International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 3 Issue 5, August 2019 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Knowledge Discovery in Classification and Distribution of Butterfly Species From Dagon University Campus, Myanmar by Rule Induction: CN2 Algorithm

Su Myo Swe[1], Khin Myo Sett[2]

[1]Lecturer, [2]Professor

[1]Department of Computer Studies, Dagon University, Yangon, Myanmar

[2]Department of Computer Studies, University of Mandalay, Myanmar

How to cite this paper: Su Myo Swe | Khin Myo Sett "Knowledge Discovery in Classification and Distribution of Butterfly Species From Dagon University Campus, Myanmar by Rule Induction: CN2 Algorithm" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-5, August 2019, pp.600-603, **https://doi.org/10.31142/ijtsrd26380**

IJTSRD26380

ABSTRACT

Rule induction is an area of machine learning in which formal rules are extracted from a set of observations. The CN2 induction algorithm is a learning algorithm for rule induction. In this research, "The Fauna of British India, Ceylon and Burma. Butterflies. Vol. I and Vol. II" written by C.T Bingham are used as the required knowledge for the resource. This research applies the CN2 algorithm to discover the knowledge of butterfly species rules for classification from Dagon University campus and a system has also developed. In this system, 29 butterflies as a zoological dataset, MS Visual Studio 2012 as a programming tool and MS SQL Server as for database development are used.

KEYWORDS: *Rule Induction, CN2 Algorithm, classification*

1. INTRODUCTION

In the current year, increasing the number of publishing paper and show great interest of engineers and scientist towards the importance of metal nano particles such as copper, silver, and gold. The cluster has a particle size of 1-10 nm, which composed of 102-104 atoms. The size of nanoparticles is 10 to 50 which contain in most cases 105-106 atoms however according to the nanoparticles dimension it has a size of 1-9 nm metal clusters are poised of metallic elements it can be single pe of the metallic element or more than one, grouping to the subclass of intermetallic or (nanoalloy) clusters. The establishment of the cluster in the larger size of metal nanoparticles from nanometer to the micro dimension known as metal colloids. They are popular among nature lovers as well as a subject for scientific study. The present research aims to discover the knowledge about. the rules of the butterflies and forecast their distribution abundance in Dagon University campus.

2. Materials and Methods

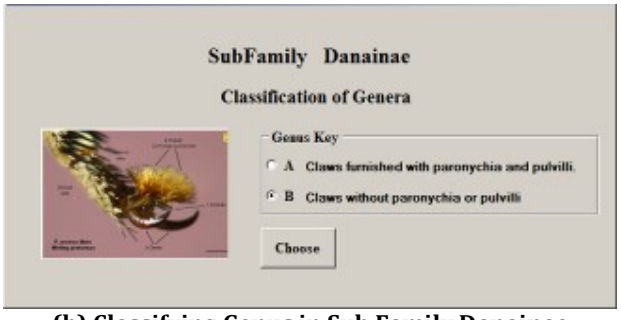
A dichotomous key is a tool that can be used to classify objects. Dichotomous means "divided in two parts" or "binary classification". Therefore, this key uses a series of yes or no questions to place objects into groups.[1] [2]The butterflies are entirely depending upon their habitat. So, there are 29 butterflies were collected in Dagon University for classifying.

This research continued the result of our previous research paper.[7] In the previous research, we mentioned the literature review about the rule induction: CN2 algorithm and about the nature of the zoological dataset. We also expressed the system flow diagram for the proposed system. The result of a previous research paper is family rules for butterfly.[7] We used these family rules for continue classification of butterflies form family to species.

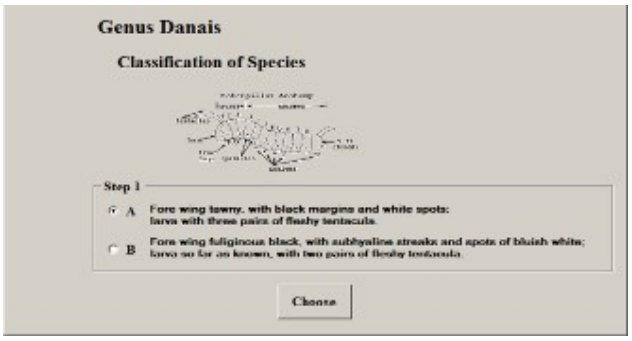
3. Proposed System

In this research paper, we proposed the following system and mentioned the classifying one family (Nymphalidae) as an illustration. Figure (3.1) shows the classification of family Nymphalidae.

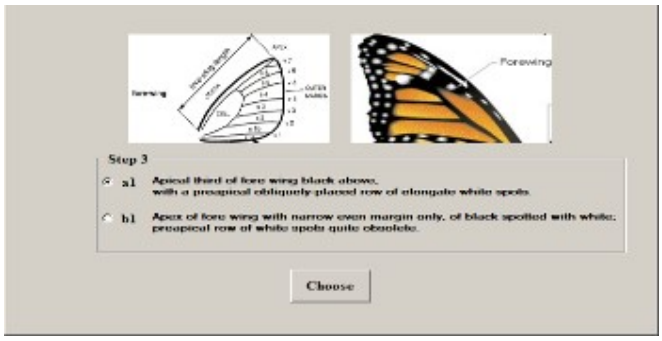
(a) Classifying SubFamily in Family Nymphalidae



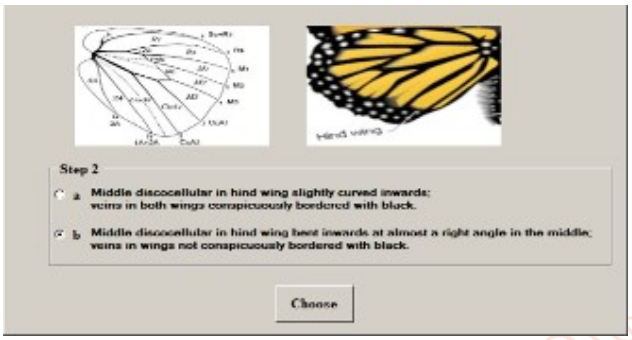
(b) Classifying Genus in Sub Family Danainae



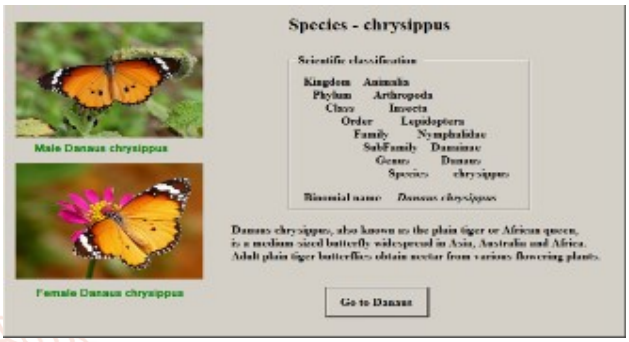
(c) Step1 for Classifying Species in Genus Danais



(e) Step3 for Classifying Species in Genus Danais



(d) Step2 for Classifying Species in Genus Danais



(f) Result species of Genus Danais

Figure (3.1) Classifying Family Nymphalidae

4. Result and Discussion

Rules of butterflies are show in figure (4.1) which is the result in the proposed system.

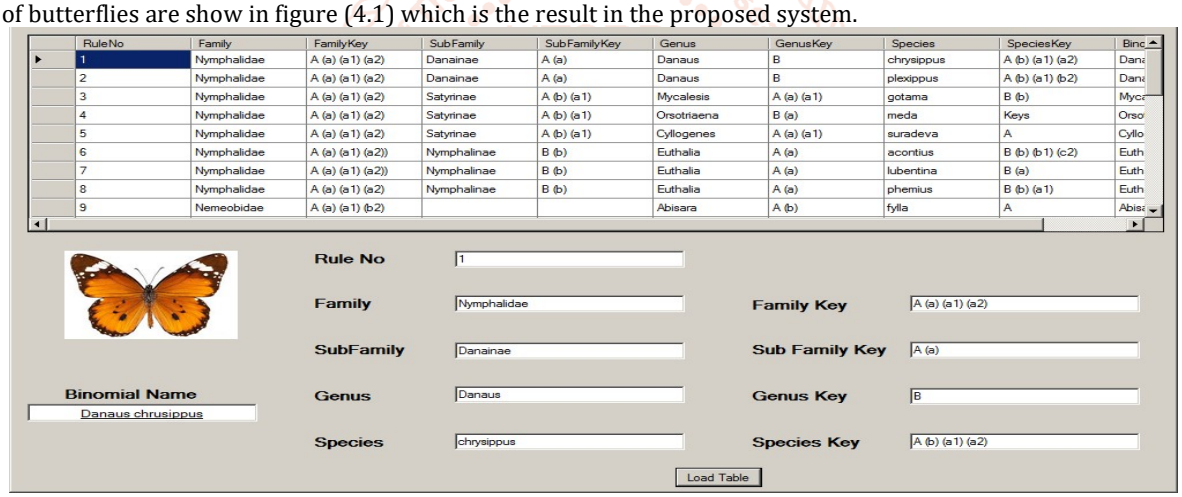


Figure (4.1) Rules of the Butterflies

4.1 Grouping the Rules of Butterflies according to with the Family Family – Nymphalidae

There are eight butterfly species (family- Nymphalidae) among collected 29 butterflies. Table (4.1) shows the rules of family Nymphalidae.

Table (4.1) Rules of Family Nymphalidae

Sr. No	Genus Rule	Genus Name	Species Rule	Species Name
1	B	Danaus	A (b) (a1) (a2)	chrysippus
2	B	Danaus	A (b) (a1) (b2)	alcippus
3	A (a) (a1)	Mycalesis	B (b)	gotama
4	B (a)	Orsotriaena	Keys	meda
5	B (b)	Cyllogenes	A	suradeva
6	A (a)	Euthalia	B (a)	lubentina
7	A (a)	Euthalia	B (b) (a1)	phemius
8	A (a)	Euthalia	B (b) (b1) (c2)	acontius

Family – Nemeobidae

There are two butterfly species (family- Nemeobidae) among collected 29 butterflies. Table (4.2) shows the rules of family Nemeobidae.

Table (4.2) Rules of Family Nemeobidae

Sr. No	Genus Rule	Genus Name	Species Rule	Species Name
1	A (b)	Abisara	A	fylla
2	A (b)	Abisara	B	echerius

Family – Papilionidae

There are five butterfly species (family- Papilionidae) among collected 29 butterflies. Table (4.3) shows the rules of family Papilionidae.

Table (4.3) Rules of Family Papilionidae

Sr. No	Genus Rule	Genus Name	Species Rule	Species Name
1	A	Memnon	A (a)	agenor
2	A	Memnon	A (b)	polymnestoroides
3	A	Memnon	B	mayo
4	B	Polytes	A	polyts
5	B	Polytes	B	pitmani

Family – Pieridae

There are twelve butterfly species (family- Pieridae) among collected 29 butterflies. Table (4.4) shows the rules of family Pieridae.

Sr. No	Genus Rule	Genus Name	Species Rule	Species Name
1	A (a) (b1)	Delias	A	eucharis
2	A (a) (b1)	Delias	B	hierta
3	B (a) (a1)	Appias	A	nero
4	B (a) (a1)	Appias	B (a)	libythea
5	B (a) (a1)	Appias	B (b) (a1)	leis
6	B (a) (a1)	Appias	B (b) (b1)	wardii
7	B (b)	Hebomoia	A	glaucippe
8	B (b)	Hebomoia	B	roepstorff
9	B (a) (b1)	Catopsilia	A (a)	crocale
10	B (a) (b1	Catopsilia	A (b)	scylla
11	B (a) (b1	Catopsilia	B (a)	pyranthe
12	B (a) (b1	Catopsilia	B (b)	florella

Table (4.4) Rules of Family Pieridae

Family – Lycaenidae

There are two butterfly species (family- Lycaenidae) among collected 29 butterflies. Table (4.5) shows the rules of family Lycaenidae.

Table (4.5) Rules of Family Lycaenidae

Sr. No	Genus Rule	Genus Name	Species Rule	Species Name
1	B (b)	Castalius	A	ananda
2	B (b)	Castalius	B	rosimon

Family – Hesperidae

There have no butterflies species (family- Hesperidae) among collected 29 butterflies aroung Dagon University campus.

4.2 Grouping the Abundance of Butterflies

Table (4.6) shows the grouping the abundance of butterflies according with their genus and species.

Table (4.6) Number of Butterflies Abundance

Sr. No	Family Name	Genus Abundance	Species Abundance
1	Nymphalidae	5	8
2	Nemeobidae	1	2
3	Papilionidae	2	5
4	Pieridae	4	12
5	Lycaenidae	1	2
6	Hesperidae	0	0
Total			29

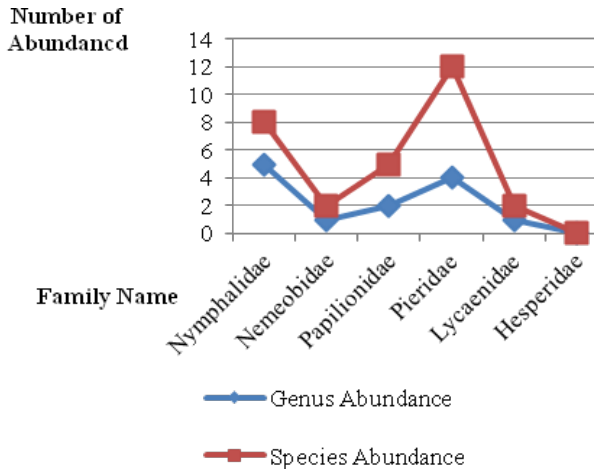


Figure (4.2) Abundance of Butterflies around Dagon University

According Figure (4.2), some butterfly families are rare abundance, some are moderate and some are more another family while one family; Hesperidae has no species occur around Dagon University.

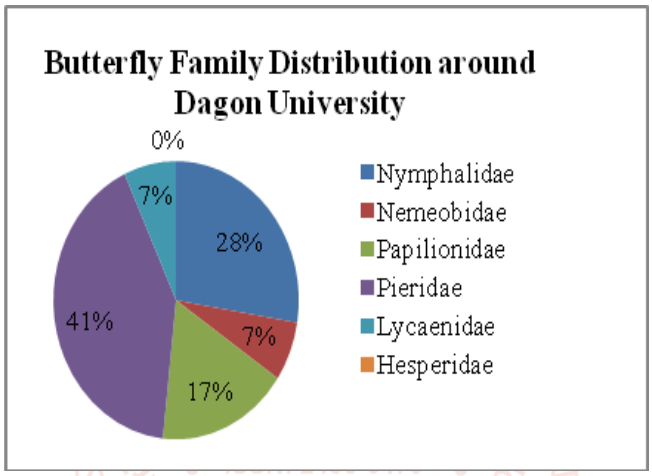


Figure (4.3) Butterflies Family Distribution around Dagon University

5. Conclusion

This research paper has introduced the idea of inducing rules from sets of examples in order to speed up the development of knowledge bases in expert systems. The rule induction with CN2 algorithm is the best technique for the classification of a biological dataset. The results give rules for classification and the implemented system is easy to use and understand for a biologist; who is willing to classify the butterflies by computerization. Additionally, the researchers can predict the distribution of fauna (butterfly) by analyzing the result of classification. Moreover, children in schools are being encouraged to study the life and habit of butterflies and in this way come to appreciate the beauty of the natural environment and become involved in efforts for its conservation. So, this research will meet the aim of correct, easy, fast, simple and useful than traditional manual methods.

References

[1] Bingham. C.T, "The Fauna of British India, Ceylon and Burma. Butterflies. Vol. I Taylar and Francis, London", 1905

[2] Bingham. C.T, "The Fauna of British India, Ceylon and Burma. Butterflies. Vol. II Taylar and Francis, London", 1907

[3] Clark Peter and Boswell Robin. "Rule Induction with CN2: Some Recent Improvement" Machine Learning, Fifth European Conference, 1991

[4] Clark Peter and Niblett T. "The CN2 induction algorithm" Machine Learning Journal, 3(4); 261-283, 1989

[5] Faruk Ertuğrul Ömer et al. "Identification of Butterfly Species by Similarity Indexes Based on Prototypes", International Journal of Computer Vision, Machine Learning and Data Mining, Volume 1, 2015

[6] Khin Mya Mya et al." Taxonomic Studies on Some Butterflies of Dagon University Campus".2001

[7] Swe Su Myo and Sett Khin Myo. "Approaching Rules Induction: CN2 Algorithm in Categorizing of Biodiversity", IJTSRD, Vol. 3, Issue. 4, May-June 2019

[8] Tan C. F. "The Application of Expert System: A Review of Research and Applications", ARPN Journal of Engineering and Applied Sciences, Vol. 11, No. 4, February 2016

[9] https://docs.orange.niolab.si/3/visual programming/widgets/model/

[10] cn2ruleinduction.html https://en.wikipedia.org/wiki