International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-5, March 2020

Mathematical Modelling for Prediction and Classification of Neonatal Infections by using Rough Set

Sipali Pradhan, Jitendu Kumar Mantri, Sujogya Mishra, P. K. Pattnaik

Abstract: ***Human contagious diseases are basically affected newly born baby. Millions of death occur due to these diseases of our intention to give a prediction of such diseases. Study report analysis the symptom of this diseases with minimum attributes by using Rough set Technique and to predict the contagious diseases are using Time series model.***

Keywords: ***AI-Aritificial Intelligence, RST-Rough Set Theory,CS-Congetive Science, ML-Machine Learning, KD-Knowledge Discovery,DA-Data Analysis, Data Minningand Time series.***

I. INTRODUCTION

RST is a computational method to handle obscurity and blurriness. TSM statistical method used for prediction. These methods are largely improved computational information by using DA and DM tools. AI and CS are alternative computational pathway to translate the evolutionary algorithm to mathematical approaches by using RST. The fundamental of RST is Lower and Upper approximation denoted as $\underline{U}(x),\overline{U}(x)$ respectively. Our intention to find result about the boundary. By using RST technique, the work basically tries to reduce the approximate errors. This is somewhat mixture of soft computing and statistical analysis. Purpose to produce good predictive result [7, 8, 9].

II. BASIC IDEAS

The concept is develop from the present medical situation around the globe. Our objective to find attribute which leads to increase in infant mortality rate. One of the major diseases which increase infant mortality rate is Neonatal infections. This paper adopt time series analysis [10, 11, 12] to know how this diseases spreading with time.

A. Information and Analysis

Bimonthly incident of NI is collected form two different medical enviroments from 2008 to 2017.

$$\text{BimonthlyIncdent} = \frac{Number\ of\ NIBimontly}{StockinMarket} \times R$$

Revised Manuscript Received on March 30, 2020.
* Correspondence Author
Sipali Prdhan, Ph.D Scholar, North Odisha University, Odisha, India.
Dr. Jitendu Kumar Mantri, Professor, Department of Computer Science, North Odisha University, Odisha, India.
Dr. Sujogya Mishra, Assistant Professor, Department of Mathematics and Computing, CET, Bhubaneswar, Odisha, India.
Dr. Pradyumna Kumar Pattnaik, Assistant Professor, Department of Mathematics and Computing, CET, Bhubaneswar, Odisha, India.

the bimonthly result given below by using Statistica software

B. Analysis Phase

The data collected for the model of the diseases NI
Being analyze using Statestica we found the following result

ANALYSIS

Table -1

TYPES	KV	LE	RR	AIC	BIC
AR(1)	1.212	1.131	60.223	326.540	330.829
AR(5)	1.412	1.187	46.492	302.848	321.108
AR(1)	0.145	0.380	-5.942	55.884	60.039
AR(2)	0.138	0.371	-4.128	54.256	60.44
AR(3)	0.116	0.340	-8.525	45.059	52.320

PREDICTIVE RESULT BIMONTHLY

Table -2

Bimonthly	True v	Expected v	Error	TRUE v	Expected v	Error
January	0.3135	0.4748	0.161	0.1373	0.067	-0.6
March	0.4876	0.4301	0.057	0.3212	0.178	0.1
May	0.4467	0.4982	0.051	0.1257	0.177	0.05
July	0.4548	0.5079	0.053	0.0727	0.178	0.1
Sep	0.4052	0.454	0.048	0.208	0.066	0.14
Nov	0.3673	0.4135	0.046	0.0773	0.067	0.1

C. Attribute Reduction Phase-I

From different medical Journal as well from different Hospital we get the information that Neonatal infections has several symptoms our intention to find which symptom are actually responsible for this diseases, initially this paper guess 6 symptoms consider as attributes (Conditional)along with it's values and get a decision.

This paper consider 6 conditional attributes as High fever, Variable body temperature, Cough, Low birth, Body stiffness, Patches in bodies and it's values are significant and insignificant and it' decision attribute values are noTable and pointless. for better presentation this paper rename the conditional attributes as I_{11}, l_{12}, l_{13}, l_{14}, l_{15}, l_{16} and it's values are a_{11}(insignificant) and b_{11}(significant) and decision attributes as d_{11} and it's values are 1(noTable) 2(point less). Analysis being applied on the 1000 records which reduced to 6 clusters.

Our intention to find the essential attributes responsible to for NI we are findinding the sets of reduct and the core[11,12,13,14,15] to find the attributes which are essential for this purpose.

DATA

Table -3

E	l_{11}	l_{12}	l_{13}	l_{14}	l_{15}	l_{16}	d_{11}
E_1	a_{11}	b_{11}	b_{11}	a_{11}	a_{11}	b_{11}	2
E_2	a_{11}	b_{11}	a_{11}	a_{11}	b_{11}	b_{11}	2
E_3	a_{11}	a_{11}	b_{11}	a_{11}	a_{11}	b_{11}	1
E_4	b_{11}	a_{11}	a_{11}	b_{11}	b_{11}	a_{11}	1
E_5	b_{11}	a_{11}	a_{11}	b_{11}	a_{11}	b_{11}	1
E_6	b_{11}	a_{11}	a_{11}	b_{11}	b_{11}	b_{11}	1

Reduct [9] as follows

1. (l_{11}, l_{12}, l_{13}, 1_{l4})
2. (l_{12}, l_{13},, 1_{l4}, l_{15})
3. (l_{11}, l_{122}, l_{14} l_{16})
4. (l_{11}, l_{12}, l_{14}, l_{15})
5. (l_{12}, l_{13}, l_{14}, l_{16})

From the reduct set it is clear that attributes (l_2, l_4) are essential attributes for the cause of Neonatal infections. As (l_{12}, l_{14}) is the core for the five reduct sets.

III. APPLICATION OF STRENGTH OF ROUGH SET IN FINDING REDUCT

The above data set of the data Table-1 again applied to find the rules in Phase-2:

a. Initialize $k_1 = \varphi$ where reduct as k_1

b. Repeat N1 (attribute sets) N_1 and $K_1 \neq \varphi \ \forall \ a \in A$ where A is attributes space.
(Conditional attribute values with respect to decision attribute values)

c. Recursively finding the reduct sets
Strength = (conditional attribute) *value* / (Decision attribute) *value* = -values D' /-values' (with respect to count) of both conditional attribute values and decision attribute values

d. Count of the ratio fall into one categories then, Reduct++
Otherwise goto step 4.
End both if and for respectively where N' & K' ∈ E' (Records).
Dvalues for values of decision attributes and Cvalues are values of the conditional attributes.

DATA

Table -4

E	l_{11}	l_{12}	l_{13}	l_{14}	l_{15}	l_{16}	d_{11}
E_1	a_{11}	b_{11}	b_{11}	a_{11}	a_{11}	b_{11}	2
E_2	a_{11}	b_{11}	a_{11}	a_{11}	b_{11}	b_{11}	2
E_3	a_{11}	a_{11}	b_{11}	a_{11}	a_{11}	b_{11}	1
E_4	b_{11}	a_{11}	a_{11}	b_{11}	b_{11}	a_{11}	1
E_5	b_{11}	a_{11}	a_{11}	b_{11}	a_{11}	b_{11}	1
E_6	b_{11}	a_{11}	a_{11}	b_{11}	b_{11}	b_{11}	1

Here in this case our target is to find the reduct using the strength $E_{noTable} = \{E_1, E_2\}$

$E_{pointless} = \{ E_3, E_4, E_5, E_6\}$ now finding E (l_{11}) $a_{11}(2)$=33%, E((l_{11}) $b_{11}(2)$=Nil similarly finding E(l_{12}) $a_{11}(2)$=100%, E (I_{12}) $b_{11}(2)$=100% ,

E (I_{13}) $a_{11}(1)$ =75%, E(I_{13}) $b_{11}(2)$ =50%, E (l_{14}) $a_{11}(1)$ =33%, E (l_{14}) $b_{11}(2)$=Nil, %,E (l_{15}) $a_{11}(1)$ =66%, E(l_{15}) $b_{11}(2)$=33%,

E (l_{16}) $a_{11}(1)$ =100%

E_1 (l_6) $b_{11}(2)$=25%

Attribute l_{11}, l_{16} produces insignificant i.e extreme result result so we drop l_{11}, l_{16} from the analysis Table-5 to get next Table -5 for subsequent analysis.

REDUCTION DATA

Table -5

E	l_{12}	l_{13}	l_{14}	l_{15}	d
E_1	b_{11}	b_{11}	a_{11}	a_{11}	2
E_2	b_{11}	a_{11}	a_{11}	b_{11}	2
E_3	a_{11}	b_{11}	a_{11}	a_{11}	1
E_4	a_{11}	a_{11}	b_{11}	b_{11}	1
E_5	a_{11}	a_{11}	b_{11}	a_{11}	1
E_6	a_{11}	a_{11}	b_{11}	b_{11}	1

After processing Table-3 the following result reflect that {l_{14}, l_{16}} fields are identical so we take one field out these two field for further processing.

REDUCTION DATA

Table -6

E	l_{12}	l_{13}	l_{14}	l_{15}	d
E1	b_{11}	b_{11}	a_{11}	a_{11}	2
E2	b_{11}	a_{11}	a_{11}	b_{11}	2
E3	a_{11}	b_{11}	a_{11}	a_{11}	1
E4	a_{11}	a_{11}	b_{11}	b_{11}	1
E5	a_{11}	a_{11}	b_{11}	a_{11}	1

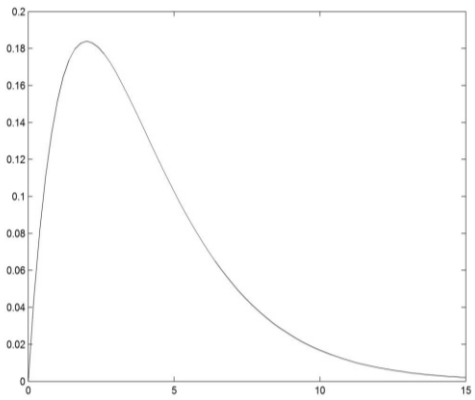

Fig. 1 Chi square

Data Table-4 is not further classified. Rule derived from Table-4 as follows

- L_{12}(insignificant),l_3(insignificant),l_{14}(significant), L_{15}(significant) →NoTable
- L_{12}(insignificant),l_3(significant),l_{14}(significant), L_{15}(insignificant) → NoTable
- L_{12}(significant),l_3(insignificant),l_{14}(significant),l_{15}(significant) →pointless
- L_{12}(significant),l_3(significant),l_{14}(insignificant) L_{15}(insignificant) →pointless
- L_{12}(significant),l_3(significant,l_{14}(insignificant),l_{15}(significant) →pointless

IV. EXPERIMENTAL SECTION

Table-7

C_1/C_2	NoTable	Pointless	Total
Success	50	20	70
Failure	20	4	24
Total	70	24	94

H_0 (Null Hypothesis) - NI does not affected by the symptom Variable body temperature and Low birth wt.
H_1 (Alternate Hypothesis)-NI diseases affected by symptom Variable body temperature and Low birth wt.

$f_{e11}= \frac{(70\times 70)}{94}$, $f_{e12}= \frac{(24\times 70)}{94}$, $f_{21}= \frac{(24\times 70)}{94}$, $f_{22}= \frac{(24\times 24)}{94}$ so we have f_{e11}=52.127, f_{e12}=17.87, f_{e21}=17.87, f_{e22}=6.127

Test of Hypothesis

$(52.127-50)^2/52.127 + (17.87-20)^2/17.87 + (17.87-20)^2/17.87 + (6.127-4)^2/6.127 = 0.0867+0.254+0.254+0.738=1.3327$

This is much below (χ^2 , 0.01, 1) value 0.841, so we reject the null hypothesis accept the alternate hypothesis.

V. CONCLUSION

This paper is deals with a predictive analysis with respect to attribute reduction using both Time series and Rough Set Approach Time Series Result Provide which diseases most no Table pediatric diseases and Rough set approach provide the attribute responsible for the diseases.

REFERENCES

1. Z. Pawlak, Rough Sets, "*International Journal of Computer and Information Sciences"*,Vol.11,pp.341–356, **1982**.
2. S.K. Pal, A. Skowron (Eds.), *"Rough Fuzzy Hybridization"*, Springer, **Berlin**, Vol.21,pp 3-98, **1999**.
3. Z. Pawlak, *"Rough Sets: Theoretical Aspects of Reasoning about Data, System Theory, Knowledge Engineering and Problem Solving*", 9, Kluwer Academic Publishers, Dordrecht, **The Netherlands**, 1991
4. Z. Pawlak, "*Decision Rules, Bayes_ rule and Rough Sets"*, in: N. Zhong, A. Skowron, S. Ohsuga (Eds.), New Direction in Rough Sets, Data Mining, and Granular-Soft Computing, Springer, **Berlin**,pp.1–9,**1999.**
5. L. Polkowski, A. Skowron (Eds.), *"Rough Sets and Current Trends in Computing"*, Lecture Notes in Artificial Intelligence, 1424, Springer, **Berlin**, **1998.**
6. L. Polkowski, A. Skowron (Eds.), *"Rough Sets in Knowledge Discovery"*, 1–2, Physica Verlag, A Springer Company, **Berlin**, **1998**.
7. L. Polkowski, S. Tsumoto, T.Y. Lin (Eds.), *"Rough Set Methods and Applications–New Developments in Knowledge Discovery in Information Systems*", Springer, **Berlin**, **2000**, to Appear.
8. N. Zhong, A. Skowron, S. Ohsuga (Eds.), *New Direction in Rough Sets Data Mining and Granular-Soft Computing*, Springer**, Berlin**, **1999**.
9. R. Vashist, M.L.Garg, *"Rule Generation based on Reduct and Core:A Rough Set Approach*",Vol-29,issue.9 IJCA, **2011**.
10. [10] M. Beaver. Beaver Introduction to Probability and Statistics Cengage 14e.
11. [11] C. Glymour, D.Madigan, D. Pregibon and P. Smith, "*Statistical inference and data mining* ",Communication of the ACM39, Vol.11,pp.35-41,**1996.**
12. [12] T.Hastie, R.Tibshirani, J.H. Friedman, "*Elements of statistical learning: data mining, inference and prediction",* New York: Springer Verlag.
13. [13] D.J. Hand,G. Blunt,M.G. Kelly, &N.M Adams, *"Data mining for fun and profit",* Statistical Science, Vol.15,pp. 111-131,**2000.**
14. [14] D.J. Hand,H. Mannila, & P.Smyth, "Principles of Data Mining", Cambridge, MA: MIT Press, **2001.**
15. G.J. Williams and S.J. Simoff, *"Data Mining Theory, Methodology, Techniques, and Applications"* (Lecture Notes in Computer Science/ Lecture Notes in Artificial Intelligence), Springer, **2006.**

AUTHORS PROFILE

Sipali Prdhan, is a Ph.D scholar in NORTH ODISHA UNIVERSITY, ODISHA, INDIA.

Dr. Jitendu Kumar Mantri, working as professor in the department of computer science (NORTH ODISHA UNIVERSITY), Dr Mantri has many high quality research papers and guided several research scholars , his area of interest is IOT, DATAMINING.

Dr. Sujogya Mishra, is working as a Assistant Professor in the department of Mathematics and Computing in CET, Bhubaneswar, Odisha. Dr. Mishra obtained his Ph.D degree from Utkal University Vani Vihar Bhubaneswar and has more than 15 research papers to his credit.

Dr. Pradyumna Kumar Pattnaik, is working as Assistant Professor in the department of Mathematics *and Computing* in CET, Bhubaneswar, *Odisha*. He obtained his Ph.D degree from Utkal University and has more than 15 published research papers to his credit.