iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 4 **Issue:** VI **Month of publication:** June 2016

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Fuzzy Association Rule by Classification Technique

Subhangi Gupta[1], Rohit Miri[2], Asha Miri[3], S R Tandan[4]
[2,4]Department of Computer Science & Engineering, Dr. C.V. Raman University, Kota, Bilaspur, India (C.G.)
[1]M.Tech scholar,Department of Computer Science & Engineering, Dr. C.V. Raman University, Kota, Bilaspur, India (C.G.)
[3]Department of Computer Science & Engineering, Govt. Polytechnic College, Bilaspur (C.G)

***Abstract*—When we use the fuzzy association rule then it is difficult to find the minimum frequency because of distinct fuzzy values of data sets. To resolve such type of problem the classification technique is used in a data set. We put fuzzy values to different classes and the no of classes act as frequency of data sets. This technique provides frequency easily for fuzzy association rule.**

***Keywords*—Classification , Data Mining , Triangular Membership function**

I. INTRODUCTION

Now days the researchers have find the different approach for solving the fuzzy association rule. The fuzzy association rule is very important for providing the data sets for the business purpose. In this approach we defined ten classes numbered from A to I. We defined the range of each classes as a single unit. The membership function takes values from 0.0 to 1. This interval is divided into 10 parts. Each of the classes has range .1 units.

Now a day's fuzzy logic and data mining are the two most important techniques on the field of research area. Fuzzy logic uses linguistic variable to provide the answer that human being are not able to give. Whereas data mining is used to find out the essential things from large data sets ,then apply some association data mining algorithm to find the relationships among the item sets. This research paper find out the range of linguistic variable by using standard deviation and mean in place of assuming its values. Secondly on fuzzy logic contains the distinct values by applying membership function. When we apply data mining association algorithm on it, it is difficult to find out the essential things from large data sets because of distinct fuzzy values will not get the minimum frequency of item sets.

Advantages of this Research are as follows:

A. More fruitful for any business prediction process.
B. It is easy to get the minimum frequency for mining algorithm

II. LITERATURE REVIEW

Mohammed Al-Maoleg1[1] et al chooses those item sets that are frequently present in particular transaction id. Its algorithm works on low support .3 and reduces the time complexity of the program very easily. Zhiyong ma [4] et al converts all the item sets into Boolean matrix by using CP tree method and reduces the time for the task.ArpnaShrivastava [5] et al , in this paper the authors have used the codes for all the items and remove the duplication by using data cleansing technique. This is also most efficient as compared to simple Apriori algorithm. K. Sathesh Kumar and M. Hemalatha [3] , this paper reduces the operational time carried out by Apriori algorithm by using artificial Bee colony optimization method (FABCO).

As we know we need some kind of association rule to perform data mining algorithm. Getting this Agrawal and his co-worker carried out some mining algorithm based on the large data sets, which is difficult to find association mining rule [9-18]. These break the mining steps into two phases. In the first phase candidate of item sets are obtained and counted by scanning the transactions. The number of item set must support the minimum pre-defined threshold value called minimum support. Then later we make the pair of item sets and apply the association rule for getting the required output. Srikant and Agrawal also proposed partitioned based mining association algorithm. Most of them are find out the range of triangular fuzzy membership function directly, means they assumed the range of linguistic variable. But on Our paper we have find out the range of linguistic variable by using mean and standard deviation.

The research papers [9], [15] are based on the fuzzy association rule. They have done on fuzzy mining association rule to reduce the computational time. They all used the simple mining association rule for doing the task, the TRApriori mining association technique is used from the paper [16].

Volume 4 Issue VI, June 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Jageshwer shriwas et all [21] also done on stock prediction based on Apriori algorithm. They suggested the stock prediction by Data Mining Techniques.

III. PROPOSED ALGORITHM

STEP-1: Find out the range of linguistic variable from the given data set as low, middle and high.
STEP-2: Find out the membership function of given data sets.
STEP-3: Find out the fuzzy values of all the attributes.
STEP-4: Define the classes for fuzzy values named from A to J. Take internal of one unit for each class.
STEP-5: Put the classes to item sets.
STEP-6: i) Select those item sets which has having maximum no of cases with repetition allowed.
ii) If both the item sets have equal no of classes then select those attributes whose sum of fuzzy values is high.
iii) Select those classes which have maximum fuzzy values.
iv) Select those item sets which has having minimum of six classes or sixty percent.
STEP-7: Obtained data set is an important or final data set and applies the fuzzy mining association rule.
STEP-8: Define the minimum threshold =.2 and minimum confidence of .70.
STEP-9: Find out the L1 tem sets.
STEP-10: Find out the C1 item Sets.
STEP-11: Repeat the step from 9 to 10.
STEP-:12 find out the support and confidence of obtained data sets.

IV. DATA ANALYSIS

The first step is to take an exemplary data set of 10 students with having their grades of different five subjects and details regarding their marks are shown in table no 1.1
First Subject: UNIX
Second Subject: CG (Computer Graphics)
Third Subject: SE (Software Engineering)
Fourth Subject: TOC (Theory Of Computation)
Fifth Subject: CN (Computer Network)

S. N	UNIX	CG	SE	TOC	CN
1	85	78	87	72	69
2	60	80	88	77	81
3	83	90	87	80	88
4	72	87	80	83	63
5	69	88	86	81	77
6	64	78	85	62	86
7	66	88	76	72	81
8	87	64	65	83	85
9	76	64	80	86	87
10	80	62	64	84	88

Table No 1.1. The set of students' course scores

Volume 4 Issue VI, June 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

STEP 1: Find out the range of linguistic variable from the given data set as low, middle and high that is shown in table No 1.2.

S. N	Attribute	low	Middle	High
1	Data Sets	(60,70,74)	(70,78,85)	(82,85,90)

Table No 1.2 Range of Linguistic Variable

STEP 2: Find out the membership function of given data sets.
The membership function is carry out by using triangular fuzzy logic number model .Shown in Figure No 1.1

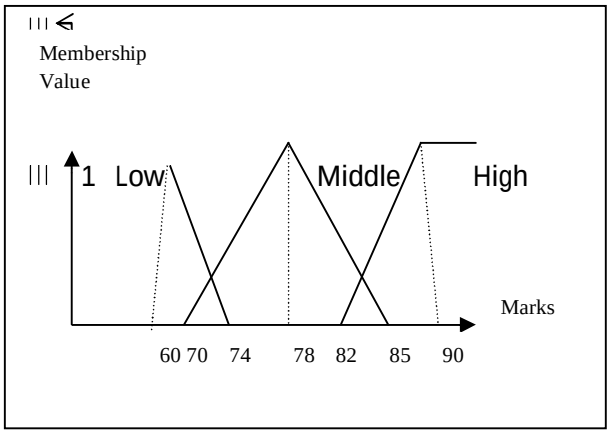


Fig.No1.1 Triangular Membership Function of Data Sets

STEP 3: Find out the fuzzy values of all the attributes.
In this, we have used triangular membership function because of its easiness and computational efficiency. We can also use Gaussian membership function.

SN	1	2	3	4	5	6	7	8	9	10	sum
UNL	.0	1	.0	.4	.6	.8	.7	.0	.0	.0	
	.2	.4	.3	.0	.0	.0	.0	.0	.8	.8	
IMXH	.7	.0	.1	.0	.0	.0	.0	.8	.0	.0	
CL	.0	.0	.0	.0	.0	.0	.0	.8	.8	.9	
	1	.8	.0	.0	.0	1	.0	.0	.0	.0	
MGH	.0	.0	1	.8	.9	.0	.9	.0	.0	.0	
SL	.0	.0	.0	.0	.0	.0	.0	.7	.0	.8	
	.4	.9	.0	.8	.1	.2	.8	.0	.8	.0	
EMH	.8	.9	.8	.0	.6	.7	.0	.0	.0	.0	
TL	.4	.0	.0	.0	.0	.9	.4	.0	.0	.0	
O	.3	.9	.8	.4	.1	.0	.0	.4	.1	.2	

Volume 4 Issue VI, June 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

M C H	.3	.0	.0	.1	.0	.0	.0	.1	.6	.5	
C	.5	.0	.0	.0	.8	.0	.0	.0	.0	.0	
L	.0	.8	.0	.0	.9	.1	.8	.2	.0	.0	
N M H	.0	.0	.9	.0	.0	.6	.0	.7	.8	.9	

Table No 1.3: The fuzzy set transformation Table.

We can categories it's as Low, Middle and high. Thus we have used three fuzzy membership values produced for each attributes. Transform the actual values of each attribute into fuzzy sets. Take UNIX marks in case 1 as an example. The marks 85 is replaced by a fuzzy set (.0/low + 0.2/middle + .7 / high) Using the given membership functions. This step is repeated for all the data sets.

STEP: 4 Define the classes for fuzzy values named from A to J. Take internal of one unit for each class.

S.N	Classes	Range
1	A	0.01-0.10
2	B	0.11-.20
3	C	0.21-0.30
4	D	0.31-0.40
5	E	0.41-0.50
6	F	0.51-0.60
7	G	0.61-0.70
8	H	0.71-0.80
9	I	0.81-0.90
10	J	0.91-1

Table 1.4 The fuzzy classification table for fuzzy values

STEP: 5 Put the classes to item sets.

N	**TName**	**Classes**
1	UNIX(L)	D,F,G,H,J
2	UNIX(M)	B , C,D,H
3	UNIX(H)	A , G ,H
4	CG(L)	H , I
5	CG(M)	A,A,H
6	CG(H)	A, H, I, I
7	SE(L)	G,H
8	SE(M)	A,B,D,H,H,H,I
9	SE(H)	F,G,H,H,I
10	TOC(L)	A, A, B, C
11	TOC(M)	A,A,B,C,D, H,I
12	TOC(H)	A,A,C,E,F
13	CN(L)	E,H
14	CN(M)	A, B, H, H, I
15	CN(H)	F,G,H,I,I

Table No 1.5 Classification Table for Data Items

STEP:6 :i) Select those item sets which has having maximum no of cases with repetition allowed.

ii) If both the item sets have equal no of classes then select those attributes whose sum of fuzzy values is high.

iii) Select those classes which have maximum fuzzy values.

iv) Select those item sets which has having minimum of six classes or sixty percent.

STEP: 7 Resulted data set is an important or final data set and applies the fuzzy mining association rule. we eliminate the unimportant data set with the help of fuzzy classification technique.

S.N	Tname	Classes
1	UNIX(L)	D,F,G,H,J
2	SE(M)	A,B,D,H,H,H,I
3	SE(H)	F,G,H,H,I
4	TOC(M)	A,A,B,C,D, H,I
5	TOC(H)	A,A,C,E,F
6	CN(H)	F,G,H,I,I

Table No 1.6 The actual Data Set

STEP 8 Define the minimum Support =.2 and minimum confidence of .70.

STEP: 9 Find out the large Item set L1

L1		
S.N	Item set	Frequency
1	A	5
2	B	2
3	C	2
4	D	3
5	F	4
6	G	3
8	H	8
9	I	5

Table 1.7.L1 Item sets

STEP 10: find out the candidate item set C1.
Find the C1 according to their classified based value on the data set in Table No 1.6.

C1		
S.N	TName	Classes Or Item sets
1	UNIX(L)	{D},{ F}, {G} , {H}
2	SE(M)	{A},{ B}, { D}, {H} ,{I}
3	SE(H)	{F},{ G}, { H}, {I}
4	TOC(M)	{{A},{ B}, { C}, {D} , {H}, {I}
5	TOC(H)	{A},{ C}, {F}
6	CN(H)	{F},{G} ,{ H}, {I}

Table No 1.8 .C1 Item sets

STEP: 11find the large item set L2

Volume 4 Issue VI, June 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

L2		
SN	Classes	Frequency
1	{A , B}	2
2	{A , D}	2
3	{A , H}	2
4	{A , I}	2
5	{B , D}	2
6	{B , H}	2
7	{B, I}	2
8	{D, H}	3
9	{D, I}	2
10	{F, G}	3
11	{G, H}	3
12	{H, I}	3

Table No 1.9 .L2 Item sets

STEP: 12 Find the candidate item set C2

C2		
S.N	TName	Classes Or Item sets
1	UNIX(L)	{D},{ F}, {G} , {H}
2	SE(M)	{A},{ B}, { D}, {H} ,{I}
3	SE(H)	{F},{ G}, { H}, {I}
4	TOC(M)	{{A},{ B}, { C}, {D} , {H}, {I}
5	CN(H)	{F},{G} ,{ H}, {I}

Table No 1. 10 .C2 Item sets

STEP: 13 Find the Large Item set L3

L3		
SN	Item set	Frequency
1	{A , H ,I}	2
2	{F , G ,H}	3
3	{G , H , I}	2

Table No 1.11 .L3 Item sets

STEP: 14 Find the candidate item set C3

C3		
SN	TNAME	CLASSES
1	SE(H)	{F},{ G}, { H}, {I}
2	CN(H)	{F},{ G}, { H}, {I}
3	SE(H)	{{A},{ B}, {D} , {H}, {I}
4	TOC(M)	{{A},{ B}, {D} , {H}, {I}

Table No 1.12 C2 Item sets

STEP: 15 Find the Large Item set L4

Volume 4 Issue VI, June 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

L4		
SN	Item set	Frequency
1	{A , B ,D, H}	2

Table No 1.12 .L4 Item sets

STEP: 16 Find the candidate item set C4

C4		
SN	TNAME	CLASSES
3	SE(H)	{{A},{ B}, {D} , {H}, {I}
4	TOC(M)	{{A},{ B}, {D} , {H}, {I}

Table No 1.13 . C4 Item sets

STEP: 17 Find the candidate item set C4 we will take the lesser membership value when we compare the two item sets

S.N	SE(H)	TOC(M)	SE(H),TOC(M)
1	.8	.3	.3
2	.9	.9	.9
3	.8	.8	.8
4	.0	.4	.0
5	.6	.1	.1
6	.7	.0	.0
7	.0	.0	.0
8	.0	.4	.0
9	.0	.1	.0
10	.0	.2	.0
sum	3.8	3.30	2.10

Table 1.14: The C2 candidate item set

STEP: 22 Find the L2Large item set
The linguistic value (SE(H)), TOC(M)) has the scalar cardinality of 2.10. The Large item set L2 shown below. The L2 large item set rides on the count value, which is greater than the minimum support.
{ (SE(H), TOC(M) 2.10)}

STEP 23: (a) Construct the association rules for all the large item set .there are three possible association rules.

SE(H) , TOC(M)/ TOC(M) = 2.10 /3.2

TOC(M), SE(H)/ SE(H) = 2.10 /3.8

The fuzzy association rules are listed on table 1.15

S.N	Fuzzy Association Rule	Confidence Value
1	SE(H)→TOC(M)	70
2	TOC(M)→ SE(H)	55

Table No: 1.15 Summary of Association rule

(b)We can also find the confidence of the entire rule. Suppose our minimum threshold is .70 for confidence. Its confidence value is calculated as:

The confidence values of the other tworuleare shown below.

"If SE = HIGH, then TOC = Middle" has a confidence value of 0.70;

"If TOC = Middle, then SE = HIGH" has a confidence value of 0.55;

V. CONCLUSION

This will produce the Fuzzy association rule by applying the classification technique.
We will easily get the frequency of fuzzy values.

VI. FUTURE SCOPE

This technique may be useful for big data.
For future work, we can predict the result of association rule without applying the Data Mining association algorithm based on the fuzzy classification technique.

VII. ACKNOWLEDGMENT

We would like to express our sincere thankfulness to our guide Dr Rohit Miri, for their helpful support and valuable advice during our research work.

REFERENCES

[1] MOHAMMED AL-MAOLEGI ,BASSAMARKOK "AN IMP ROVED APRIORIALGORITHM FOR ASSOCIATION RULES" INTERNATIONAL JOURNAL ON NATURAL LANGUAGE COMPUTING VOLUME 3 , NO-1 , FEBUARY 2014.

[2] N S NITHYA AND K DURAISWAMI "GAIN RATIO BASED FUZZY WEIGHTED ASSOCIATION RULE MINING CLASSIFIER FOR MEDICAL DIAGNOSTIC INTERFACE" VOL 39 PART 1 ,FEBUARY 2014 , PP 39-52 .

[3] K SATHESH KUMAR AND M HEMLATHA "AN INNOVATIVE POTENTIAL ON RULE OPTIMIZATION USING FUZZY ARTIFICIAL BEE COLONY" RESEARCH JOURNAL OF APPLIED SCIENCE , ENGINEERING AND TECHNOLOGY.. VOLUME 2627-2633 , ISSN:2040-7459, E-ISSN:2040-7467, APRIL 05 2014.

[4] ZHIYONG MA , QINGYINGQIU AND PEIEENFENG "CPM ALGORITHM FOR MINING ASSOCIAOTION RULES FROM DATABASES OF ENGINEERING DESIGN+N INSTANCES." JOURNAL OF INFORMATION SCIENCE AND ENGINEERING VOL30 , 463-481 (2014)

[5] ARPNASHRIVASTAVA , R . C JAIN AND A K SHRIVATAVA "GENERATING 3RD LEVEL ASSOCIATION RULE USING FAST APRIORIIMPLEMENTATION."BRITISH JOURNAL OF MATHEMATICS AND COMPUTER SCIENCE. VOLUME 4(2):241-251, 2014

[6] NEELUKHARE ,NEERUADLAKHA, K R PARDASANI "AN ALGORITHM FOR MINING MULTIDENTIONAL FUZZY ASSOCIATION RULE" INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND INFORMATION SECURITY VOL 5, NO-1 ,2009

[7] THOMAS SUDKOMP "REFINEMENT OF TEMPORAL CONSTRAINTS IN FUZZY ASSOCIATION " VOLUME 48(2008) 14 JUNE 2007

[8] TIBERIU BAN , FUZZY COMPUTING FOR COMPLEXITY LEVEL OF EVALUATION TESTS. STUDIAUNIV, BABES-BOLAI, INFORMATICS, VOLUME LVIII NUMBER 1, 2013

[9] TZUNG-PEI HONGA,CHAN-SHENG KUOB,SHYUE-LIANG WANGC "A FUZZY APRIORITID MINING ALGORITHM WITH REDUCED COMPUTATIONAL TIME" VOLUME 5, ISSUE 1, DECEMBER 2004, PAGES 1–1

[10] TZUNG-PEI-HANG, GUO-CHENG LAN, YI-HSINLIN AND SHING-TAI PAN. AN EFFECTIVE GRADUAL DATA REDUCTION STRATEGY FOR FUZZY ITEM SET MINING INTERNATIONAL JOURNAL OF FUZZY SYSTEMS ,VOL 15 NO2 JUNE 2013.

[11] USHA RANI, R VIJAYAPRAKKASH, DR A. GAVARDHAN,MINING MULTI-LEVEL ASSOCIATION RULES USING FUZZY, INTERNATIONAL JOURNAL OF EMERGING TECHNOLOGY AND ADVANCED ENGINEERING. VOLUME 3, ISSUE 8, AUGUST 2013.

[12] RUCHIBHARGAVA ,SHRIKANT LADE , EFFECTIVE POSITIVE NEGATIVE ASSOCIATION RULE MINING USING IMPROVED FREQUENT PATTERN TREE. INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN COMPUTER SCIENCE AND SOFTWARE ENGINEERING.VOLUME3 , ISSUE 4 APRIL 2013 .

[13] DHARMENDRAKUMAR , DEEPAK BHARDWAJ RISE OF DATA MINING: CURRENT AND FUTURE APPLICATION AREAS. IJCSI INTERNATIONAL JOURNAL OF COMPUTER SCIENCE ISSUE VOL 8 ISSUE 5 NO1 SEPTEMBER 2011.

[14] KANUPATEL ,VATSALSHAL , JITENDRAPATEL , JAYNA DONGA COMPARISION OF VARIOUS ASSOCIATION RULE MINING ALGORITHM ON FREQUENT ITEM SETS. INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN COMPUTER SCIENCE AND SOFTWARE ENGINEERING VOL-3 ISSUE II NOV 2013 .

[15] JRSHIANCHEN HUNG LIEHCHOU, CHINGHSUE CHENG, JEN-YA WANG CPDA BASED FUZZY ASSOCIATION RULES FOR LEARNING ACHIEVEMENT MINING, 2009 INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND COMPUTING IPCSIT VOL3 (2011) IACSIT PRESS SINGAPUR.

[16] E RAMARAJ, K RAMESHKUMAR, N VENKATESAN"A BETTER PERFORMED TRANSACTION REDUCTION ALGORITHM FOR MINING FREQUENT ITEM SET FROM LARGE VOLUMINOUS DATABASE"VOLUME 5, ISSUE 1, DECEMBER 2004, PAGES 1–10

[17] R. AGRAWAL, T. IMIELINKSI AND A. SWAMI, "MINING ASSOCIATION RULES BETWEEN SETS SUBMITTED MANUSCRIPT 24 OF ITEMS IN LARGE DATABASE," THE 1993 ACM SIGMOD CONFERENCE, WASHINGTON DC,USA, 1993.

[18] R. AGRAWAL, T. IMIELINKSI AND A. SWAMI, "DATABASE MINING: A PERFORMANCE PERSPECTIVE," IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 5, NO.6, 1993, PP. 914-925.

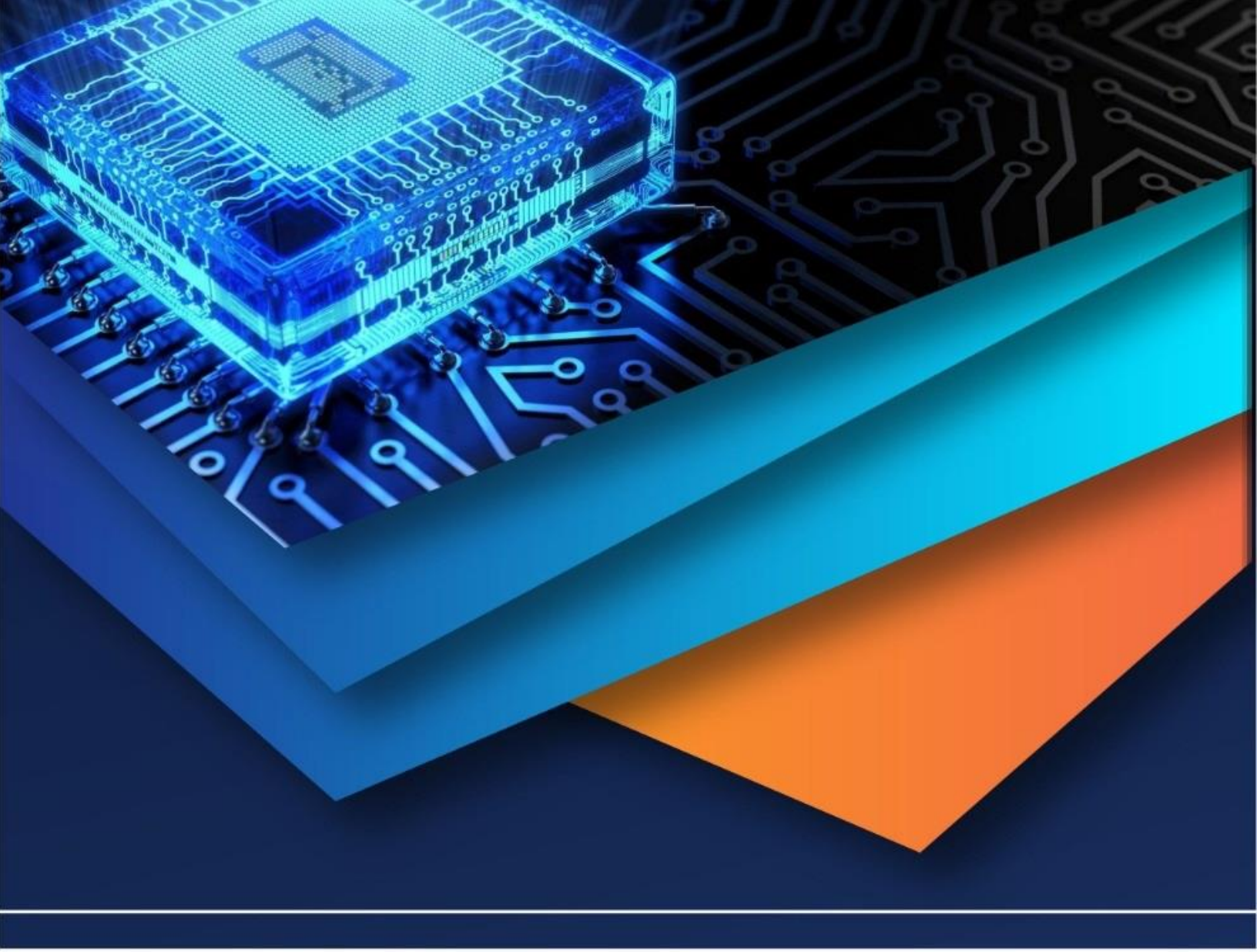
[19] R. AGRAWAL, R. SRIKANT AND Q. VU, "MINING ASSOCIATION RULES WITH ITEM CONSTRAINTS," THE THIRD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY IN DATABASES AND DATA MINING, NEWPORT BEACH, CALIFORNIA, AUGUST 1997.

[20] R. AGRAWAL AND R. SRIKANT, "FAST ALGORITHM FOR MINING ASSOCIATION RULES," THE INTERNATIONAL CONFERENCE ON VERY LARGE DATA BASES, 1994, PP. 487-499.

[21] JAGESHWERSHRIWAS, SHAGUFTAFARZANA," USING TEXT MINING ANDRULED BASED TECHNIQUEFOR PREDICTION OF STOCK MARKET PRICE" INTERNATIONAL JOURNAL OF EMERGING TECHNOLOGY AND ADVANCED ENGINEERING (ISSN 2250-2459, ISO 9001:2008 CERTIFIED JOURNAL, VOLUME 4, ISSUE 1, JANUARY 2014

[22] JAGESHWERSHRIWAS, SHAGUFTAFARZANA," PREDICTION OF STOCK MARKET PRICE USING CLASSIFICATION RULES AND GRAPH BASED ANALYSIS", CIIT DATA MINING AND KNOWLEDGE ENGINEERING, VOL 5, NO 7 (2013)

[23] JAGESHWERSHRIWAS, DR SAMIDHADWIVEDI SHARMA," STOCK PRICE PREDICTION USING HYBRID APPROACH OF RULE BASED ALGORITHM AND FINANCIAL NEWS" INT.J.COMPUTER TECHNOLOGY &APPLICATIONS,VOL 5 (1),205-211,JAB-FEB 2104.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY