INTERNATIONAL JOURNAL OF RESEARCH CULTURE SOCIETY ISSN: 2456-6683 Volume - 5, Issue - 12, Dec - 2021
Monthly Peer-Reviewed, Refereed, Indexed Journal Scientific Journal Impact Factor: 5.743
Received on : 26/11/2021 Accepted on : 13/12/2021 Publication Date: 31/12/2021

Incomplete information system to complete information system

[1]B. Venkataramana, [2]L. Padmasree, [3]M. Srinivasa Rao [4]G. Ganesan
[1]Department of Computer Science & Engineering, Holy Mary Institute of Technology and Science, Bogaram, Kesara Mandal, Telengana, India,.
[2]Department of Electronics & Communications Engineering, VNR Vignana Jyothi Institute of Engineering & Technology, Bachupalli ,Nizampet, Telengana, India,
[3]School of Information Technology, Jawaharlal Nehru Technological University Hyderabad, Kukatpally,
[4]Department of Mathematics, Adikavi Nannaya University, Rajahmundry, Andhra Pradesh, India,
Email: [1]bandaruramana1@gmail.com, [2]padmasree_l@vnrvjiet.in, [3]srmeda@gmail.com, [4] prof.ganesan@yahoo.com

Abstract: *The main objective of this paper to complete information system from incomplete information system. Development of algorithms for Indexing large data either of complete or incomplete, using Fuzzy concepts and evaluate their applicability to regular data and for image data.An information system is a relational database in which data are stored as tables. In real time decision tables, sometimes the attributes values may be missing due to two reasons: One is due to 'Loss of Data' and another is due to 'Don't Care attributes'. The information system which consists of such missing data is often termed as the Incomplete Information systems. There are good computing tools in finding minimal features of the incomplete information systems are currently in use. However, since this work is focusing on the approaches of rough sets in incomplete information systems by assuming all 'Loss Data' and 'Don't Care attributes' as the separate clusters, this thesis did not deal the various approaches in computing the reducts.*

Key Words: *Fuzzy, Don't Care attributes, clusters.*

1. INTRODUCTION:

Several solutions to the problem of knowledge discovering from an IIS have been proposed in the area of artificial intelligence[1-5]. rough set theory (also called rough set) is a mathematical tool which was proposed by professor pawlak in 1982 and is able to make quantitative analysis and processes for inaccurate, inconsistent, and incomplete information and knowledge. the original prototype of rough set theory resulted from a simple information model, and its basic idea was defined by a relational database used to form concepts and rules and realize knowledge discovery by the classification of the equivalence relation and of objective approximation. the core foundation of rough set theory and its application is a couple of approximate operators derived from approximation space, namely the upper approximation operator and lower approximation operator (also called the upper and lower approximation set). the obscurity relationship of the classic pawlak model is an equivalence relation with high demands, and it restricts the application of the rough set model. therefore, how to expand and define the approximate operator has become an important area of rough set theory research.

The probability distribution of unknown value over the possible values from the domain. it is suggested in[5] to predict an attribute value based on the values of other attributes of the object. a rough set [6] approach to knowledge acquisition is based on the assumption that two objects are indiscernible with regard to attributes set if they have the same value for each attribute from the attributes set. in the case of IIS, such a definition is not sufficient. An information system is a relational database in which data are stored as tables. it is defined by a set of attributes some of which are conditional attributes and some decision attributes. in incomplete information system attribute values are missing either because they were lost or they are "do not care" conditions. a "do not care" condition is, original values not being recorded due to irrelevant and the decision to which concept a case belongs to was taken without this information.

2. EXISTING SYSTEM:

Incomplete Information System

An integrated set of components for collecting, storing, and processing data and for providing information, knowledge, and digital products. An Incomplete Information Systems is a system where we may get missing / noisy values.

Row	A1	A2	A3	Member ship
1	40	60	30	0.4
2	25	60	35	0.5

3	65	25	35	0.6
4	60	-1	75	0.4
5	48	65	35	0.5
6	-1	60	30	0.7
7	25	45	35	0.4
8	65	35	-1	0.1
9	60	45	68	0.9
10	48	65	35	0.3

3. PROPOSED SYSTEM:

The main objective of this project is to complete information system from incomplete information system. Development of algorithms for Indexing large data either of complete or incomplete, using Fuzzy concepts and evaluate their applicability to regular data and for image data.

It helps you to reduce the total number of I/O operations needed to retrieve that data, so you don't need to access a row in the database from an index structure. Offers faster search and retrieval of data to users. Indexing also helps you to reduce table space as you don't need to link to a row in a table, as there is no need to store the ROWID in the Index and so it reduce the table space. You can't sort data in the lead nodes as the value of the primary key classifies it.

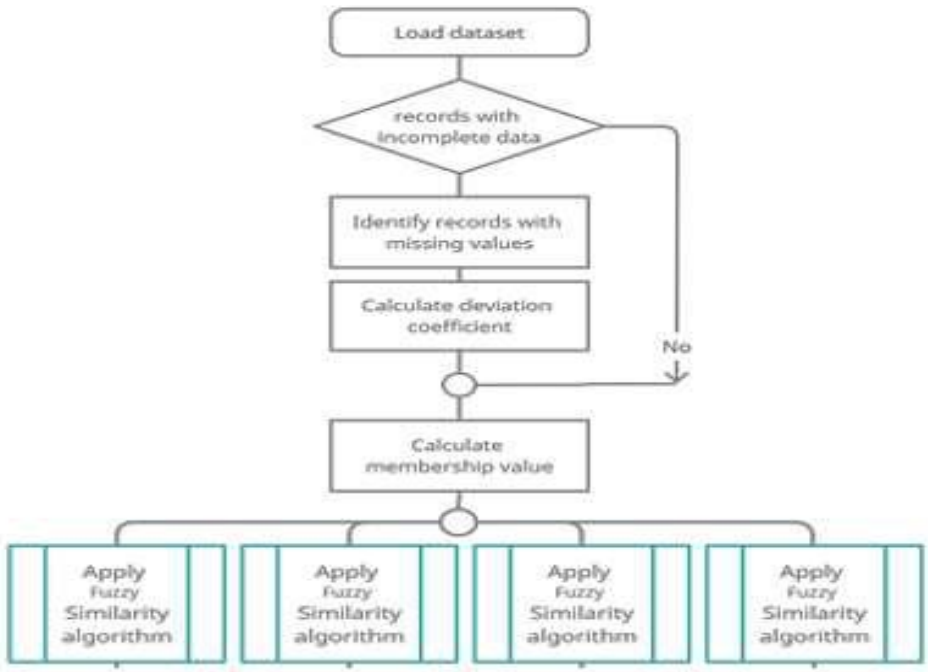
4. METHODOLOGY:

Indexes in Large Databases

Indexing of records in any database or information system enhances the efficiency of searching and hence, several researchers focus on developing indexing procedures in these systems an indexed file is a file with an index that allows easy random access to any record given its file key. The key must be such that it uniquely identifies a record. If more than one index is present the other ones are called alternate indexes. The indexes are created with the file and maintained by the system. An index on a file speeds up selections on the search key fields for the index. Any subset of the fields of a relation can be the search key for an index on the relation. Search key is not the same as a key in the Database. Convert Incomplete Information system to Complete information system using…

1. Fuzzy Similarity algorithm
2. Fuzzy Dissimilarity algorithm
3. Intuitionistic Fuzzy Similarity algorithm
4. Intuitionistic Fuzzy Dissimilarity algorithm

Architecture Diagram

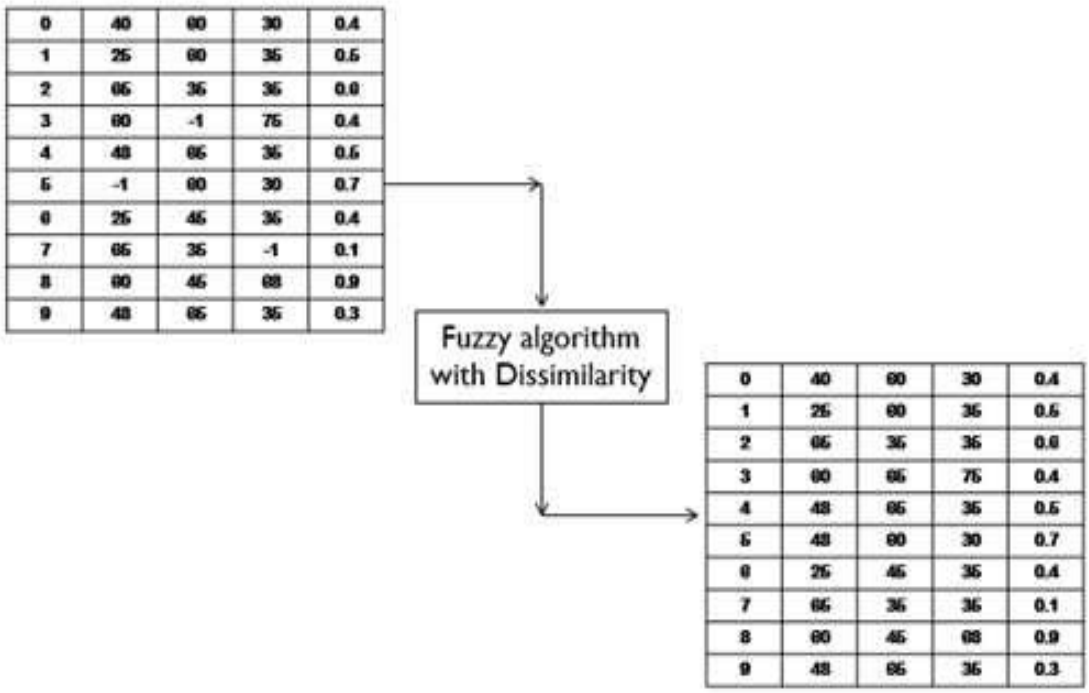


INTERNATIONAL JOURNAL OF RESEARCH CULTURE SOCIETY ISSN: 2456-6683 Volume - 5, Issue - 12, Dec - 2021
Monthly Peer-Reviewed, Refereed, Indexed Journal Scientific Journal Impact Factor: 5.743
Received on : 26/11/2021 Accepted on : 13/12/2021 Publication Date: 31/12/2021

IMPLEMENTATION

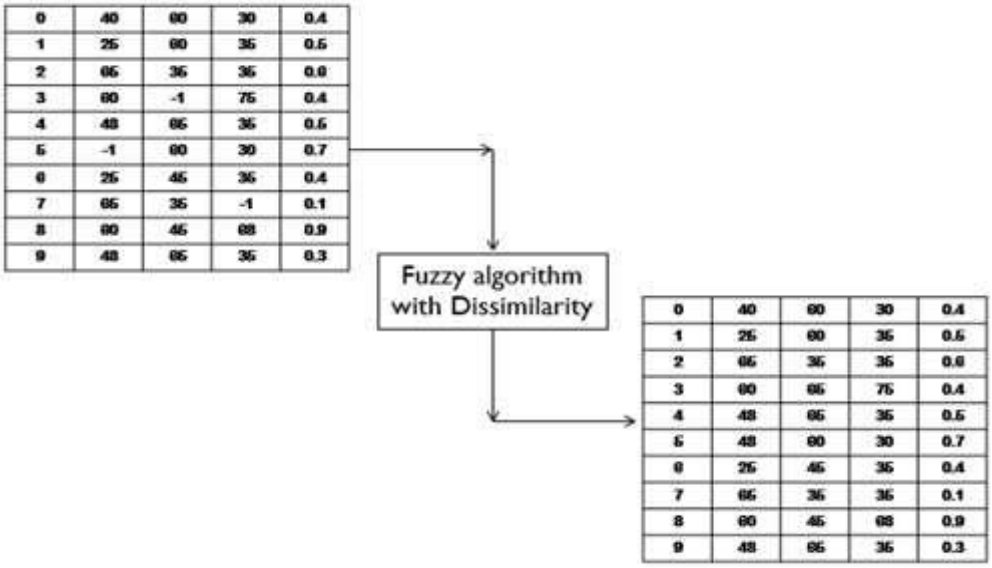
Fuzzy Similarity algorithm :

1. Read input alpha
2. Calculate cut value as 0 or 1 for all records
3. If (membership ≥ alpha) then cut = 1
   else cut = 0
4. Identify rows with missing values
5. Repeat for identified rows
6. Select rows with similar values comparing to the current row
7. Calculate deviation values
8. Find row with minimum deviation value
9. Replace missing value with the corresponding value in the above row.



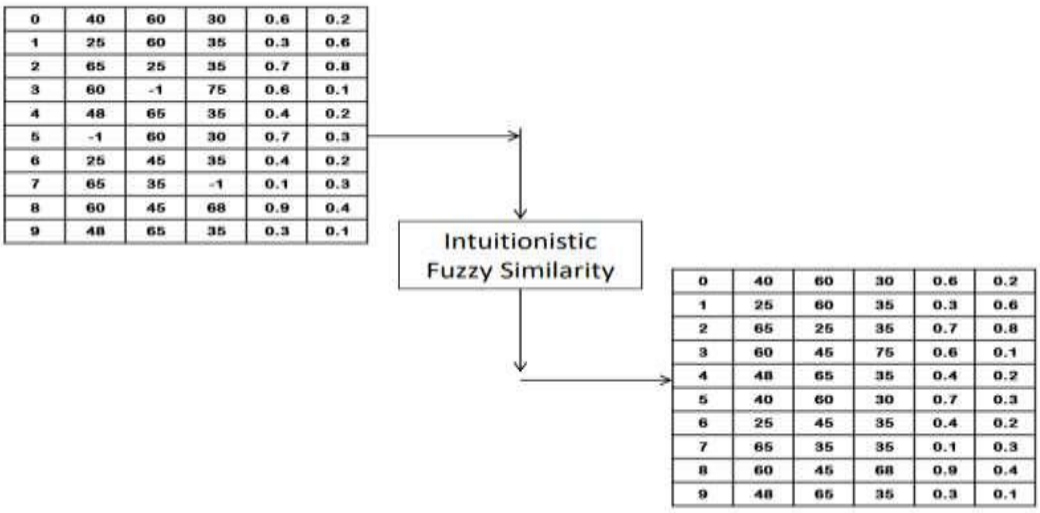
Fuzzy Dissimilarity algorithm

1. Read input alpha
2. Calculate cut value as 0 or 1 for all records
3. If (membership ≥ alpha) then cut = 1
   else cut = 0
4. Identify rows with missing values
5. Repeat for identified rows
6. Select rows with dissimilar values comparing to the current row
7. Calculate deviation values
8. Find row with maximum deviation value
9. Replace missing value with the corresponding value in the above row



Intuitionistic Fuzzy Similarity algorithm

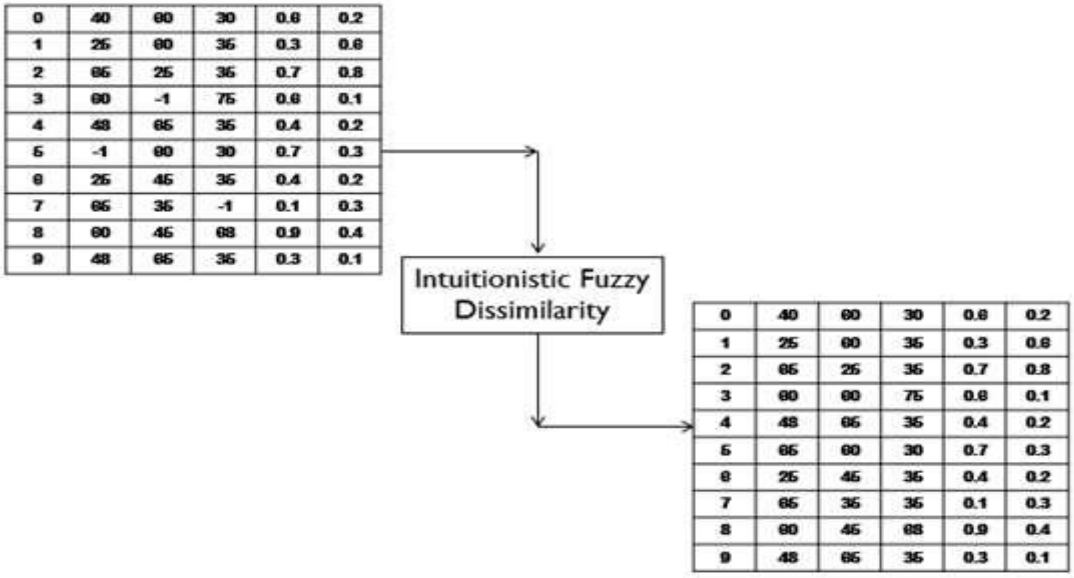
1. Read input alpha, beta
2. Calculate cut value as 0 or 1 for all records
3. If (des1 ≥ alpha and des2 ≤ beta) then cut = 1
   else cut = 0
4. Identify rows with missing values
5. Repeat for identified rows
6. Select rows with similar values comparing to the current row
7. Calculate deviation values
8. Find row with minimum deviation value
9. Replace missing value with the corresponding value in the above row



Intuitionistic Fuzzy DisSimilarity algorithm

1. Read input alpha, beta
2. Calculate cut value as 0 or 1 for all records
3. If (membership ≥ alpha and non-membership ≤ beta) then cut = 1
   else cut = 0
4. Identify rows with missing values
5. Repeat for identified rows

6. Select rows with dissimilar values comparing to the current row
7. Calculate deviation values
8. Find row with maximum deviation value
9. Replace missing value with the corresponding value in the above row



5. RESULT:

Row index having Missing value	Fuzzy Similarity	Fuzzy Dissimilarity
3	45	65
5	40	48
7	35	35

Table 1 : Fuzzy with similarity and dissimilarity

The results in the above table shows that both Fuzzy with similarity and dissimilarity have generated unique values for the random samples. But with dissimilarity the range is more compared to similarity.

Row Index having Missing value	Intuitionistic Fuzzy Similarity	Intuitionistic Fuzzy Dissimilarity
3	45	60
5	40	65
7	35	35

Table 1.1: Intuitionistic Fuzzy with similarity and dissimilarity

The results clearly indicates that in case of Intuitionistic Fuzzy with similarity, the algorithm has generated unique values for the random samples. But with dissimilarity it has generated index with higher range comparing to similarity.

Row Index having Missing value	Fuzzy Similarity	Fuzzy Dissimilarity	Intuitionistic Fuzzy Similarity	Intuitionistic Fuzzy Dissimilarity
3	45	65	45	60
5	40	48	40	65
7	35	35	35	35

Table 1.2 : Consolidated Results

The results shows that Fuzzy Similarity and Intuitionistic Fuzzy similarity algorithms generated unique values for the random samples also conclude that similarity algorithms generate index with lower range comparing to Dissimilarity algorithms.

6. CONCLUSION:

Our main objective is to complete information system from incomplete information system. Development of algorithms for Indexing large data either of complete or incomplete, using Fuzzy concepts and evaluate their applicability to regular data and also for image data. By Fuzzy Similarity and Intuitionistic Fuzzy similarity algorithms generate unique values for the random samples. The results conclude that similarity algorithms generate index with lower range comparing to Dissimilarity algorithms.

REFERENCES:

1. Ljubljana ,I. Kononenko, I. Bratko and E. Roskar (1984): Jozef Stefan Institute : Experiments in Automatic Learning of Medical Diagnostic, Yugoslavia.
2. Morgin Kaufmann and J.K. Quinlan (1990): Introduction of decision trees, In Reading in incline learning, (Edited by J.W.Shavlik and T.G.Dietterich), pp(57-69).
3. Springer and J. Stefanowski (1994): Stefan Institute Handling various types of uncertainty in rough set approach ,in rough sets, Fuzzy Sets and Knowledge Discovery (RSKD'93),(edited by w.ziarko),Berlin.
4. M.R. Chmielewski, J.W. Grzymala-Busse, N.W. Peterson and S. Than(1994): The rule induction system LERS-Aversion for personal computers, Foundations of Computing and Decision Sciences 18 (3-4),181-212.
5. R. Slowinski and J. Stefanowski (1993): Rough Classification in Incomplete Information System, Math,Comput Modelling 12(10-12),1347-1357.
6. Z. Pawlak (1982): Rough sets, International Journal of Computing and Information Science 11,(341-356).