice result primarily based solely on one coaching sample, that isn't the foremost reliable scenario in most cases (for example, if the sole instance offered is very inconsistent, providing the incorrect data to be aggregated). This example isn't fascinating and DRA has been designed to avoid it by checking out as massive $\Delta_\lambda$hypertuples as potential. DRA could be a technique that searches for a balance between these three principles, by calibration the worth of A so as to pick out the foremost acceptable size for the hypertuple to be processed in any steps. Therefore, the DRA technique is situated between the rule activation method (located within the content a part of} the system) and rule aggregation method (located within the abstract thought engine part of the DSS). Figure three illustrates the final structure of the choice model utilized in this analysis when adding DRA for pre-processing:

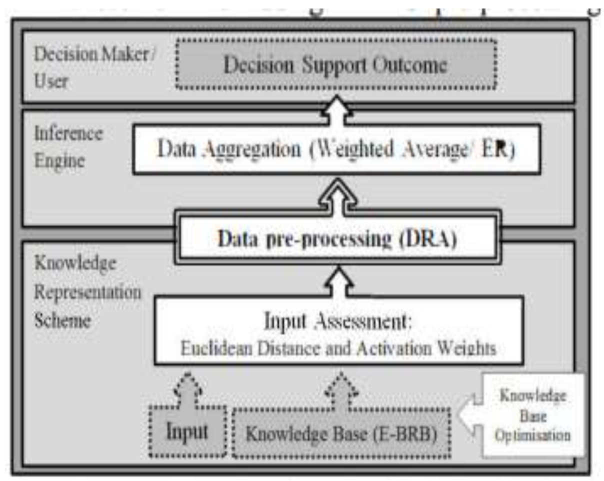

Figure 3.Architecture of the projected two-step decision model exploitation DRA as an information pre-processing technique.

Next Section demonstrates the connation of DRA, by comparing the performance of two abstract thought models used with and while not DRA embedded in their pre-processing step.

3. EXPERIMENTAL RESULTS AND DISCUSSION

Twenty-two benchmark datasets were collected from the ucirvine Machine Learning Repository. The EBRB-WA(performing the Weighted Average—WA) and RIMER+(performing the Evidential Reasoning—ER) decision models were tested with and without applying DRA as an activation algorithm (see Section 5).

3.1. False positive rate (FPR) Comparison

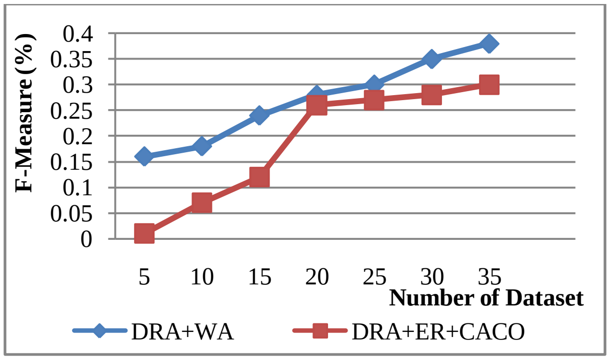

Fig.3. False Positive Rate Comparison

The graph of false positive rate in percentage is shown in Figure.3. The proposed method DRA+ER with CACO have low false positive rate 10% which is significantly enhance accuracy rate than the DRA+WA.

3.2. Accuracy Comparison

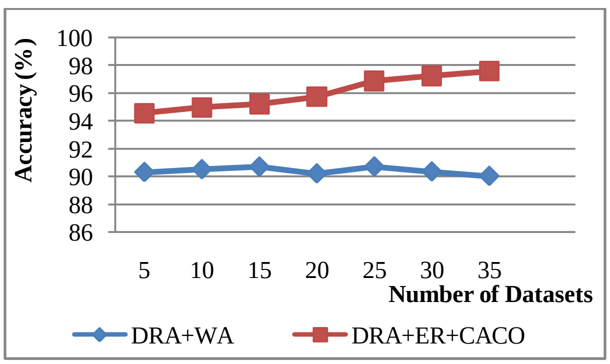

Fig.4. Accuracy Comparison

The accuracy of the proposed DRA+ER with CACO method achieves higher accuracy than the existing classification method of DRA+WA and the accuracy result is illustrated in Figure 4 as the proposed methods selects the important and optimal rules.

3.3. Execution Time comparison

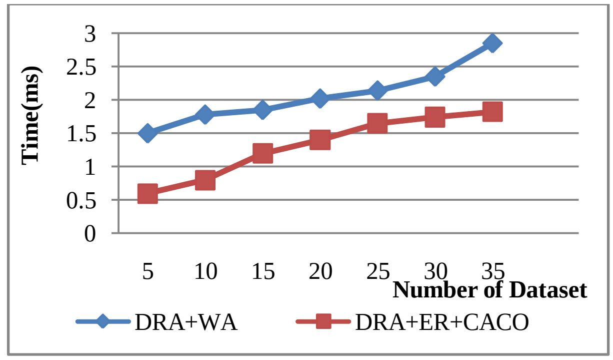

Fig.5.Execution Time comparison

Figure 5 shows the execution time comparison results of the proposed DRA+ER with CACO approach and the existing method DRA+WA. The proposed DRA+ER with CACO approach have less execution time to detect the rules, since the proposed method can effectively predict the optimal rules through CACO.

3.4. Number of rules

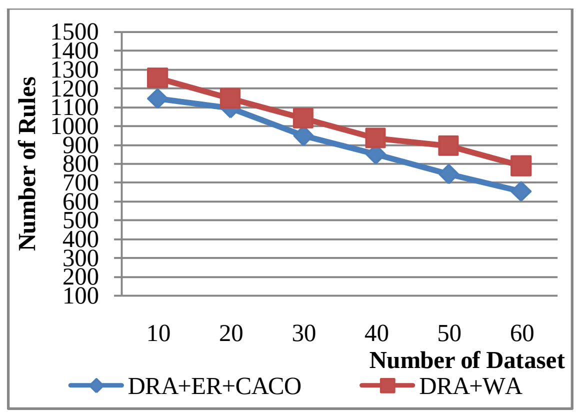

Fig.3. Number of rules vs. Support

Experimental results for number of results generated has been shown graphically in figure. From the Fig.3 it can be concluded that rules generated using proposed methodology are less than the existing method which indicates that proposed algorithm is better than existing approach.

4. CONCLUSION

A new Dynamic Rule Activation (DRA) technique was planned as a pre-processing step to pick out the foremost relevant data to be aggregative within a rule-based decision model. This analysis has incontestable however focusing simply on the foremost relevant data (discarding non-relevant, inconsistent rules) for every system execution not solely simplifies resulting aggregation ways, however conjointly will increase the accuracy of the system. At identical time, DRA provides a sublime, however effective method of tackling wholeness and inconsistency things right away, employing a easy, flexible, efficient and effective methodology. Note that DRA not solely opens the chance of achieving higher overall accuracies, however conjointly provides mechanisms to feature additional practicality to the choice model.

REFERENCES

1. Liu S.F., Duffy A.H.B., Whitfield R.I., Boyle I.M. (2010), "Integration of decision support systems to improve decision support performance", Survey Paper. Knowledge and Infonnation Systems, Volume 22, Issue 3, Pages 261-286
2. Phillips-Wren G., Mora M., Forgionne G.A., Gupta J.N.D. (2009), "An integrative evaluation framework for intelligent decision support systems", European Journal of Operational Research, Volume 195, Issue 3, 16 June 2009, Pages 642-652
3. Angelov P. And Yager R., (2010), "A simple fuzzy rule-based system through vector membership and kernel-based granulation". In Proceeding of 2010 IEEE International Conference on Intelligent Systems, pages 349-354, 7-9 July, London, UK.
4. Huysmans J., Dejaeger K., Mues C., Vanthienen J., Baesens B. (2011), "An empirical evaluation of the comprehensibility of decision table, tree and rule based predictive models", Decision Support Systems, Volume 51, Issue 1, April 2011, Pages 141-154
5. Liu J., Martinez L., Wang H., Calzada A, and Chen S.W. (2012), "A new belief rule base representation scheme and its generation by learning from examples", Proceeding of the 10th International FUNS Conference on Uncertainty Modelling in Knowledge Engineering and Decision Making (FLINS2012), Istanbul, Turkey, 27-29th, August, 2012.
6. Wang H., Diintsch I., Gediga G., Skowron A. (2004), "Hyperre1ations in version space", International Journal of Approximate Reasoning, Volume 36, Issue 3, July 2004, Pages 223-241
7. Frank, A. And Asuncion, A. (2010). "UCI Machine Learning Repository", Irvine, CA: University of California, School of Information and Computer Science.
8. Oong T.H., Isa N.A.M. (2012), "One-against-all ensemble for multiclass pattern classification", Applied Soft Computing, Volume 12, Issue 4, April 2012, Pages 1303-1308
9. Cha S.H. (2007), "Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions", International Journal of Mathematical Models and Methods in Applied Sciences, Volume 1, Issue 4, Pages 300-307
10. Wang H., (2012), "Contextual Probability and Neighbourhood Counting", International Journal of Software and Informatics, Volume 6, Issue 3, Pages 435-452.