FUZZY SOFT SET BASED RAIN FALL PATTERN MINING FROM NON-STATIONARY STREAM DATA

Sudhanshu Kumar, G.S.Thakur
Department of Mathematics &Computer Applications,
Maulana Azad National Institute of Technology, Bhopal,M.P. INDIA
Email: ksudhanshu20@gmail.com, ghanshyamthakur@gmail.com

Abstract:
This paper presents a Fuzzy Soft set based model for forecasting. Four districts of Madhya Pradesh state viz. Bhopal, Gwalior, Indore and Jabalpur are selected for the rainfall forecasting. The results shows accuracy and efficiency of proposed Fuzzy Soft approach outperforms.

Keywords: Fuzzy soft set; Pattern Mining; Stream Data, Rainfall Pattern Forecasting
DOI Number: 10.48047/nq.2022.20.22.NQ10101 **NeuroQuantology 2022; 20(22):1244-1252**

1. Introduction

The data are growing exponentially every day in science, engineering, economics, sociology, medical science, environmental science, and many more fields. A challenging issue to deal with the complexities of modeling the uncertainties of rainfall data. Rough set[8], Fuzzy set[10] and probability theory are used for handling the uncertainties of data.

Rainfall Pattern Forecasting is a very challenging task in the world. Rainfall pattern forecasting is a challenging issue to the Government and Societies. This problem motivates us to do the research work in this area. This research work is based on new approach Fuzzy Soft set theory . Four districts of Madhya Pradesh state viz. Bhopal, Gwalior, Indore and Jabalpur are selected for the Rainfall Pattern Forecasting.

Fuzzy sets

The problem of forecasting is related with decision making, where the cloud cover, humidity, wind, temperature and other related parameters are responsible for the rainfall. In the rainfall cases the available information are often ambiguous, indistinct and vague. Examples are climate, weather condition, wind, session, humidity, cloud cover, tempreture etc. The above example shows that diverse uncertainties are present in the parameters of the rainfall cases. The accurate Rainfall Pattern Forecasting model is difficult to develop using the classical approaches.

L.A. Zadeh proposed Fuzzy set theory[10], in 1965 as an extension of classical set theory. The applications of fuzzy set theory has been widely used in the various areas like: System control , aircraft control, data analysis and pattern recognition. Fuzzy set theory could be used as a powerful tool to handle uncertain and imprecise data. Fuzzy approach is useful in real life decision making problem like rainfall.

Definition1. [10]: Let X be a nonempty set. A fuzzy set A in X is characterized by its membership function

$A:X \rightarrow [0,1]$

and A(x) is interpreted as the degree of membership of element *x* in fuzzy set A for each $x \in X$.The value zero is used to represent complete non-membership,the value one is used to represent complete membership,and values in between are used to represent intermediate degrees of membership.The

mapping A is also called the membership function of fuzzy set A.

Soft sets

There are many mathematical tools are developed for modeling the uncertainties such as Rough Set[8],Fuzzy set theory[10], probability theory, interval mathematics etc. All these approaches suffer in handling the uncertain data in the various areas like, economics, sociology, science, engineering, medical science, environmental science and other fields. Thus the soft set theory came into the existence to overcome the above said difficulties and it is widely applicable to resolve the problems.

The soft set theory [1-6], proposed by Russian researcher Molodtsov[3,4], in 1999 is a new generic mathematical tool for dealing with the uncertain data.

2. Preliminaries and Basic Definitions [3,4]

Definition 2. Let *U* be an initial universe set and *E* be a set of parameters. Let *P (U)* denotes the power set of *U* and *A* ⊂ *E*. A pair (*F*, *A*) is called a soft set over *U*, where *F* is a mapping given by *F: A→P (U).* In other words, a soft set over *U* is a parameterized family of subsets of the universe *U.* For ε ∈ A, F (ε) may be considered as the set of ε-approximate elements of the soft set (F, A).

Example 1. Let the initial universe U={**L_1,L_2,L_3,L_4,**}be the four selected district of Madhya Pradesh i.e. Bhopal, Indore, Gwalior & Jabalpur and E={**E_1,E_2,E_3,E_4**} be the generic parameters where **E_1,E_2,E_3,E_4** are Temperature, Climate, Cloud Cover ,Humidity and wind respectively.

F (E_1) ={L1,L3, L_4}

F (E2) = {L2,L3}

F (E3) = {L1, L_4} and

F (E4) ={L_2,L_3, L_4 }

Now (F,E) is a parameterized family { *F* (Ei): i=1,2,..4} which give a collection of approximate descriptions of an element.Each approximation has two parts: predicate part and approximate value set.Consider *F(E_1),here* predicate name is Temperature and the value set is{L1,L3, L_4} (Table 1)

Table 1: Tabular representation of Example1 (Soft Set).

U	E_1	E_2	E_3	E_4
L_1	1	0	1	0
L_2	0	1	0	1
L_3	1	1	0	1
L_4	1	0	1	1

Fuzzy soft sets

The real world is inherently uncertain, imprecise and vague. Traditional mathematical tools cannot deal with such problems. The fuzzy sets theory is widely employed in such kinds of problems . Maji and Biswas et al. [7,9] proposed the notion of fuzzy soft sets and an example of decision-making was discussed in [5] as well.

Definition 3. [7,9]: Let U be an initial universe set and E be a set of parameters. Let P (U) denotes the set of all fuzzy sets of U and A ⊂ E. A pair (F, A) is called a fuzzy soft set over U, where F is a mapping given by F: A→P (U).

Example 2 : Consider an example of Fuzzy Soft Set (F,E)={ {E_1},{E_2},{E_3},{E_4},{E_5}}

where

F(E_1) = { L_1/0.5, L_2/0.7, L_3/0.6, L_4/0.8}

F(E_2) = { L_1/0.9, L_2/0.4, L_3/0.8, L_4/0.3}

F(E_3) = { L_1/0.5, L_2/0.4, L_3/0.8, L_4/0.5}

F(E_4) = { L_1/0.9, L_2/0.5, L_3/0.7, L_4/0.6}

F(E_5) = { L_1/0.7, L_2/0.4, L_3/0.9, L_4/0.5}

Also, we can express it in a tabular form in Table 2.

Table 2 :Tabular representation of Example 2 (Fuzzy Soft Set).

U	E_1 Temperature	E_2 Climate	E_3 Cloud Cover	E_4 Humidity	E_5 wind

L_1	**0.5**	**0.9**	**0.5**	**0.9**	**0.7**
L_2	**0.7**	**0.4**	**0.4**	**0.5**	**0.4**
L_3	**0.6**	**0.8**	**0.8**	**0.7**	**0.9**
L_4	**0.8**	**0.3**	**0.5**	**0.6**	**0.5**

3. Fuzzy soft forecast model

The Fuzzy soft based rainfall forecasts model is used rainfall data year-2010 of four districts Bhopal,Gwalior,Indore,Jabalpur of Madhya Pradesh,India. The rainfall data are in the numerical form. The rainfall data are collected by MID over the time period in uniform interval. These datasets become a time series.

Definition 4. A time series $\hat{S}=\{s_1 s_2 s_3 \ldots s_n\}$ is an ordered set of real values. s_i is the i^{th} element of sequence $\hat{S}$ and the length of $\hat{S}$ is denoted as n.

Definition 5. Let (F,A) be a fuzzy soft over the universal set $U=\{L_1,L_2,L_3, L_4, L_5, L_6, \ldots\ldots L_n\}$,where U denotes different locations and $A=\{S_1,S_2, S_3 \ldots S_n\}$ be a set of individual forecasts. The mapping *F* is given by
F: A→P (U).

Definition 6. The fuzzy membership $f(x_{ij})$ is defined as follows:

$$f(x_{ij}) = [1-(|S^m_{i=1} - A^m_{i=1}|)/A^m_{i=1}]$$

where
n,m-- are the integer numbers

Definition 7.(Comparison Matrix) Comparison Matrix C[i,j] is represented as

$$C[i,j]= \sum_{k=1}^{m} (f_{ik} - f_{jk})$$

where
k -is a counter variable & varies from1 to m

Definition 8. The vector v_j is defined as follows:

$$\upsilon = \sum_{i=1}^{m} f(x_{ij})$$

Definition 9: The weight w_j is defined as follows:

$$w_j = \frac{\upsilon_j}{\sum_{i=1}^{n} \upsilon_i}$$

Proposed forecasting model (PFM):

Lot of research work has been done for rainfall forecasting using the classical approaches. This paper presents a new rainfall forecasting model based on fuzzy soft theory.

Let A be an actual time series over the time period in the uniform intervals. And A_t (t=1 to n) is its data points obtained through repeated measurements in uniform interval. Let $P=\{P_1,P_2, P_3 \ldots P_n\}$ be a set of individual forecasts. Table 3 shows the rainfall data(Time Series),year-2010 of four districts: Bhopal, Gwalior, Indore, Jabalpur of Madhya Pradesh, India. The L_1,L_2,L_3,L_4 denotes Bhopal, Gwalior, Indore and Jabalpur respectively.

The **Proposed forecasting model (PFM)** f_c is defined as follows

$$f_c = \sum_{j=1}^{n} w_j * p_{jt}$$

where
i --is a counter variable & varies from 1 to m
j --is a counter variable & varies from 1 to n
n- is an integer number

The fuzzy soft set[9-10] is used to develop an effective algorithm for reliable output. The inputs for the proposed model are selected locations i.e Bhopal,Gwalior,Indore,Jabalpur and their rainfall datasets. Forecasting is the output of the proposed model shown in table 6.

Algorithm1

Inputs:

A. The Actual time series A_t

B. The forecast it with each individual forecast S_{ti} {t=1,2,3,…….m and i=1,2,3…….n}

Output: forecast time series f_c

1: Construct Fuzzy Soft set(F,A)
Loop : i=1 to m
Loop:j=1 to n

$$f(x_{ij}) = [1-(|S_{i=1}^{m} - A_{i=1}^{m}|) / A_{i=1}^{m}]$$

2: Calculate Comparison Matrix C[i,j]
Loop : i=1 to m
Loop:j=1 to n
Loop:k=1 to n

$$C[i,j] = \sum_{k=1}^{m} (f_{ik} - f_{jk})$$

3: Calculation of v_j Vector
Loop : j=1 to n

$$\upsilon = \sum_{i=1}^{m} f(x_{ij})$$

4: Find the weight w_j
Loop : j=1 to n

$$w_j = \frac{\upsilon_j}{\sum_{i=1}^{n} \upsilon_i}$$

5: Calculate f_c
Loop : j=1 to n

$$f_c = \sum_{j=1}^{n} w_j * p_{jt}$$

End:

The output of the algorithm1 is Table 4 Fuzzy representation, Table 5 Comparison Matrix representation and Table 6 Forecasting results. Vector values V={V_1,V_2,V_3} where V_1=1.117772 ,V_2=0.95892 and V_3 =1.230998. Weight of vector W={W_1,W_2,W_3} where W_1=0.337431,W_2=0.289406 and W_3 =0.372162.

Table 3:Rainfall data (Time Series)

U	**A**	**S_1**	**S_2**	**S_3**
L_1	**740**	**1203.2**	**786.1**	**1174.7**
L_2	**670.6**	**965.1**	**958.8**	**1015.8**
L_3	**768.8**	**516.6**	**921.3**	**780.6**
L_4	**1201.1**	**1122**	**1398.4**	**1372.9**

Table 4:Tabular representation of Example 2 (Fuzzy Soft Set).

S_1	S_2	S_3
0.374324	0.937783	0.405811
0.560841	0.576236	0.485237
0.657168	0.828012	0.993003
0.934144	0.835734	0.856964

5. Experiment and results

Experimental Setup and Data sets: All the experiments are done on a PC with MS windows XP operating system, Matlab tools and 512 MB RAM, Pentium IV processor with 2.53 Ghz. The rainfall data are collected by IMD [11] over the time period in uniform interval .The rainfall data are in the numerical form. These datasets become a time series. The data of four districts Bhopal, Gwalior,Indore,Jabalpur of Madhya Pradesh,India is collected for forecasting. Table 3 shows the Rainfall data(Time Series) of four districts of year 2010. The L_1,L_2,L_3,L_4 denotes Bhopal,Gwalior,Indore and Jabalpur respectively.

Table 5:Comparison Matrix C[i,j] representation

S_1	**S_2**	**S_3**
0	0.079426	0.587193
0.079426	0	0.507766
0.587193	0.507766	0
0.457154	0.371727	0.136034

Table 6:Forecasting results

U	M_1	M_2	M_3	**PFM**
L_1	0.224157	0.23557	0.214017	0.241557
L_2	0.201812	0.221581	0.215012	0.215812
L_3	0.350563	0.335635	0.325765	0.345635
L_4	0.321054	0.300854	0.310854	0.310854

Table 6 is a fuzzy representation of the rainfall data.Algorithm1 is used to for forecasting using the fuzzy soft concepts. Table 5 represents the matrix representation of fuzzy data table 4.The results of the proposed fuzzy soft are shown in Figure-1, Figure-2 and Figure-3.

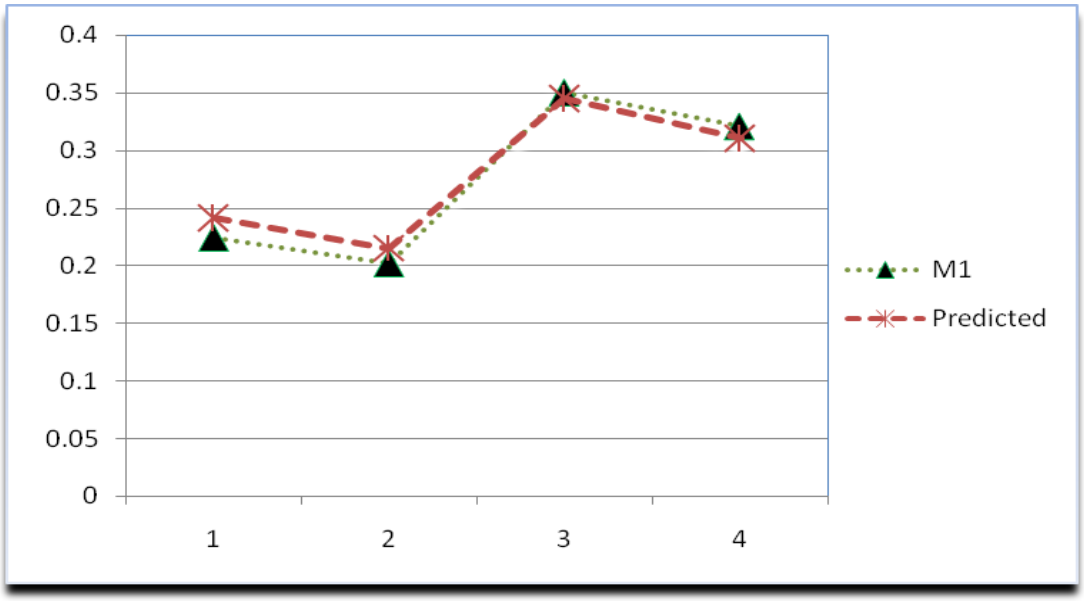

Figure-1:Comparision of M_1 and PFM

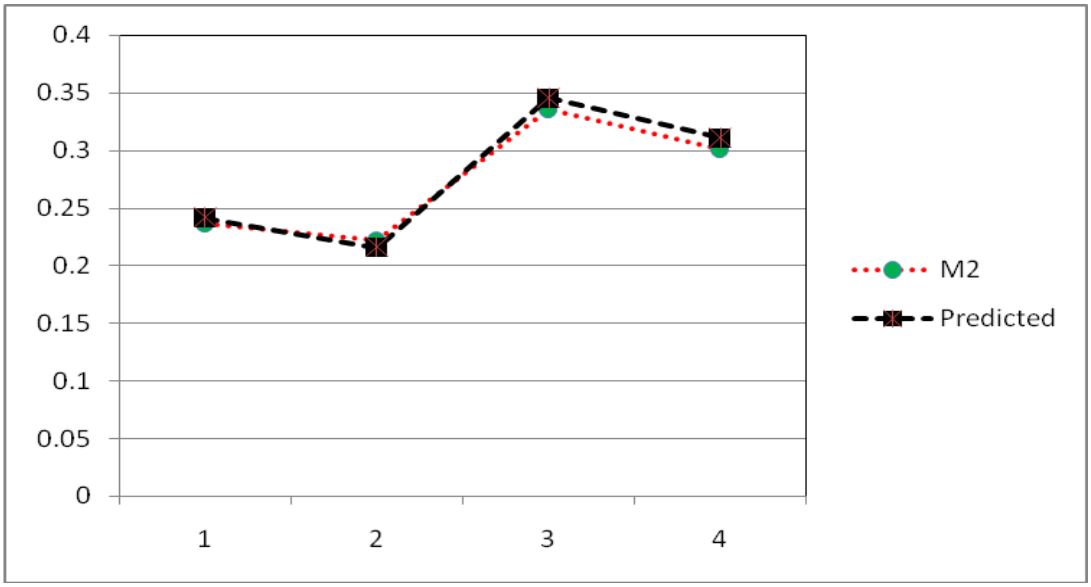

Figure-2: Comparision of M_2 and PFM

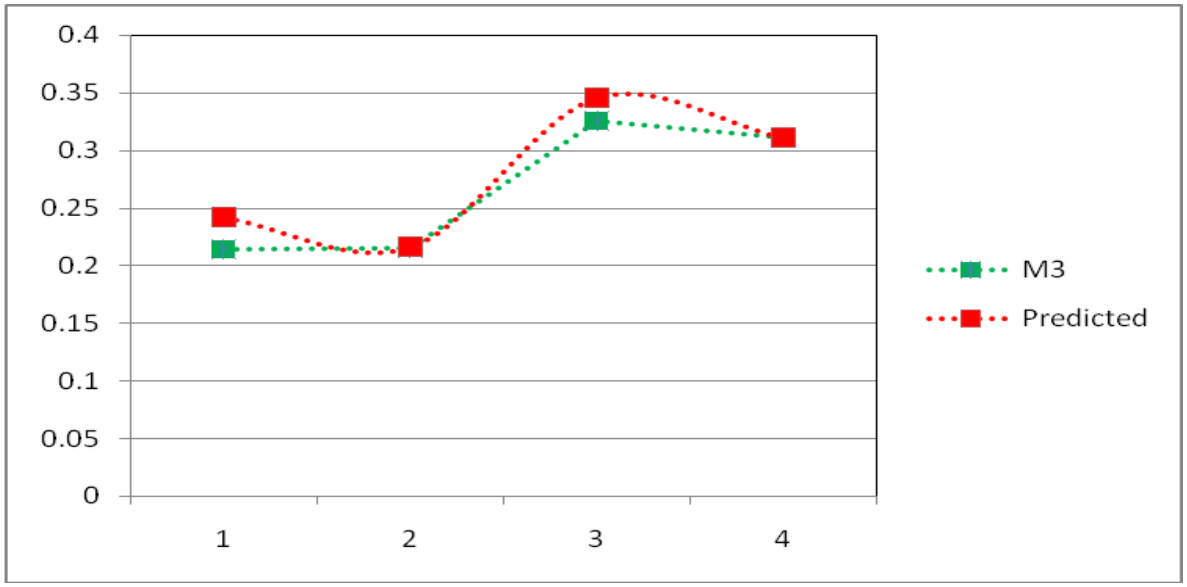

Figure-3: Comparision of M_3 and PFM

6. Conclusion

In this research paper the fuzzy soft based rainfall pattern forecasting model is proposed. The accuracy and efficiency of the proposed model outperforms. Bhopal, Gwalior, Indore, Jabalpur four districts of Madhya Pradesh, India are taken for rainfall forecasting. This model is also useful for all states rain fall pattern forecasting.

References

[1]. H. Aktas, N. Cagman, Soft sets and soft groups, Inform. Sci. 177 (2007) 2726–2735.
[2]. D. Chen, E.C.C. Tsang, D.S. Yeung, X. Wang, The parameterization reduction of soft sets and its applications, Comput. Math. Appl. 49(2005) 757–763.
[3]. D. Molodtsov, The Theory of Soft Sets, URSS Publishers, Moscow, 2004 (in Russian).
[4]. D. Molodtsov, Soft set theory—first results, Comput. Math. Appl. 37 (1999) 19–31.
[5]. P.K. Maji, A.R. Roy, R. Biswas, An application of soft sets in a decision making problem, Comput. Math. Appl. 44 (2002) 1077–1083.
[6]. P.K. Maji, R. Bismas, A.R. Roy, Soft set theory,Comput. Math.Appl.45 (2003) 555–562.
[7]. P.K. Maji, R. Biswas, A.R. Roy, Fuzzy soft sets, J. Fuzzy Math. 9 (3) (2001) 589–602.
[8]. Z. Pawlak, Rough sets, Int. J. Inform. Comput. Sci. 11 (1982) 341–356.
[9]. A.R. Roy, P.K. Maji, A fuzzy soft set theoretic approach to decision making problems, J. Comput. Appl. Math. 203 (2007) 412–418.
[10]. L.A. Zadeh, Fuzzy Sets, Inform. Control 8 (1965) 338–353.
[11]. http://www.imd.gov.in
[12]. Harshita Patel, G.S. Thakur, “An Improved Fuzzy K-NN Algorithm for Imbalanced Data using Adaptive Approach”, published in IETE Journal of Research in Vol. 65, Issue 6, pp. 780-789, ISSN: 0974-780X
[13]. Brijesh Bakariya and G.S. Thakur, "An Efficient Algorithm for Extracting Infrequent Itemsets from Weblog", The International Arab Journal of Information Technology (IAJIT),2019 (2)
[14]. G.S. Thakur, "Fuzzy Soft Traffic Accident Alert Model", National Academy Science Letters June 2014, Volume 37, Issue 3, pp 261-268(Springer)
[15]. Brijesh Bakariya and G.S. Thakur, "Mining Rare Itemsets from Weblog Data",National Academy Science Letters, 2015 page 1-5
[16]. Brijesh Bakariya and G.S. Thakur, "Pattern Mining Approach for Social Net-

work Service", National Academy Science Letters, June 2017, Volume 40, Issue 3, pp 183–187

[17]. Krishna K. Mohbey & G. S. Thakur,Interesting User "Behaviour Prediction in Mobile E-commerce Environment using Constraints",published by Taylor & Francis in IETE Technical Review,18 Nov 2014

[18]. Brijesh Bakariya, G.S. Thakur,"An Efficient Algorithm for Extracting High Utility Itemsets from Web Log Data",published in The Institution of Electronics and Telecommunication Engineers (IETE) Technical Review SCI DOI: 10.1080/025646 02.2014.1000396, Volume-32, Issue-02, March- 2015, Pages 151-160. ISSN 0256-4602 impact factor1.304

[19]. G.S.Thakur,Ravi Singh, "New Hesitant Fuzzy Operators", in Fuzzy Information and Engineering Volume 6, Issue 3, September 2014, Pages 379–392

[20]. Yachana Bhawsar, G.S. Thakur,"Performance Evaluation of Link Prediction Techniques Based on Fuzzy Soft Set and Markov Model", Fuzzy Information and Engineering,Volume 8, Issue 1, March 2016, Pages 113–126

[21]. Krishna K. Mohbey and G.S. Thakur, "Constraint based Interesting Location and Mobile Web Service Sequence Mining in M-commerce Environment", published in Journal of Theoretical and Applied Electronic Commerce Research, ISSN 0718-1876, 11(1):84-95, 2016

[22]. Ghanshyam Singh Thakur and Rekha Singh Thakur," BSclassifier for Balance Scale Weight and Distance Database", International Journal of Soft Computing Year: 2010 | Volume: 5 | Issue: 6 | Page No.: 211-213.

[23]. Thakur G.S. and Jain R.C. "Association Rule Generation from textual data", In International Journal of Soft Computing 2(2), 346-348, 2007.

[24]. D. S. Rajput, R. S. Thakur, G. S. Thakur, "A Computational Model for Knowledge Extraction in Uncertain Textual Data using Karnaugh Map Technique", International Journal of Computing Science and Mathematics (InderScience) ISSN: 1752-5055, Jan 2016, Vol. 7, Issue 2, pp. 166-176

[25]. Harshita Patel, G.S. Thakur, "Classification of Imbalanced Data using a Modified Fuzzy-Neighbor Weighted Approach", published in IJIES in Vol. 10, Issue 1, 2017, pp. 56-64

[26]. Harshita Patel, G.S. Thakur, "Improved Fuzzy-Optimally Weighted Nearest Neighbor Strategy to Classify Imbalanced Data", published in IJIES in Vol. 10, Issue 2, 2017, pp. 156-162

[27]. D.S Rajput, R.S. Thakur, G.S. Thakur, "An Integrated Approach and framework for Document Clustering using Graph based Association Rule Mining" published in Advances in Intelligent Systems and Computing, Springer, ISBN 978-81-322-1602-5, Vol. 236 pp. 1421- 1438.

[28]. G. Vishwakarma and G. S. Thakur, "Hybrid System for MPAA Ratings of Movie Clips Using Support Vector Machine," in Advances in Intelligent Systems and Computing book series (AISC, volume 817), 2019, pp. 563–575

[29]. G.S. Thakur,R.S. Thakur, "Design of 2-Level Clustering Framework for Time Series DataSets", In an International Conference on Soft Computing for Problem Solving (SocPros11) ,IIT Roorkee India, December 16-18, 2011, Springer Berlin/Heidelberg

[30]. D.S Rajput, R.S. Thakur and G.S. Thakur, "Fuzzy Association Rule Mining based Frequent Pattern Extraction from Uncertain Data" presented in IEEE 2nd World Congress on Information and Communication Technologies (WICT-2012) October 30-November 02, 2012 in IIITM Trivandrum, ISBN: 978-1-4673-4804-1 pp 709-714

[31]. Krishna Kumar Mohbey, G.S.Thakur, "New Architecture for Fuzzy Association Rule Mining in Ecommerce Environment" presented in IEEE International conference on control, computing, communication and materials (ICCCCM 2013), Allahabad, August 3-4, 2013.

[32]. G.S. Thakur,Neeraj Sahu,Swatranta Sahu,"Hesitant Fuzzy Linguistic Term Set Based Document Classification" ,In IEEE Third International Conference on Communication Systems and Network Technologies April 6-8,2013,Gwalior,India

[33]. Neeraj Sahu ,G.S. Thakur and R.S.Thakur , "Numerical Result Analysis of Document Classification for Hesitant k-Nearest Neighbor Classifier" , The Second International Conference on "Soft Computing for Problem Solving (SocProS 2012)" December 28 - 30, 2012. Published by the AISC Series of Springer Berlin/Heidelberg.

[34]. G.S. Thakur,R.S. Thakur, "Design of 2-Level Clustering Framework for Time Series DataSets", In an International Conference on Soft Computing for Problem Solving (SocPros11) ,IIT Roorkee India, December 16-18, 2011, Published Springer Proceeding.