International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-5, June 2019

Imprecise Data Assessment for Information retrieval using Fuzzy Associative Memory Mechanisms

P. Mohan Kumar

Abstract: . Query processing is the primary task in information retrieval.Either user information is a keyword or a sentence based the answer display to the user is only what the resultant obtained from query execution. If the correct information to be executed was given the processing time minimized to the maximum extent by avoiding unnecessary formalisms, tuning, identifying the nearest relevance, etc. instead directly the query can be allowed to process. In this paper, an approach was proposed to assess the incoming user data imprecise extent and its reutilization level for future use instead of avoiding as wrong or irrelevant.Fuzzy associative memory mechanism was deployed to assist and associate user compatibility while retrieving information

Index Terms: Imprecise, fuzzy, semantic query, data assessment, reuse

I. INTRODUCTION

Human create fast call supported imprecise information as assumptions that keep information represent the right set of world information and work with the information or web in existence for survivals. Typically the data is also irrelevancy except for alternative input, it should be precise, thus it should not take away instead managed to utilize fitly for minimizing the interval. Here data analysis was performed on user given irrelevance input by measuring the precision-imprecision level and assess how far relevant to the existing database information and essential for query processing. In this paper solution for this is provided. The roadmap of this paper work flow is initially the imprecise data form was defined and followed managing impreciseness for query execution /information retrieval and semantic mapping in order to acquire knowledge during execution. Then the Fuzzy model to represent the imprecise data and processing methods to information using the proposed FAM approach.

II. DATA ANALYSIS

Data analysis is required in reference to question process, to provide knowledge outline data with the kind of rules or assertions that permit linguistics question optimization or direct question responsive while not consulting the information itself. The aim of analysis thought of during this paper is to enhance the speed of responsive queries thereon data.The outline data creat ed by instrument will either answer a question while not consulting the information itself, instead modify the question to a type the information server are able to method additional quickly. Here this outline is employed to spot precise or general and to assess whether or not it's relevant for the process question or not. additional knowledge analysis on observe that forms of errors and inconsistencies square measure to be removed, betting on the quantity of knowledge sources, their degree of heterogeneousness and also the "dirtiness" of the information, transformation progress and mapping rules should be outlined. Some errors solely become apparent once applying some transformations that the correctness and effectiveness of a metamorphosis progress and also the transformation definitions ought to be tested and evaluated in verification part. Once subtlety is over the cleansed knowledge ought to replace the dirty knowledge within the original sources so as to avoid redoing the cleanup work for future knowledge extractions. Knowledge delicacy meth od [1,2]]is as shown in below diagram; here we tend to concentrating in abundant depth regarding solely the essential problems as associated with our half i.e. at instance level as previous side is taken into account.

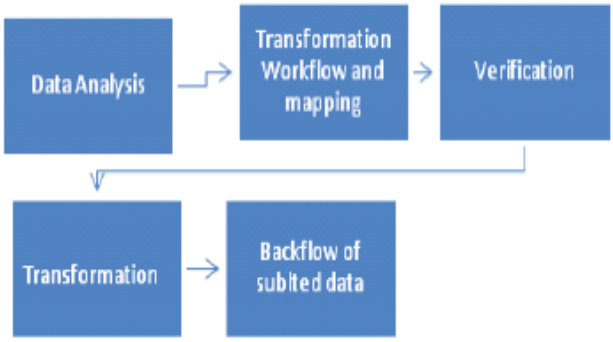

Fig 1 Data Delicacy Process

III. PROPOSED WORK

A. *Information retrieval using FAM mechanism*

The generalized FAM architecture was discussed in order to bring out the work flow of the proposed work then followed the fuzzy approach to represent data and query relevant to imprecision[3] and operation performed over defined datasets.

Manuscript published on 30 June 2019.
* Correspondence Author (s)
Mohan Kumar P, Department of Digital Communications, VIT University, Vellore (Tamil Nadu), India.

B. *Fuzzy system Architecture*

Fuzzy Associative Memory (7) will be viewed because the combination of associative memory and symbolic logic. FAM roughly refers to the depot of fuzzy association s every of that encodes a fuzzy rule. Normally fuzzy systems S maps families of fuzzy sets to families of fuzzy sets .i.e

$$S=I^{n1}xI^{n2}\ldots\ldots I^{nr}\rightarrow I^{p1}xI^{p2}x\ldots.I^{Pn}$$

(1)

In account of this mapping between fuzzy sets to fuzzy sets that fuzzy systems behave like associative recollections. Hence, fuzzy systems area unit said as fuzzy associative recollections. A FAM not like standard neural network that acquire data through coaching directly encodes structured data of the shape. If X could be a then Y is B wherever A and B area unit n-dimensional and P-dimensional fuzzy sets severally and X, Y area unit fuzzy variables. It's been extended as a Fuzzy illation System and used wide and with success within the areas of pattern recognition and management issues. during this paper we tend to illation a model as extension of FAM that uses fuzzy pure mathematics to map inputs (features within the case of fuzzy classification) to outputs (classes within the case of fuzzy classification). Fuzzy illation model [8] are common computingframeworksthat supported the ideas offuzzy puremathematics, that are applied successf ully in several fieldslike management, call support, system identification Here membership functions obtained from training data, there are two important factors for increasing the accurately inferred results. First is collecting a correct training data set and second is the method to form membership functions (MFs). It means a group of defect coaching information willcertainly influe nce inferred results of a fuzzy abstract thought. The input layer simply accepts input values. Every node of input layer is employed for one feature or input variable. The fuzzification layer contains membership functions of input variables. Every input variable xi produces pi fuzzy sets. Thus, the add of nodes during this layer becomes Σin nodes. The output of this layer is that the match values of input values to associated membership functions. The antecedent layer contains antecedent components of fuzzy rules that have the shape of logical AND of individual fuzzy terms. This layer needs N = Π p nodes. We have a tendency to permit each attainable combination of fuzzy sets. Every incoming link will haea weight that represents the degree of the quality of associate fuzzy set. Every node of this layer simply compares incoming values and takes minimum of them. The output of this layer associating with weights of links between antecedent layer and ensuing layer then becomes the values of input values to the resultant layer contains consequent components of fuzzy rules.

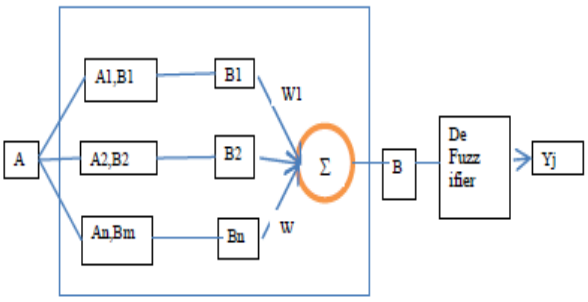

Fig 2 Generalized FAM System

1. The analysis on imprecision -based queries has sometimes been conducted either for fuzzy databases, for classical ones and for info retrieval systems additionally. For the fuzzy relation-based approach, [9] has represented a fuzzy relative calculus, (10) have developed a fuzzy relative pure mathematics extending the classical relative pure mathematics with the specification of associate degree accretive threshold on the attribute fuzzy degrees. In [11], the authors have introduced a whole fuzzy relative calculus supported a fuzzy extension of the classical domain calculus. For the chance based mostly approach, [12.13], have developedafuzzy relative calculus supported thetwin idea s of chance and necessity.an extension of SQL with fuzzy querying capabilities, known as SQLf, is planned. The modeling of fuzzy quantifiers like'about 10' or 'much but 50'has been pursued in [14].

2. Data and question Model

In this section however the info is diagrammatic and the way it will be processed in by a question a model was mentioned.

3. Precise questionuser question that needs information precisely matching the question limit could be a precise question. Certain questions contain solely crisp conditions over the attributes

4. General question

Auser question that doesn't impose precise match(and sol ely needs information closely matching the question constraint)is associatedegree general question. So associ ate degree general question contains similarity expressions that rank the results. Answers to such {a question |a question |a question} should be hierarchal in line with their closeness/similarity to the query constraints. as an example, the question Q:- Stud DB(subject like DBMS) is associate degree general question, the answers to that should have the attribute subject certain by a price almost like database management system. The diagram shows the projected design for the work given. It evokes the bottom diagram with extension the content unit and therefore the deciding unit. The system generally willbe viewedas symboliclogic controll er.Herethe management was determined as 3 basic purpo seful units one fuzzification [15] wont to calculate the fuzzy input.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

(i.e. to judge the input variables with relevancy corresponding terms within the condition side) second the fuzzy illation in our case the conjointly data ,database and rule based mostly to calculate the activation strength of each rule base and mix their action aspect and third defuzzification to calculate the particular output i.e. to convert the fuzzy output into precise price. The work flow of the system was supported the algorithmic rule declared below. The expected member perform is shown within the below table. The experimental take a look at is performed to realize this member perform connection.

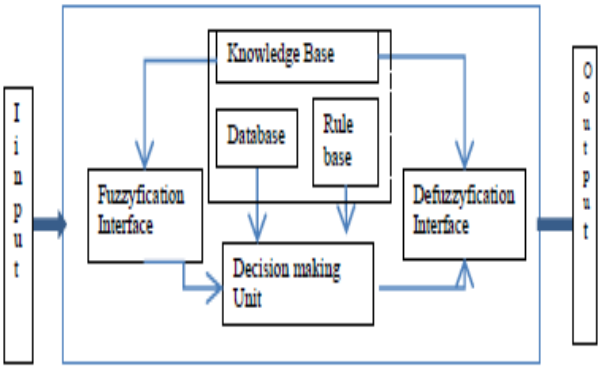

Fig3 Proposed FAM MODEL

C. *Generalized Algorithm*

Input training datasets with associated rules and weights
Output Expected member function (crisp)
Compute memberships of current input in the relevant antecedent fuzzyset of rule

i) Compute minimum of minimum memberships found in step 1.
   a. (since the antecedents are in conjunctive form and AND operation is replaced by minimum).
ii) Scale or clip the consequent fuzzy set of the rule by minimum value found in step(Since this gives the small degree to which the rule must fire)
iii) Repeat step 1-3 for each rule in the rule base.

Super pose the scale or clipped consequent fuzzy sets and compute the centroid pf composite set formed by such a super position

IV. EXPERIMENTAL SETUP

The basic work of this paper is to assess the user imprecise input utilization for answering future queries i.e. for retrieving other new coming relevance information instead of neglecting it as wrong input. The FAM application sample problem balancing an inverted pendulum as reference for work proof convenience. Inspiring the application of FAM we took the sample to test our proposed concept as follows. We took the academic system so called FFCS where the students and faculty are allowed to choose their work based on convenience. Particularly students given priority to choose their subject, faculty, class timings and exam mode based on their convenience. Here progress may be university core or department wise for example university means the students can register for their subject irrespective of their department. The experimental setup is as similar shown in generalized architecture. Here each element is a processing unit and storage i.e. (a cpu+ memory/cache) connected as shown and allowed to process in parallel. Fuzzy interface unit consists the user given input which may have the relevance or similar or exact to the existing database to processed and retrieved. With respect to the sample taken some sort of processed results and the query predicates with semantics form is stored. The process idea is at first viewed with relevance single association FAM to grasp the extended for multiple cases. One association FAM associates one antecedent Ai with the only ensuant Bj encoded in an exceedingly fuzzy rule. generally one association FAM system encodes and method n parallel a FAM bank of rules (Ai,Bj),-------(An,Bm).The given input is allowed to pass all the weather in antecedent layer and allowed to method.The obtained relevance is considered as associated weight vector and combines with the consequent layer to check the rules satisfaction. Based on the weight vectors and rules satisfaction response the network is retrained if necessary otherwise the obtained output is defuzzifed and given as output. Training in sense comparing the obtained information and striving for exactness with respect to existing(stored relevance on behalf of semantics) and rules satisfaction, the knowledge base used to assess the information is imprecise or not and how it can be managed for future use based on acquired weight vectors, else the exact is retrieved from database. just in case of rule base the fuzzy answer associate the input vector A with every of N FAM matrices Mi and additively lay over the N known as vectors B1 I, i=1,2…..N. therefore the recalled vectors area unit super obligatory instead of N matrices Mi. The input vector A activates N completely different decree parallel however to variable degrees. The recalled vector B equals the weighted add of the individual recalled vectors B1K.

$$B=\sum_{I=1}^{N} w_k B^1_k \qquad (2)$$

Where the weights wk indicate the quality or the strength of Kth FAM rule(Ak,Bk).The recalled vector B is that the normalized add of the match vector B1K.Now the defuzzified output B to end in one crisp worth of the output universe of discourse Y=.Any of the defuzzification schemes may well be utilized. The applying sample is to conclude the work proposed; initial vector is zero i.e. to specify the apparatus position is vertical. Primarily based within the inexactness of the incoming input the apparatus position could vary from left or right. Coming back to our application take a look at the challenge is to create the system to live through the imperfections i.e. not merely correction however build compatible.
Linguistics mapping is for user input to retrieve the relevant data. Typically, the data could also be irrelevancy except for alternative input it should precise therefore we must always not take away it instead, managed to utilize suitably to reduce the interval.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

Connection might not be previous outlined however even be taken as user feedback supported the get output. Typically, the data could also be of unknown connection and illustrious irrelevancy. Intimate ly data may be found in below URL in reference.www.vit.ac.in. The workflow control of the proposed system is shown in Fig 4.

A sample case was taken from these surroundings for testing the performance of projected system. If a student type CSE branch want to register course computer code maintenance when authentication of the priority user the system shows possibility seeable course page course displayed supported course ID, Course Title or school Exits to show. Likely students might not alert to coursed or school base so that they tend to enter title for if students enter information system displays all the prevailing connection. It should have illustrious non-relevance conjointly as told previous sessionCase1: User input course title- computer code engineering. System displays connection data exists in cache as {software system, computer code construction, computer code management system, computer code necessities, computer code principles style, computer code maintenance, computer code reengineering --} here the output consists connection and illustrious non connection. Most of the algorithms work on ranking supported user feedback and exactitude is retrieved

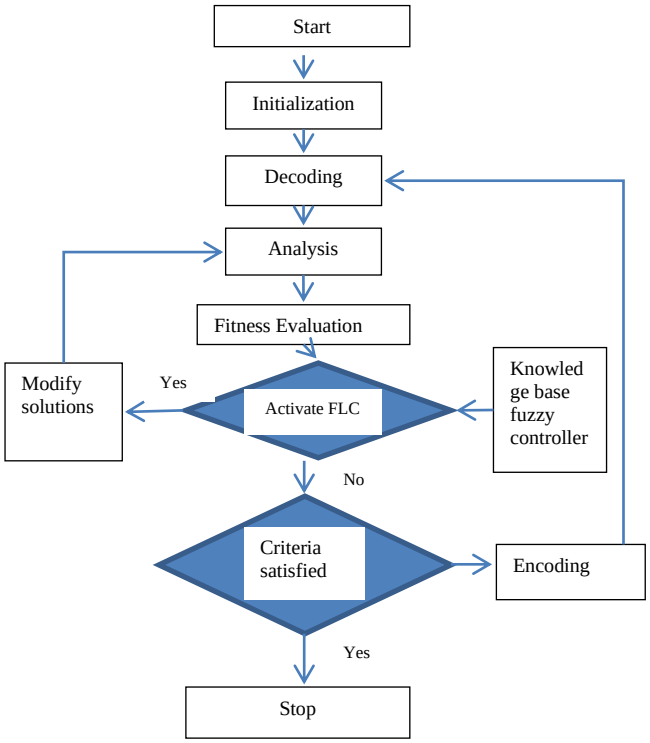

Fig 4 Workflow representation.

Case 2: User input branch name: B.Tech System displays-{software system, computer code construction, computer code management system, computer code necessities, computer code principles style, computer code maintenance, computer code re-engineering} i.e. illustrious connection and possibly illustrious irrelevancy.For case one several solutions is evidenced. Case a pair of chosen for assessment take a look at i.e. managing precisenesstowards inexactitude and contrariw ise by indirect connection feedback, wherever user clicks a link in an exceedingly list of answers, heuristically, the user is probably going curious about the document may be thought to be a regeneration by assumption the snipping is informative enough as overall exploitation click-through knowledge.

Case 3: User input branch name: B. Tech System displays-{software system, software construction, software management system, software requirements, software principles design, software maintenance, software re-engineering} i.e. known relevance and probably known irrelevance. For case 1 and case 2 most solutions proved.Case 3 chosen for assessment test i.e. managing precision towards imprecision and vice versa by indirect relevance feedback, where the user clicks a link in a list of answers, heuristically, the user is likely interested in the document can be regarded as a positive feedback by assumption the snippet is informative enough as overall using click-through data.

V. GRAPHICAL REPRESENTATION

The observed result was plotted as graph shown in Fig 6 and Fig 7 represents the performance status comparatively the proposed approach with existing system. In the graph Y-axis represents the response time in milli seconds and X-axis the user query as input. Here the work was viewed in two aspects one the existing system without assessment and FAM supportive the result obtained .That is shown in Fig 6, Precise from imprecise user input. Secondly the proposed one with imprecise assessment with FAM deployment shown in Fig 7. As observed by analysis our system come out with a difference of significance. This was based on a frequencies test by taking query at various instance and application relevance amidst students under academic system.

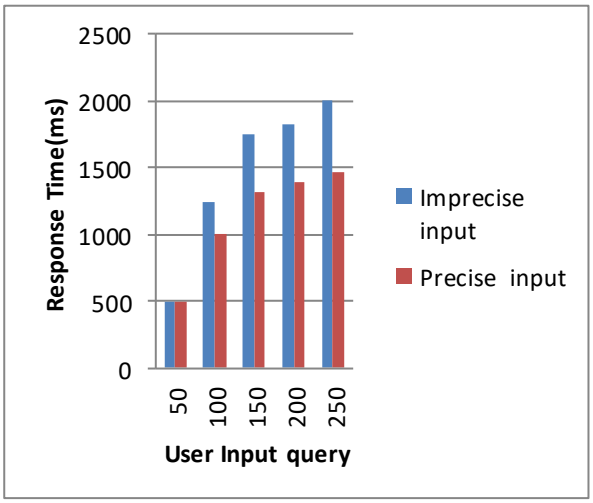

Fig 6 Precise from imprecise information response time.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

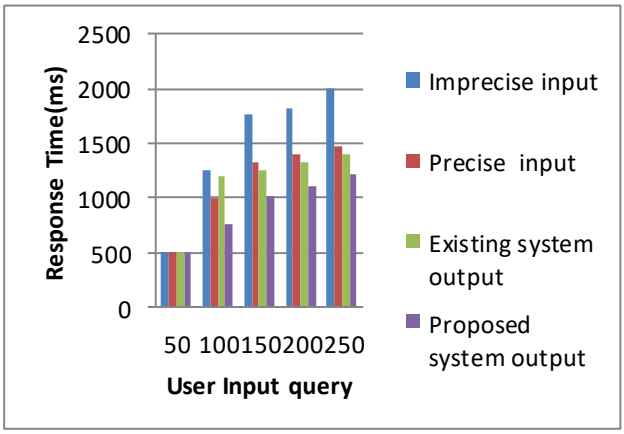

Fig 7 Proposed system

VI. CONCLUSION

Database system users works with assumptions that keep information represents the right set of globe information and create fast call supported in exact information in day to day life for survivals. The challenge is to form the system to endure the imperfections i.e. not merely correction however create compatible and convenient. That's to urge happy output. In this paper the work was strived for this. Initially the data analysis was made to identify the level of data precise imprecise that user possibilities to post as input. In this paper the FAM technique was deployed we observed a quite difference with the existing approaches comparatively. This can be extended for cloud and ubiquitous base real time applications such as hospitals and railways where enormous dataflow is in day to day survivals.

REFERENCES

1. Mohankumar, P. and Vaideeswaran, J., Assessment on precision-imprecision essentials in Semantic query processing. *Indian Journal of Science and Technology*, *8*(13) 2015.
2. Mohankumar, P. and Vaideeswaran, J.,. Imprecision Delicacy in Semantic Based Query Processing. *International Journal of Applied Engineering Research*, *9*(18), 2014 pp.5181-5196.
3. Ren, Q. and Dunham, M.H., *Semantic Caching and Query Processing. Southern Methodist University*. TR-98-CSE-04 1998.
4. Paulheim, H. and Pan, J.Z., 2012. Why the semantic web should become more imprecise. *What will the Semantic Web look like, 10 2012.*
5. Nambiar, U.,. *Answering imprecise queries over autonomous databases* (Doctoral dissertation, Arizona State University) 2005.
6. Ahmad, S.F., Sadreddini, M.H. and Jahromi, M.Z.,IQPI: An incremental system for answering imprecise queries using approximate dependencies and concept similarities. In *Proceedings of the World Congress on Engineering* (Vol. 1).2007
7. Kosko, B., Fuzzy associative memories.1991
8. Zemankova-Leech, M. and Kandel, A., . *Fuzzy Relational Data Bases: A Key to Expert Systems* (Vol. 84). Verl. TÜV Rheinland.1984
9. Ton-That, A.H., Cao, N.T. and Choi, H.I., Fuzzy Inference Systems based on Fuzzy Associative Memory with Adjusting Algorithm for Selecting Optimal Membership Functions. *International Journal of Intelligent Information Processing*, 2013 *4*(3), p.46.
10. Baldwin, J.F. and Zhou, S.Q.,A fuzzy relational inference language. *Fuzzy sets and Systems*, *14*(2), 1984 pp.155-174.
11. Kerre, E.E. and Chen, G., An overview of fuzzy data models. In *Fuzziness in Database Management Systems* (pp. 23-41). Physica-Verlag HD. 1996
12. Prade, H. and Testemale, C., Generalizing database relational algebra for the treatment of incomplete or uncertain information and vague queries. *Information sciences*, *34*(2),1984 pp.115-143.M.
13. Van Schooten, A., 1988. *Design and implementation of a model for the presentation and manipulation of uncertainty and imprecision in databases and expert systems* (Doctoral dissertation, Master's thesis, University of Gent).1988
14. Buckles, B.P., Petry, F.E. and Sachar, H.S., A domain calculus for fuzzy relational databases. 1989.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)