Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2016, 8(12):5-12

Review Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Classification of Disease Data Using Bijective Soft Set

Jagadeesh Gopal[1*], Jai Ganesh K[2] and Salman Ahmed G[2]

[1]*Associate Professor, VIT University,Vellore, Tamilnadu, India*
[2]*M.S Software Engineering, VIT University,Vellore, Tamilnadu, India*

ABSTRACT

Arrangement is a problem in the today's trend world where data should be handling in more accurate way there is lot of software available to store the data in a systematic way but the size of data is too large to keep. There are server problem issues and cost issues. Most of the research is going on big data. Classification of data is also a big problem. When data are not systematic it is very problem to search and analyze the data from a group of data. Soft computing play an important role in classification of data. Artificial Intelligent techniques play a important role in classifying the data. There are many approaches such as neural network, fuzzy set; decision trees and other expect system. To maintain the data in medical field is a huge confidential and classification of data should be done with certain good approach. Artificial Intelligent always play a important role in these fields. In these we are going to use one of the new algorithm for classification is bijective softest. This algorithm helps to give more accuracy then previous traditional methods. We are going to predict the level of diseases such as Kidney cancer, fever, and diagnosis. The software helps as to predict the diseases level accuracy from the data which is classified by using the bijective. It helps to reduce paper work and cost estimation and time for the customer.

Keywords: Bijective soft set; Online prediction of diseases; Decision rules generation; Classification of data

---

INTRODUCTION

In todays world data play a important role for understanding and manipulating any system. But major problem is that data is too large to keep. If we classify the data means the data size can be reduce like uncertain data and data which are outdated can be remove easily. So we can maintain the consistent data. Bijective soft set helps in classifying the data with more accurate results. For classifying the data in medical fields play an important role because if data is not accurate means total system will be collapse. In medical field patient data play a key role.

Doctors need patient update data and previous persistent data which form a certain pattern. This patten may be same for certain patient. So maintaining thousands of patient details. We need to form a rule and pattern so the data size can be. This can be done with the help of Bijective soft set. For checking the level of diseases for kidney, heart diseases, liver diseases. For our system helps in showing the level of diseases to the patient if they have a accurate data which is taken from lab report. Instead of consulting doctors they can interact with our system to know the level. Our system will form a rule depends on current data and rule generation for bijective softest. In this paper we will understand how to form an algorithm from bijective soft set and prediction of diseases from this system.

SIMILAR WORK

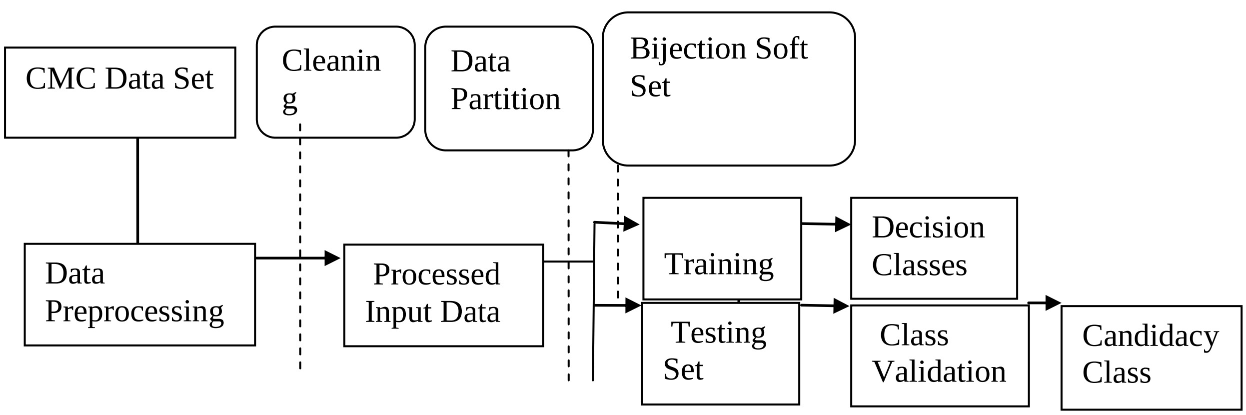
Molodstov first started to work on Soft set theory to solve data problem which is not solved by previous methods He has apply Soft set theory in many practical problems in field such as engineering and medical science. About Ella Hassanien proposed the artificial intelligent techniques by analysis breast cancer based on rough set model. Aktasand Cagman has done in the concept of rough set with the help of softest theory for the purpose of prediction. Maji also proposed a model for decision making with additional studied on softest. Pawlak's and Iwinski"s proposed a model in rought set with according to the soft set rule generation. K.Gong et al have proposed a operation set with the help of soft set.

MOTIVATION FACTOR

The motivation of this paper is started from CMC hospital. Where, huge number of patient visits all over the world. Storing the patient details in the hospital plays a very important role. But the data is too large to store in the server and checking patient details takes lot of time. Same Patient visit often to hospital for different diseases related causes;

But certain patient can have same disease related to one another. Some patient can have same inheritance relation. So classifying the data helps the hospital to search and reducing the sector of data play a very important role. So this helps me to classify the data with an idea of new approach bijective soft set. All the sample data is taken from CMC database. The data was taken from diseases like kidney Cancer, liver diseases, fever and diagnosis etc.

Overall structure



PRELIMINARIES

Soft set theory
Let's understand the basic notions of soft sets approach. Let we can take U be the initial universe of any objects and e can be considered as set of parameters in relation to objects in U. Parameters are often, characteristics, attributes or properties of objects[10].

Definition 1: A pair (f, a) is called a soft set over U, where f is a mapping given by

$$f : a \rightarrow p(U)$$

In other words, a soft set over U is a parameterized family of subsets of the universe U. For $e \in a$, f(e) may be considered as the set of e - elements of the soft set (f, a) or as the set of e - approximate elements of the soft set.

Bijective soft set theory
In overall section U refers to an initial universe, e is a set of parameters, p(U) is the power set of and $a \subseteq e$.[4]

Definition 2: Let (f, b) be a soft set over a common universe U, where f is a mapping f : b →p(U) and b is nonempty parameter set. We say that (f, b) is a bijective soft set, if (f , b) such that
a) Ue∈b f(e) = U.
b) For any two parameters ei, ej b, ei ≠ ej, f(ei)∩f(ej) = ∅
c) In other words, Suppose y ⊆p(U) and y = {f(e1), f(e2), . . . , f(en)}, e1, e2, . . . , en ∈b. From definition 2, the mapping f : b →p(U) can be transformed to the mapping f : b →y, which is a bijective function. i.e. for every y ∈Y, there is exactly one parameter e ∈b such that f(e) = y and no unmapped element remains in both b and y.[1]

And operation
Definition 3: (See [10]) And operation on two soft sets. If (f, a) and (g, b) are two soft sets then ''(f, a)And(g, b)'' denoted by (f , a) ∧(g, b) is defined by (f , a) ∧(g, b) = (h, a × b), where H(α, β) = f(α) ∩g(β), ∀(α, β) ∈a × b.[4]

Definition 4: Suppose that (f, e) and (g, b) are two bijective soft sets over common universe U. (h, c) = (f, e) ∧(g, b) is a bijective soft set.[4]

Restricted and operation
Definition 5: (Restricted And Operation on a Bijective Soft Set and a Subset of Universe). Let U = {o1, o2, . . . , on} be a common universe, x be a subset of U, and (f , e) be a bijective soft set over U. The operation of ''(f , e) restricted AND x'' denoted by
(f , e)∧~x is defined by Ue ∈ e {f(e): f(e) ⊆x}[4].

BIJECTIVE SOFT BASED ON CLASSIFICATION

Bijective soft set based on classification - proposed algorithm
In the new Bijective soft set based classification algorithm, the rules generation is achieved by using And, Restricted And operations in bijective soft set theory. It starts with constructing bijective soft set and apply And, Restricted And operations the finally propose classification rules.

Algorithm Bijective – Soft set Classification (BSCLASS)

Input: Data set D
Output: A set of Rules R
Step 1: Construct bijective soft set for all conditional attributes (fi,ei) for i=1 to n-1, n is the number of attributes.
Step 2: Construct bijective soft set for decision attribute (g, b).
Step 3: Apply AND operation on the bijective soft set (fi,ei). And result to be stored in (H,C).
Step 4: Apply the Restricted AND operation (h, c) and decision attribute (g, b).
"(fi, ei) restricted AND (g, b)" , for i=1 to n-1 and store the result in (f, d)
Step 5: Obtain the decision rules from each bijective soft set (f, d)

Numerical Example

Table 1: Sample Report Data

Object	Attribute-1	Attribute-2	Attribute-3	Decision
Patient-1	10.2	10	10.4	LEVEL-1
Patient-2	11	11.5	11.3	LEVEL-3
Patient-3	11.8	110.6	10.9	LEVEL-2
Patient-4	10.4	11	10.8	LEVEL-1
Patient-5	10.3	10.5	11	LEVEL-1
Patient-6	10	10.8	10.7	LEVEL-1
Patient-7	11.7	11	11	LEVEL-3
Patient-8	11.8	10.5	10	LEVEL-2
Patient-9	10.9	11.7	10.5	LEVEL-1
Patient-10	11.5	11.2	10.3	LEVEL-3

Let's take p1,p2,p3,p4,p5,p6,p7,p8,p9,p10 are patient conditional attribute.

Step 1: Conditional attributes helps to construct bijective softest
(s1, Attribute-1) = {p1, p4, p5, p6, p9} {p2, p3, p7, p8, p10}
(s2, Attribute-2) = {p1, p3, p5, p6, p8} {p2, p4, p7, p9, p10}
(s3, Attribute-3) = {p1, p3, p4, p6, p8, p9, p10} (p2, p5, p7}

Step 2: Decision attributes helps to construct bijective softest
Decision (1) = {p1, p4, p5, p6, p9}
Decision (2) = {p3, p8}
Decision (3) = {p2, p7, p10}

Step 3: Apply and operation on the bijective soft set (fi, ai). Table 2 shows tabular form of this And operation.

Table 2: Tabular Form of (s1, p1) ^ (s2, p2) ^ (s3, p3)

Objects	R1	R2	R3	R4	R5	R6	R7	R8
P1	Yes	No	Yes	No	No	No	No	No
P2	No	No	No	No	No	Yes	No	Yes
P3	No	No	No	No	Yes	No	Yes	No
P4	No	Yes	Yes	No	No	No	No	No
P5	Yes	No	No	Yes	No	No	No	No
P6	Yes	No	Yes	No	No	No	No	No
P7	No	No	No	No	No	Yes	No	Yes
P8	No	No	No	No	Yes	No	Yes	No
P9	No	Yes	Yes	No	No	No	No	No
P10	No	No	No	No	No	Yes	Yes	No

Figure: Yes=match, No=No match

Where R1 = {P1, P4, P5, P6, P9} ∧ {P1, P3, P5, P6, P8} = {P1, P5, P6}
R2 = {P1, P4, P5, P6, P9} ∧ {P2, P4, P7, P9, P10} = {P4, P9}
R3 = {P1, P4, P5, P6, P9} ∧ {P1, P3, P4, P6, P8, P9, P10} = {P1, P4, P6, P9}
R4 = {P1, P4, P5, P6, P9} ∧ {P2, P5, P7} = {P5}
R5 = {P2, P3, P7, P8, P10} ∧ {P1, P3, P5, P6, P8} = {P3, P8}
R6 = {P2, P3, P7, P8, P10} ∧ {P2, P4, P7, P9, P10} = {P2, P7, P10}
R7 = {P2, P3, P7, P8, P10} ∧ {P1, P3, P4, P6, P8, P9, P10} = {P3, P8, P10}
R8 = {P2, P3, P7, P8, P10} ∧ {P2, P5, P7} = {P2, P7}

(f1, a1) ∧ (f2, a2) ∧ (f3, a3) = (h, c)
(h, c) = {{P1, P5, P6}, {P4, P9}, {P1, P4, P6, P9}, {P5}, {P3, P8}, {P2, P7, P10}, {P3, P8, P10}, {P2, P7}}

Step 4: Apply the restricted And operation (h, c) and decision attribute (g, d).
(h, c) ∧~d(1) = {{P1, P5, P6}, {P4, P9}}
(h, c) ∧~d(2) = {P3, P8}
(h, c) ∧~d(3) = {P2, P7, P10}
(f, b) = {{P1, P5, P6}, {P4, P9}, {P3, P8}, {P2, P7, P10}}

Step 5: Generating decision rules based on restricted And result (f, b).
The following deterministic decision rules are extracted:
If Attribure-1 = 10.0 to 10.9 and Attribure-2 = 10.0 to 10.9=>decision = 1
If Attribure-1 = 10.0 to 10.9 and Attribure-2 = 11.0 to 11.9=>decision = 1
If Attribure-1 = 11.0 to 11.9 and Attribure-2 = 10.0 to 10.9=>decision = 2
If Attribure-1 = 11.0 to 11.9 and Attribure-1 = 11.0 to 11.9=>decision = 3

BIJECTIVE SOFT SET BASED ON CLASSIFICATION

Classification is a data mining technique used to predict group membership for data instances. The classification algorithm learns from the training set and builds a model. The model is used to classify new objects. In this work, classification accuracy of the proposed approach is compared to two different classifiers Naïve Byes and decision table used for computing the Precision, Recall and F-measure.

Data description

Data sets are taken from the CMC hospital where classifying the data is a huge problem.
Diseases data such as fever, kidney cancer, liver diseases and other diagnosis data sets. For experimental analysis we have taken this data set for practical propose.

VALIDATION MEASURES

Precision

It is defined by accuracy provided that a specific class has been predicted.

$$\text{Precision}=tp / (tp+fp)$$

Where tp and fp are the numbers of true positive and false positive respectively;

Recall

Recall is also known as sensitivity. It is measure of the ability of a prediction model to select instance of a certain class from a data sets. Its value is corresponds to the true positive rate.

$$\text{Recall (or) Sensitivity} = tp / (tp + fn)$$

F-measure

It is defined as combining precision and recall is the harmonic mean of precision and recall, the traditional F-measure or balanced F-score of

$$\text{F-Measure} = (2*\text{Precision}*\text{Recall}) / (\text{Precision} + \text{Recall})$$

Table 3: Performance analysis of classification algorithms on breast cancer data set

Accuracy Measures	Classification Algorithm		
	Bijective soft set based classification	**Decision Table**	**Naïve Bayes**
Precision	0.98	0.44	0.9
Recall	1	0.667	0.889
F-Measure	0.99	0.533	0.882

Table 4: Performance analysis of classification algorithms on Pima Indian diabetes data set

Accuracy Measures	Classification Algorithm		
	Bijective soft set based classification	**Decision Table**	**Naïve Bayes**
Precision	0.96	0.78	0.84
Recall	0.92	0.64	0.73
F-Measure	0.94	0.71	0.65

Tables 3, 4 and 5 helps to predict performance analysis of the classification algorithm on breast cancer from other two dataset such as Pima Indians diabetes and liver-disorders dataset; Bijective soft set performs better approaches then other two approaches.

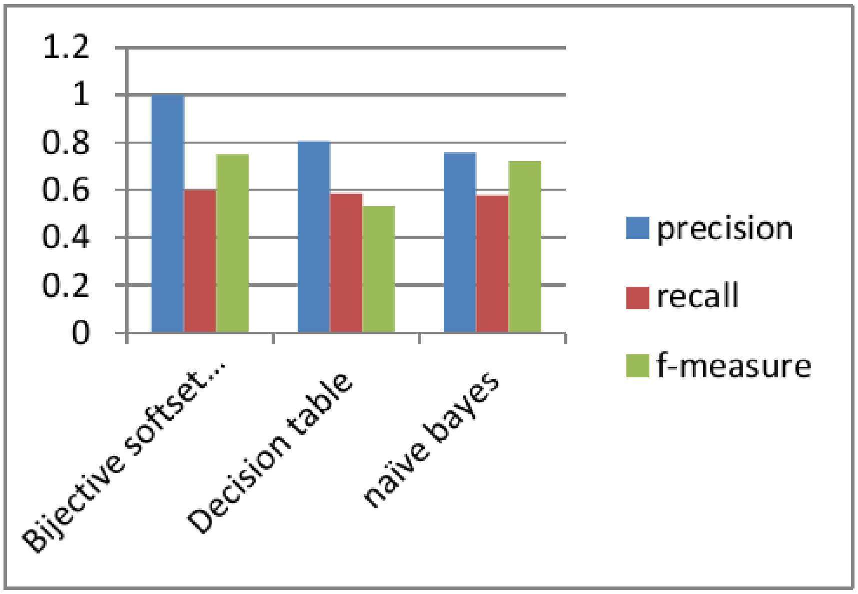


Figure 1: Comparative analysis of classification algorithms for breast cancer data set

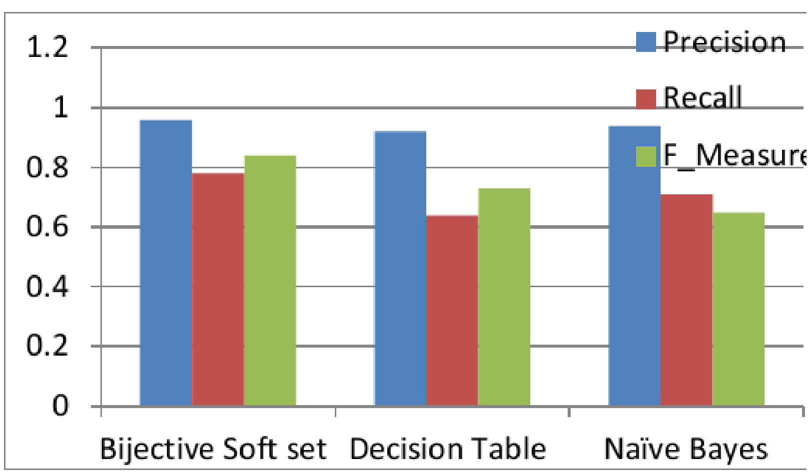


Figure 2. Comparative analysis of classification Algorithms for Pima Indians Diabetes dataset

Figure 2: Comparative analysis of classification algorithms for Pima Indian diabetes data set

Table 5: Components and their classifications with normal values

Attributes	Normalized value	Classification
Haemoglobin (HB)	09-Dec	Low (1)
	07-Sep	Medium (2)
	05-Jul	High (3)
	< 5	Very high (4)
Serum urea (SU)	40-80	Low (1)
	80-120	Medium (2)
	120-200	High (3)
	>200	Very high (4)
Serum creatine (SC)	01-Mar	Low (1)
	03-May	Medium (2)
	05-Oct	High (3)
	> 10	Very high (4)
Protein (Prot)	0.031-0.99	Low (1)
	1-1.99	Medium (2)
	2-2.99	High (3)
	> 3	Very high (4)
Weight loss (WL)	5%-7%	Low (1)
	7%-12%	Medium (2)
	12%-15%	High (3)
	>15%	Very high (4)
High blood pressure (HBP)	120 to 139 / 80 to 90	Normal (1)
	140 to 159 / 90 to 99	Medium (2)
	160 to 179 / 100	High (3)
	to 109	Very high (4)
	>180 / 110	
Abdominal pain (AP)		High (1)
	----	Medium (2)
		Low (3)
Colour of urine (COU)		Reddish yellow (1)
		Pale yellow (2)
	----	Colour less (3)
Level of Kidney cancer (KCL)		Level (1)
		Level (2)
	----	Level (3)
		Level (4)

Table 6: Patients' component values

Patient	HB	SU	SC	Protein	WL	HBP	AB	COU	KCL
1	6	126	2	2.55	13	150/95	2	2	1
2	8	44	6	1.33	6	135/80	3	1	2
3	11	135	8	1.99	12	130/90	1	2	2
4	7	90	4	0.031	5	125/80	1	1	1
5	6.5	85	7	2.27	11.5	140/98	1	2	2
6	7.5	55	8	1.25	8	120/80	3	1	2
7	12	60	6	1.58	9	138/90	2	1	2
8	9	115	5.5	0.136	15	165/105	3	3	1
9	5	175	6.5	0.156	11.5	150/95	1	3	2
10	6.5	95	7.5	0.5	13	140/90	2	1	3

CONCLUSION

Classification plays an important role in today's data formation. Bijective soft set helps them to classify the data with help of rules. In future Big data and digital technology play a very important role, So maintaining a data is a huge problem and removing the data which is outdated is most important and classifying it play a very important role for searching, modifying, classifying. These helps to reduce cost, time and size of data. So in digital world bijective soft set can play a very important role in future publication.

REFERENCES

[1] AE HAssanien. Intelligent data analysis of breast cancer based on Rough set theory. *International Journal on Artificial Intelligence Tools* **2003** 12(4): 465-479.
[2] Aktaş.H, Çagman.N. Soft sets and soft groups. *Information Science*. **2007** 177 (13): 2726–2735.
[3] Bana Handaga and Mustafa Mat. Similarity Approach on Fuzzy Soft Set Based Numerical Data Classification", ICSECS, Part II,CCIS **2011** 180: 575–589.
[4] Gong K et al. The bijective soft set with its operations. Computers & Mathematics with Applications **2010.**
[5] Herawan.T, Deris.M.M. A direct proof of every rough set is a soft set", *Third Asia International Conference on Modelling & Simulation* **2009** 4: 119–124.
[6] Ivo D and Gunther G. Statistical Evaluation of Rough Set Dependency Analysis. *International Journal of Human-Computer Studies* **1997** 46: 589- 604.
[7] Kalaiselvi.N, Hannah Inbarani.H. Fuzzy Soft Set Based Classification for Gene Expression Data. *International Journal of Fuzzy Set Theory* **2003** 1(2): 23-34.
[8] Maji.P.K, Biswas.R and Roy.A.R. Fuzzy Soft Sets. *Journal of Fuzzy Mathematics*. **2001** 9: 589-602
[9] Maji PK, Roy A, Biswas R. An application of soft sets in a decision making problem. *Computer Mathematical Application* **2002** 44: 1077–1083.
[10] Molodtsov D. Soft set theory-first results. *Computers & Mathematics with Applications*. **1999** 37: 19–31.
[11] D. Pei, D. Miao, "Soft sets to information systems. Granular Computing, IEEE International Conference, **2005**.