Territorial Behavior based Algorithm for Data Classification

Supriya Singh¹, Dr. Mahendra Tiwari²
¹Student of M. Tech Computer Science, INDIA
²Sr. Assistant Professor (UCER), INDIA

ABSTRACT
Territorial behavior of lizard and presented
induction rules to evaluate the social behavior of Sinai
lizard. The algorithm is presented in this paper is based on
generating ordinal classification rules and comparatively
better with stat-of-art algorithms which are working in
categorical data.

Keywords----- Data mining, Artificial Intelligence, Ant
colony

I. INTRODUCTION

Behavioral study is scientific study of learning of
behavior in natural computing which investigates models
and computational techniques for information processing.
It is interdisciplinary field that joins natural science and
computational science for exploring new and hidden
pattern of social behavior of mammals. Experimental
analysis of Behavior analysis describes and develops
methods of changing behavior as well as provides services
to meet diverse behavioral needs [1] [4].

There are various computing paradigm abstracted
from natural phenomena which studies individual and
group behavior of mammals. To analyze behavior of
complex systems, different nature inspired models have
been used. Evolutionary computation [1] is computation
paradigm based on Darwinian principle of evolution.
Evolutionary system evolves evolution strategy [6] to
solve parameter optimization. Similarly evolutionary
programming [1] uses evolutionary techniques to obtain
goal of Artificial Intelligence by evolving population of
intelligent agent model like finite state machine. Moreover
genetic algorithm [3] features a population of individual
encoded as fixed strings. Social computing [4] is also
nature inspired computation paradigm within the purview
of social behavior of organism and it encompasses
intertwining of social behavior and computational system.
Scope of data mining and Knowledge discovery has been
rapidly increased in last 10 years and techniques of
natural computing plays important role to seek new
knowledge in some application domain [1]. Rule induction
process in data mining plays important role in knowledge
discovery task of social behavior . In literature, different
types of knowledge for behavior intelligence can be
encountered as Association rules, classification rules,
characteristics rules , clustering etc. Although, many
parties do not common in data mining and social behavior.
However, they can be used jointly to form a model which
often directs to the result, and in exhaustive literature
survey we observed that there are many classification and
clustering methods proposed to tackle optimization
problem in social behavior domain.

In literature, Ant Colony Optimization (ACO) [8]
algorithm is introduced to Analyze foraging behavior of
ants and ant miner is the first application of ACO proposed
for discovery of supervised classification rules. Similarly
cellular automata based adaptive artificial ant clustering
algorithm (AC) [7] is described forwdw active and
sleeping state of ant. Moreover, Ant part [8] and A Cluster
[8] algorithms are used in unsupervised classification
technique in social behavior of ant. However,artibee
Colony (ABC) [7]. Mosquito Swarm algorithm (MSA) [8],
and Shuffled Frog Leaping algorithm (SFLA) [7] are
another examples of foraging behavior. Furthermore,
Particle Swarm Optimization (PSO) [8] is stochastic,
population based global optimization technique inspired by
social behavior of bird flocking and Fish Swarm
Algorithm (FSA) [7] is another example of PSO that is
based on fish schooling behavior. In literature, there is
various PSO based classification and clustering methods
are proposed and hybrid PSO/ACO algorithm [7] is one of
them used for supervised classification rules.
Hunting behavior of animals is another elaboration of social behavior and there are some herd and pack algorithms contributed in this regard. Wolf pack search algorithm [6], feral dogs herd algorithm (FDHA) [8] and Dolphins herd algorithm (DHA) [8] are few examples in this domain.

In exhaustive literature survey, we observed that there is plethora of work related to swarm based social behavior. However it has been revealed that individual behavior based computational model seldom exist in literature. Territorial behavior of lizards can be characterized as individual social behavior. Precisely territory traditionally indicates an area fortified for exclusive use [1]. In literature there are three Hypotheses proposed regarding territory: defense of food, defense of mates, and defense of basking site [8]. There are certain related traits observed and analyzed during demonstrating of territorial behavior like body position, color change head bob, and push up. Some experimental data results described that intra specific aggression is important social behavior in all lizards. It demonstrates from specific site defense to home range defense [8].

It is explicit from extensive literature survey that many of swarm based algorithms suffer from drawbacks of poor performance in social behavior optimization. Just as, Slow convergence rate and local searching is weak in PSO and Ant system (AS) Likewise A common problem of ACO is stagnation or premature converge to a local optimum and it is designed for discrete search space and parameter initialization happens throughwdwd trial and errors. Similarly weak balance between exploration and exploitation is displayed in ABC and ACS. Moreover, in State transition rule of ACS, best state is selected by probability distribution and pheromone level is adjusted by applying global and local update rule. In addition, limited states of ant sleeping model (ASM) ] are defined by little local information and fitness function. However, in SFLA, due to lack of uniformity, population diversity is decreased. It is worthwhile to mention that our proposed algorithm is free from such drawbacks. In social behavior optimization of bio-inspired computing, Our objective in this paper is to introduce a new concept of rule induction method.

II. TERRITORIAL BEHAVIOR OF LIZARD

Geographic or microhabitat variation is one of the selective pressures for phenotypic divergence & specification. Spatial distribution of lizards has typically been described in terms of home range, indicating an area defended for exclusive use [4]. In experimental observation of defending home range, certain traits are analyzed like body position, color change, & head bob. Change of color is the most striking feature of lizards and colors are used in quite complex ways both physiologically and ecologically[3]. Three primary roles of color change in reptiles are thermoregulation, crypt city (camouflage), and signaling .Lizards have evolved physiological mechanism of thermoregulation. By temperature variation, Lizards are able to regulate their body temperature with managing the time they spend in sun and shade. In literature, there are some evidence suggesting that color associated with thermoregulation [5]. Similarly, according to Crypt city hypothesis, lizards change color to camouflage with background to escape predators. Signaling is almost opposite of crypt city. Lizards use bright color to communicate their social status. PseudotrapelusSinaitus () agamid shows rapid color change from brown to vivid blue. It is known to perch conspicuously on high rocks and display distinctive head bobbing motion. It is hypothesized that this color change is used as a form of social signaling.

Some fieldwork was carried out in exploring territorial behavior of lizard in south Sinai, Egypt. South Sinai is mountainous region interspersed with dried-up riverbeds [8].Lizard’s behavior was observed while walking along paths throughout the study areaDSGFEWgEea (rocks). An ethogram is constructed by observing and recording all of the activities of lizard. In order to correctly record the behaviors of lizard Focal-ani mal sampling method was used.

Territorial behavior of lizard was placed in four discrete categories: Lying flat, sitting up, head bob, & lying flat with head down [3,6]. Color is also associated to evaluate territorial behavior. When a lizard began to change color the underside of the throat was the first to turn blue and gradually, this blueness spread across the whole back. Blue is conspicuous color from a predation & camouflage perspective and it is also the key to an effective form of communication.

From experimental analysis, some displayed behaviors have been observed for defending the territory. These are relax, alert, aggressive, and submissive display.

III. BEHAVIORAL MODAL OF LIZARD

Lizard’s territorial behavior can be demonstrated by our proposed behavioral model. This behavioral model is based on party-based coupling. Party-based coupling reveals one-party-multiple operations which depict multiple behaviors are performed by same actor.

In literature, very limited research have been observed on formalizing and representing the concept of behavior, and no any formal behavior representation models exist for comprehensive understanding of behavior constitution. Although, traditional behavior model relies on qualitative methods borrowed from behavior and social science, but it misleads and ineffective because of partial analysis in understanding of social activities. Our proposed behavior model tries to remove such research limitations by formalizing a new social behavior framework.
IV. ALGORITHM

The proposed territorial based algorithm is introduced for data classification problem. Figure 1 presents the algorithm. Let $\mu$, $\delta$, $\rho$, and $\gamma$ are body positions of lizard. $\mu$ represents lying flat position and it represents the relaxed behavior of lizard. Similarly, $\delta$ demonstrates sitting up position and it shows alert behavior of lizard. Further, $\rho$ represents more sitting up position of lizard that shows aggressive behavior. Furthermore, $\gamma$ denotes lying flat with head down. Generally this crouched or submissive position shows losses of its territory by opponent. Moreover, $\alpha$, $\sigma$, and $\Upsilon$ represent head bob motion of lizard. When lizard is in relaxed basking position and does not feel any opponent approaching to him. He very seldom does head bob. $\alpha$ presents this secure feeling with no head bob motion. However, $\sigma$ shows 1 to 3 head bobs in some interval of time, from experimental analysis, it is found that this head bob represents alert behavior. Furthermore, when lizard encounters the opponent, it starts head bobbing rapidly and shows dominant position with 1 to 8 head bobs. We represent this aggressiveness by symbol $\Upsilon$.

Change of color is also an effective social signaling approach. Let $\beta$, $\Delta$, $\xi$, and $\eta$ are colors displayed by lizard according to situation. $\beta$ represents brown color which is associated with color of rock. At rest, lizard tends to be brown color that matches with environment. Similarly, $\Delta$, and $\xi$ represent normal and vivid blue respectively. Moreover, $\eta$ represents the grey color which displays the submissive behavior of lizard.

### Data
- Different behavior of Sinaitus lizard in microhabitat

### Result
- Relaxed display of lizard in FEW microhabitat
- Alert display of lizard in microhabitat
- Aggressive display of lizard in microhabitat
- Submissive display of lizard in microhabitat

### Begin
Step 1: Analysis of relaxed display

For (each display in microhabitat) do

Find body position, head bob, color change of lizard in microhabitat

If (body_position=µ && head_bob=α && color_change=β) then

Declare it as relaxed display

End if

End for

Step 2: Analysis of alert display

For (each display in microhabitat) do

Find body DSposition, head bob, & color change of lizard in microhabitat

If (body_position=δ && head_bob=σ && color_change=∆) then

Declare it as alert display

End if

End for

Step 3: Analysis of aggressive display

For (each display in microhabitat) do

Find body position, head bob, & color change of lizard in microhabitat

If (body_position=ρ && head_bob=τ && color_change=£) then

Declare it as aggressive display

End if

End for

Step 4: Analysis of submissive display

For (each display in microhabitat) do

Find body position, head bob, & color change of lizard in microhabitat

If (body_position=γ && head_bob=α && color_change=η) then

Declare it as alert display

End if
End for  

Step 5: Collect all the classification rules  
endASC

Figure 1: Territorial behavior based algorithm

Table 1: Displayed operations in different behaviors

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Body Position(radian)</th>
<th>Head bob(acceleration in m/s²)</th>
<th>Color (frequency in NHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relaxed</td>
<td>0.523 (30 degree)</td>
<td>0.01</td>
<td>0.480 (brown)</td>
</tr>
<tr>
<td>Alert</td>
<td>0.785 (45 degree)</td>
<td>0.03</td>
<td>0.610 (blue)</td>
</tr>
<tr>
<td>Aggressive</td>
<td>1.134 (65 degree)</td>
<td>0.07</td>
<td>0.620 (vivid blue)</td>
</tr>
<tr>
<td>Submissive</td>
<td>0.261 (15 degree)</td>
<td>0.00</td>
<td>0.430 (grey)</td>
</tr>
</tbody>
</table>

In table 1 various operations of lizard is measured and mentioned. Body position of lizard is measured in radian. Similarly, head bob is calculated in terms of m/s² and color is measured in NHz. Let body position can be considered as x and head bob is considered as y. Moreover, color is considered as z.

The correlation between these variables are found out and tried to prove that these operations are positively correlated.

First x and y variables are calculated with Pearson product correlation coefficient formula. The result of this calculation is 0.8915/0.1210=7.3677 which indicates that both values are positively correlated.

Again, we took another variable z and correlate it with y. The result is 0.7062/0.1116=6.3231. This result is also proven that, the operations are performed by lizard is positively correlated.

V. EXPERIMENTAL SETUP

We used Weka 3.6.8 to evaluate and compare of performance of territorial behavior based algorithm with state-of-the-art algorithms. The data for comparison is obtained from UCI repository. The details of dataset are given below:
Table 2: Details of dataset

<table>
<thead>
<tr>
<th>Name of Dataset</th>
<th>Type of file</th>
<th>Number of attributes</th>
<th>Number of instances</th>
<th>Attribute characteristics</th>
<th>Dataset characteristics</th>
<th>Missing value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zoo</td>
<td>CSV (comma separated value)</td>
<td>17</td>
<td>101</td>
<td>Categorical, Integer</td>
<td>Multivariate</td>
<td>No</td>
</tr>
<tr>
<td>Letter Image Recognition</td>
<td>CSV (comma separated value)</td>
<td>17</td>
<td>174/20000</td>
<td>Categorical, Integer</td>
<td>Multivariate</td>
<td>No</td>
</tr>
<tr>
<td>Labor</td>
<td>ARFF (Attribute Relation File Format)</td>
<td>17</td>
<td>57</td>
<td>Categorical, Integer</td>
<td>Multivariate</td>
<td>No</td>
</tr>
<tr>
<td>Supermarket</td>
<td>ARFF (Attribute Relation File Format)</td>
<td>217</td>
<td>4627</td>
<td>Categorical, Integer</td>
<td>Multivariate</td>
<td>No</td>
</tr>
</tbody>
</table>

Modeling and evaluation involves four steps: Classification and prediction are two forms of data analysis that can be used to extract models describing important data classes or to predict future data trends. While classification predicts categorical labels (classes), prediction models continuous functions.

Table 3: The details of applied classifier:

<table>
<thead>
<tr>
<th>Classification Techniques</th>
<th>Classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Tree</td>
<td>Decision stump, REP, CART</td>
</tr>
<tr>
<td>Lazy learner</td>
<td>IBK, Kstar</td>
</tr>
<tr>
<td>Rules based</td>
<td>One R, Zero R</td>
</tr>
<tr>
<td>Naïve bayes</td>
<td>Naive</td>
</tr>
<tr>
<td>Regression</td>
<td>Linear regression</td>
</tr>
</tbody>
</table>
To evaluate the accuracy of data mining algorithms on selected datasets, several experiments are conducted. For evaluation purpose, two test modes are used:

K-fold cross-validation (k-fold CV) mode. Percentage split (holdout method) mode.

The k-fold CV randomly divides in to k disjoint blocks of objects, then the data mining algorithm is trained using k-1 blocks and the remaining block is used to test the performance of the algorithm, this process is repeated k times. At the end, the recorded measures are averaged. It is common to choose k=10 or any other size depending mainly on the size of the original dataset. In percentage split (holdout method), the database is randomly split in to two disjoint datasets. The first set, which the data mining system tries to extract knowledge from called training set. The extracted knowledge may be tested against the second set which is called test set, it is common to randomly split a data set under the mining task in to 2 parts. It is common to have 66% of the objects of the original database as a training set and the rest of objects as a test set.

Experimental Result and discussion:

The performance of classifier can most simply be measured by counting the proportion of correctly predicted examples in training and test dataset. This value is the accuracy. The accuracy of a classifier on a test set is the percentage of test set tuples that are correctly classified by the classifier. It is also referred as recognition rate of classifier.

<table>
<thead>
<tr>
<th>Model</th>
<th>instance</th>
<th>Correctly classified</th>
<th>Incorrectly classified</th>
<th>TP rate</th>
<th>FP rate</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree Decision stump</td>
<td>1573/4627</td>
<td>1027(65.28%)</td>
<td>546(34.71%)</td>
<td>0.719</td>
<td>0.386</td>
<td>0.526</td>
<td>0.719</td>
</tr>
<tr>
<td>SMO</td>
<td>1573/4627</td>
<td>986(62.68%)</td>
<td>587(37.31%)</td>
<td>0.627</td>
<td>0.627</td>
<td>0.393</td>
<td>0.627</td>
</tr>
<tr>
<td>Lazy-IBK early1573/627</td>
<td>602(38.27%)</td>
<td>971(61.72%)</td>
<td></td>
<td>0.383</td>
<td>0.37</td>
<td>0.678</td>
<td>0.383</td>
</tr>
<tr>
<td>NaiveBayes</td>
<td>1573/4627</td>
<td>986(62.28%)</td>
<td>587(37.71%)</td>
<td>0.627</td>
<td>0.627</td>
<td>0.393</td>
<td>0.627</td>
</tr>
<tr>
<td>Rules-OneR</td>
<td>1573/4627</td>
<td>1030(65.48%)</td>
<td>543(34.52%)</td>
<td>0.655</td>
<td>0.655</td>
<td>0.642</td>
<td>0.655</td>
</tr>
<tr>
<td>Tree-CART</td>
<td>1573/4627</td>
<td>986(62.28%)</td>
<td>587(37.31%)</td>
<td>0.627</td>
<td>0.627</td>
<td>0.393</td>
<td>0.627</td>
</tr>
<tr>
<td>Rules-ZeroR</td>
<td>1573/4627</td>
<td>986(62.28%)</td>
<td>587(37.71%)</td>
<td>0.627</td>
<td>0.627</td>
<td>0.393</td>
<td>0.627</td>
</tr>
</tbody>
</table>

Methods of estimation:
Accuracy can be estimated with:
Correctly classified instances
Confusion Matrix
Cost Matrix
Error Estimation
Classifier models can predict accuracy on basis of correctly classified instances from training as well as test dataset, the error rate of model describes the incorrectly classified the instances. The accuracy is also estimated by computing an error based on the difference between the predicted value and the actual known value for each of the test tuples.

VI. CONCLUSION

Performance of Data Mining algorithms are analyzed and evaluated by accuracy, clarity, ability to handle corrupted data, scalability, and speed. Accuracy estimated by accuracy measurement tool and method of estimation Kstar, Naïve, and SMO are performed better than other algorithms while using labor data set, cross validation method, and correctly classified tool used but, if percentage split method used Naïve performed good. Accuracy of OneR is better than other when supermarket dataset used with correctly classified tool along with cross validation as well as percentage split method used. When accuracy is evaluated with confusion matrix the result is changed, Kstar and Decision stump performed better than others with labor data set and cross validation and percentage split used. While OneR and Decision stump are good with supermarket dataset. In cost matrix tool
accuracy of kstar and Naïve are better than others using labor data set and cross validation and percentage split method used accordingly

REFERENCES

[19]www.behaviorinformatics.org