ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 11, Issue 4, pp: (29-36), Month: October - December 2023, Available at: www.researchpublish.com

Theoretical framework towards Crop Yield Classification using Fuzzy SQL and Fuzzy Relational Database

[1]Kile, Awuna Samuel, [2]Ukeyima, Dooshima Joyce, [3]Adegoke, O. Folake

[1]Department of Mathematical Sciences, University of Maiduguri, Maiduguri, Nigeria.

[2]Department of Computer Science, Benue State University, Makurdi, Nigeria.

[3]Department of Computer Science, Prince Abubakar Audu University, Anyigba Kogi State, Nigeria.

DOI: https://doi.org/10.5281/zenodo.10068420

Published Date: 03-November-2023

Abstract: **Inefficient classification of agricultural products has resulted to failure or poor market value for many crops. The need for using technology in enhancing crop yield classification cannot be overemphasized. This work used fuzzy logic system in performing crop yield classification and considered yam for the representation. Attributes of yam like size, length, age, date of maturity, and others are considered in this work. Fuzzy rules for the classification are formed. A fuzzy SQL expression using Create, Insert, Select and Classify statements for the implementation are considered. The evaluation shows that classifying crop yield using fuzzy systems adds great value to post harvest yield management and enhances crops marketing and storage value with high efficiency.**

Keywords: **Crop Yield, Fuzzy Classification, Fuzzy Logic, SQL, System.**

I. INTRODUCTION

Agriculture has been the main human occupation since the creation of man. Presently, the farming population produces 98% of the total food consumed in Nigeria and has the capability of reducing poverty in Nigeria to about 7%. This informs why efforts should be made to strengthen and standardized every agricultural process, as such, the need to implement systems that will properly perform best crop yield post-harvest processes for effective utilization. This has been done using various approaches but some other methods can still be employed. This study considers the use of fuzzy SQL in this direction.

According to [1], expert systems and process control use fuzzy logic as a form of rule-based decision making. In contrast to conventional Boolean logic, fuzzy logic permits partial membership in a set. The two-valued nature of conventional Boolean logic refers to the fact that a member of a set either belongs to it or does not. The membership of a member to the set is represented by values of one and zero, with one denoting absolute membership and 0 denoting no membership. Fuzzy logic allows for partial membership, or a degree of membership, which can range in value from zero to one. The result is the development of a fuzzy system.

Ganeshkumar [2], a fuzzy system is a set of variables that are connected when fuzzy logic is used. Based on the current input variable values, a fuzzy controller controls a fuzzy system using predefined rules. The three fundamental components of fuzzy systems are rules, membership functions, and linguistic variables. Fuzzification is a necessary step for a fuzzy system to function properly.

Fuzzification is the process of changing a linguistic variable (a fuzzy number) into a numerical variable (a real number or crisp variable). This is so because linguistic variables, not numerical ones, are used in fuzzy logic.

In this study, agricultural yield is managed by dividing the yields into several classes and categories for sale or storage. Kudlacik [3] expressed that Zadeh introduced fuzzy sets in 1965. It can be used in practical applications to deal with ambiguous and imperfect knowledge. According to Jain and Abraham [4], it has shown to be an effective tool for making decisions as well as for handling and manipulating erratic and noisy data. Since the Select statement extracts important information from databases and tables, efforts in this research will be focused on building a fuzzy relational database, categorizing the yield attributes, and utilizing fuzzy SQL to perform the select operation on the database.

Kuncheva [5] stated that a fuzzy classifier is an algorithm that determines an object's class based on its description. The classifier is also said to predict the class label.

Crops are mostly used as food and also utilized to generate foreign income, which results in high GDP ratings for our nation, Nigeria. Some crops are used as raw materials by various businesses. Numerous agricultural export initiatives have failed, costing the nation millions of naira. Furthermore, because crops are frequently not properly classified and graded, the nutritional content of the goods has frequently not been adequately investigated. Once more, due to lack of classification, grading and categorization, industrial raw materials are often deficient or do not meet the needed standard.

The fuzzy classification method used in this study aims to grade crop yields in order to improve the commercial, dietary, and industrial worth of crops. Additionally,[6] opined that geography benefits from having a standard, well-defined system of crop classification because it makes it evident where each crop will be grown.

Majorly**,** this study aims at designing a crop yield management system using fuzzy relational database and performing some fuzzy SQL operations on the database, with objectives to be achieved as: formulating fuzzy rules for the crop yield fuzzy classification system; designing a fuzzy relational database for the crop yield system and performing some fuzzy SQL operations like ‘Create’, ‘Insert’, ‘Select’ and ‘Classify’ on the database

The study is important to the following categories of people in the following ways:

1. The farmers can easily classify and categorize their farm products for effective marketing and use by using the fuzzy classification system

2. Government can as well use it to determine and set a standard for crop yield to be accepted for exports, industries and preservation in storage barns.

3. The industries can also use the fuzzy classification system as a standard for measuring and determining both quality and quantity of crop yields to be accepted for usage in the industry

II. LITERATURE REVIEW ON FUZZY CLASSIFICATION PROCESS

Fuzzy data are categorized using the fuzzy classification procedure. According to Hudec and Vujosevic [7], users categorize data to determine to which class each categorized record (territorial unit, client, etc.) belongs. They added that fuzzy classes allow decision-makers or analysts to define input attributes and output classes more empathetically utilizing linguistic variables, overlapping classes, and approximation reasoning. Fuzzy classes also better mirror reality. The way that objects are handled varies depending on how many classes they partially belong to.

Hudec and Vujosevic [8] claimed that it is important to fuzzify attributes, define all IF-THEN rules (rule base), process them, and to offer results in a useable and intelligible form when addressing a fuzzy classification problem within a knowledge-based fuzzy inference system (FS).

The broad definition of a fuzzy classifier, as expressed by [5], suggests a multitude of potential models. He said that a fuzzy if-then system, similar to that used in fuzzy control, can be built by specifying classification rules and is the most basic fuzzy rule-based classifier.

Jain and Abraham [4] said the fundamental benefit of fuzzy rule-based systems is that they do not need a lot of memory storage, have a quick rate of inference, and allow users to thoroughly review each fuzzy if-then rule. He said that a fuzzy system is defined by a collection of linguistic claims based on expert knowledge, typically expressed as "if-then" rules.

They again provided a rule for the if-then construction, stating that the result is innocuous if variables 1 and 2 are low and high, respectively.

They continued by saying that a case or an item may be categorised using a set of fuzzy rules based on the linguistic values of its characteristics in a fuzzy classification system. Every rule has a weight, which is a value in the range of 0 and 1, and this weight is applied to the antecedent's number. It consists of two separate sections; the first part is assessing the antecedent, fuzzifying the input, and using any required fuzzy operators.

According to [9], a collection of classes C is presumed in many fuzzy classification applications. There are numerous approaches to obtain such a set of classes. Once such a family of classes is formed, the initial goal is to ascertain the degree c(x) to which each item x under investigation, x of X, belongs to class c of C.

They went on to say that the idea of partition, which creates an organized family of classes, is a crucial concept in categorization. Each class has a strong relationship with the others, demonstrating a certain structure.

Werro [10] stated that fuzzy classification is a logical progression from traditional classification, much like fuzzy sets are a logical progression from classical sets. Each object in a sharp classification is given a single class, indicating that its membership degree is 1 in that class and 0 in all the others. Therefore, the membership of the objects in the classes is mutually exclusive. In contrast, a fuzzy categorization allows things to simultaneously belong to many classes. Each object also contains membership degrees that indicate how much of an object is a member of each class.

Once more, [9] stated that a fuzzy categorization system will be useless in some situations if it does not include object differentiation. This might occur if no explanation can be found or if all classes completely overlap. No two classes should completely overlap, and classification systems should be able to explain anything about each object.

According to [2], fuzzy logic substitutes linguistic variables for numerical ones and refers to the process of fuzzification as the transformation of a numerical variable (a real number or crisp variable) into a linguistic variable (a fuzzy number).

When executing operations on things to be categorised, their attributes, which are properties of these items, are taken into account and divided into four datatypes. Type 1, Type 2, Type 3, and Type 4 are among them.

Type 1: According to Galindo [11], type 1 qualities are those that have "precise data", are traditional or crisp (i.e., they are devoid of imprecision). Nevertheless, we are able to define linguistic labels within its domain and apply them in fuzzy searches. This kind is beneficial for expanding conventional databases and enabling fuzzy queries regarding conventional data.

Type 2: He added that type 2 attributes collect "imprecise data over an ordered referential." In the form of possible distributions over an underlying ordered dominion, these qualities allow for both crisp and fuzzy data (fuzzy sets). It is an improvement over Type 1 that now permits the preservation of vague information, such as "he is about 2 meters tall."

The fuzzification method also makes use of a membership function. A triangular membership function is one of the many membership functions, and it is frequently employed due to its usability and simplicity.

For [2], the flow of processes in fuzzification is depicted in the diagram figure 1.0.

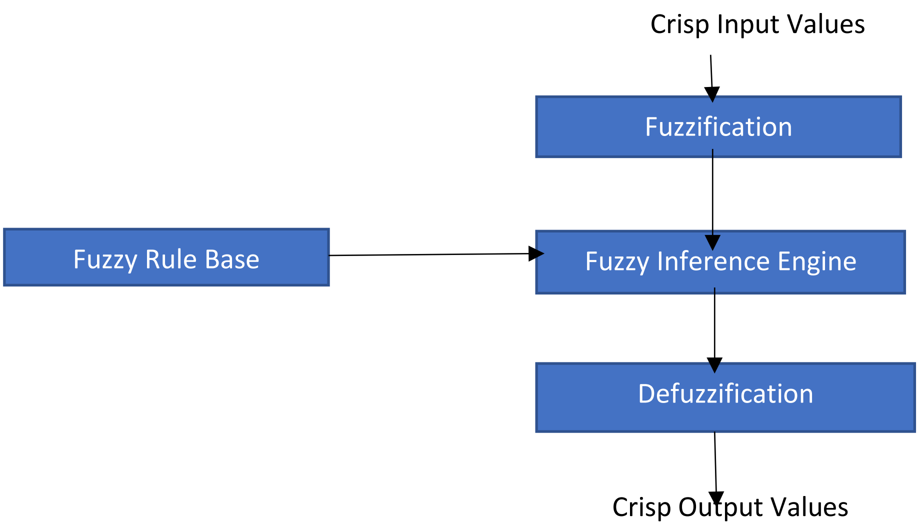

Figure 1.0: Fuzzification Process.

Figure 1.0 shows the fuzzification process. A fuzzy inference system, according to [4], is a rule-based system that uses fuzzy logic rather than Boolean logic to reason about data. Its fundamental composition consists of four key parts. A knowledge base contains both an ensemble of fuzzy rules, known as the rule base, and an ensemble of membership functions, known as the database. It also contains a fuzzifier, which converts crisp (real-valued) inputs into fuzzy values, an inference engine, which uses a fuzzy reasoning mechanism to produce a fuzzy output, a defuzzifier, which converts this latter output into a crisp value, and a fuzzifier, which converts fuzzy values back into crisp values. The inference engine makes decisions by applying the rules found in the rule base. These ambiguous rules specify the relationship between input and output fuzzy variables.

III. METHODOLOGY

The crop yield fuzzy classification system in this research will consider yam crop for classification and the properties of yam that will be used for classification process are its size, length and age. Yam can be of big, small and medium sizes, with some intermediate sizes like very big and very small. The length will be short, medium and long. The type1 and type2 datatypes for the fuzzy attribute will be applied in this research so that both crisp and fuzzy data will be accommodated.

Yam is considered as a case study for this research because it is one of the most table foods produced in Nigeria, especially Benue State and consumed across the globe and recently, there is a high demand for it to be exported to other countries. Efforts in this direction have been made but it resulted to a huge loss partly due to poor product classification and packaging issues.

The diagram in figure 2.0 shows the architecture of the proposed system.

Figure 2.0: Crop Yield Fuzzy Classification System

Figure 2.0 captures the basic structure and operations of the proposed system. The fuzzy classifier will accept inputs in the form of crisps values that will represent the respective sizes, lengths and ages of the crop in question, that is yam. These sizes will be in the form of big, medium, and small. The lengths too will be in a similar form as long, medium, and short. Inputs to the system will be the basic yam attributes considered for this research which are:

(1). Size: It depicts the size of the yam which will be measured in kilograms and lies between 1.5 to 2.0kg.

(2). Length: It depicts the length of the yam and is measured in centimeter and lies between 15 to 40cm.

(4). Diameter: It depicts the diameter of the yam and is measured in centimeter and lies between 3.5 to 4.5cm.

(5). Color: It will show the color of the yam and mostly, colors of yam are white, yellow, brown and others.

(6). Name: It is the name given to a particular variety of yam by farmers/producers.

(6). Time of maturity: It shows the time it will take for a particular specie of yam to mature to consumption stage. Various varieties of yam have various times of maturity ranging from 6 to 12months.

(7). Age of the Crop: The age of the crop in this case stipulates the time of harvest of the crop to when it should be disposed of, that is, consumed or marketed. This study adopts 0 to 6months.

These inputs dimensions are standards for yam marketing as released by Nigerian Agricultural Quarantine Services, (NAQS) and reported by [12].

The last three inputs to the system, that is color, name and maturity time for yams are static. The other inputs, yam name, size, length and age vary and will be used in determining the final output from the system.

These inputs will be acted upon by the processes at the fuzzy classifier and it will result to an output which will be the classified yam grade, using basically the rules for the fuzzy crop yield classification system.

3.1 Fuzzy Classification Rules for the System

The following rules will be applied in the process of the crop yield classification

If yam size is small and length is short, the yam is not good for marketing

If yam size is medium and length is short, the yam is not good for marketing

If yam size is big and length is short, the yam is good for marketing

If yam length is short and size is big, the yam is good for marketing

If yam length is medium and size is medium, the yam is good for marketing

If yam length is long and size is small, the yam is good for marketing

If yam age is less than 6 months, then yam is good for marketing

If yam age is more than 6 months, then yam is not good for marketing.

3.2 Fuzzification Process

The process of fuzzification will start now that the fuzzy categorization rules have been established. The fuzzy values will be changed into crisp values using a membership function. Galindo, Urrutia, and Piattini [11] provided the triangular membership function that should be employed in this situation. Its lower limit a, upper limit b, and modal value m serve as its definitions, such that a m b. When the value equals the value m - a, we refer to the value as the b-m margin. The triangular function's expression is as follows:

$$A(x) = \begin{cases} 0 & \text{if } x \leq a \\ (x-a)/(m-a) & \text{if } x \in (a, m] \\ (b-x)/(b-m) & \text{if } x \in (b, m) \\ 1 & \text{if } x \geq b \end{cases} \qquad (1)$$

For the inference mechanism, this work considers the Sugeno method as expressed by [13].

Since we have eight rules for our classification system, the Sugeno method will be represented as:

$$\left.\begin{aligned} \alpha_1 &= A_1(x_0) \vee B_1(y_0) \\ \alpha_2 &= A_2(x_0) \vee B_2(y_0) \\ \alpha_3 &= A_3(x_0) \vee B_3(y_0) \\ \alpha_4 &= A_4(x_0) \vee B_4(y_0) \\ \alpha_5 &= A_5(x_0) \vee B_5(y_0) \\ \alpha_6 &= A_6(x_0) \vee B_6(y_0) \\ \alpha_7 &= A_7(x_0) \vee B_7(y_0) \\ \alpha_8 &= A_8(x_0) \vee B_8(y_0) \end{aligned}\right\} \qquad (2)$$

From equation (2) above, α_1 to α_8 stands for the eight rules to be used for the classification, the A's represent the size and the B's represent the length.

The individual rule outputs for the fuzzy crop yield classification system are derived by the expressions below:

$$\left.\begin{aligned} C'_1(\text{Ꞷ}) &= (\alpha_1 \vee C_1(\text{Ꞷ})) \\ C'_2(\text{Ꞷ}) &= (\alpha_2 \vee C_2(\text{Ꞷ})) \\ C'_3(\text{Ꞷ}) &= (\alpha_3 \vee C_3(\text{Ꞷ})) \\ C'_4(\text{Ꞷ}) &= (\alpha_4 \vee C_4(\text{Ꞷ})) \\ C'_5(\text{Ꞷ}) &= (\alpha_5 \vee C_5(\text{Ꞷ})) \\ C'_6(\text{Ꞷ}) &= (\alpha_6 \vee C_6(\text{Ꞷ})) \\ C'_7(\text{Ꞷ}) &= (\alpha_7 \vee C_7(\text{Ꞷ})) \\ C'_8(\text{Ꞷ}) &= (\alpha_8 \vee C_8(\text{Ꞷ})) \end{aligned}\right\} \qquad (3)$$

Then the overall system output is calculated by *ANDing* the individual rule outputs

$$\left.\begin{aligned} C_1(\text{Ꞷ}) &= C'_1(\text{Ꞷ}) \wedge C'_2(\text{Ꞷ}) \wedge C'_3(\text{Ꞷ}) \wedge C'_4(\text{Ꞷ}) \wedge C'_5(\text{Ꞷ}) \\ &= (\alpha_1 \vee C_1(\text{Ꞷ})) \wedge (\alpha_2 \vee C_2(\text{Ꞷ})) \wedge (\alpha_3 \vee C_3(\text{Ꞷ})) \wedge (\alpha_4 \vee C_4(\text{Ꞷ})) \wedge (\alpha_5 \vee C_5(\text{Ꞷ})) \\ &\quad \wedge (\alpha_6 \vee C_6(\text{Ꞷ})) \wedge (\alpha_7 \vee C_7(\text{Ꞷ})) \wedge (\alpha_8 \vee C_8(\text{Ꞷ})) \end{aligned}\right\} \quad (4)$$

3.3 Defuzzification Mechanism

The Centroid defuzzification approach will be used to complete the defuzzification task, which aims to identify a single crisp value that best characterizes the fuzzy collection. The formula in [13] description of this strategy says:

$$Z_0 = \frac{\int Ax(x)dx}{\int Ax(x)dx} \quad (5)$$

where Z_0 is defuzzified output, A_i is a membership function and x is output variable.

3.4 Fuzzy Relational Database and Fuzzy SQL

Since the attributes for the system have been known, a relational database can be created for the system using fuzzy SQL. The syntax that will create the relation database is:

CREATE DATABASE Crop_Yield;

This syntax will create the database for the system, named *Crop_Yield.*

Table(s) will be created from the database. The table to be created next using fuzzy SQL will contain basic attributes of yam as the chosen crop considered for this research. These attributes are name, size, length, color, age, and maturity date. The name chosen for this table is *Yam.*

Syntax:

```
CREATE TABLE Yam (
        fc:id fv:1 INT FTYPE1 PRIMARY KEY AUTO_INCREMENT,
        fc:Name  fv:2 FTYPE2,
        fc:Size fv:3 FTYPE2,
        fc:Length fv:4 FTYPE2,
        fc:Color fv:5 FTYPE1,
        fc:Age fv:6 FTYPE1,
        fc:Maturity_Date fv:7 FTYPE1
 );
```

When the table is created, a Fuzzy SQL *Insert* operation will be done on the table. The syntax for the fuzzy SQL *Insert* statement is:

```
INSERT INTO Yam (fc: id, Name, Size, length, Color, Age,  Maturity_Date)
    values (fv: 1,2,3,4,5,6,7);
```

After creating the insert statement, a fuzzy SQL Select statement will now be created. The syntax is as thus:

```
    SELECT fc: Name, Size, Length, Age
    FROM Yam
WHERE size≤2.0kg, length≤40cm, and Age≤6months;
```

When the Select operation is done, fuzzy classify operation will now be performed which classifies the yams so that the right category will go to the market as expected, majorly the export market or for storage in barns.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 11, Issue 4, pp: (29-36), Month: October - December 2023, Available at: www.researchpublish.com

The syntax for the fuzzy Classify statement is shown below:

CLASSIFY into Grade_A
FROM Yam
WITH Size is≤3.5, Length is≤2.0 and Age is≤6

This fuzzy query will return yams of size less or equal to 3.5kg, length less or equal to 2.0cm and age less or equal to 6 months and classify them as grade A. This category of yams will be the most suitable category for marketing or export.

From the above processes, an algorithm can be established for the implementation of a fuzzy SQL classifier. The algorithm is represented as thus:

Step 1: Identify crop(s) to classify

Step 2: Consider the classification index(es) or input(s) parameters (attributes) to use for the classification

Step 3: Formulate fuzzy rules to be applied for the fuzzy classification process

Step 4: Perform Fuzzification

(i) Translate the real values into fuzzy values

(ii) Transmit the fuzzy values to the inference engine to obtain fuzzy output

(iii) Defuzzify the output by converting them into crisp values

Step 5: Create a fuzzy relational database for the classification system

(i) Create fuzzy relation(s) for the fuzzy classification system

(ii) Perform insertion of fuzzy values on the fuzzy columns created

(iii) Perform a fuzzy selection operation

Step 6: Classify the selected output(s) using fuzzy classification query

IV. EVALUATION AND CONCLUSION OF THE FUZZY CROP YIELD CLASSIFIER

Fuzzy SQL is known to return results of queries that add more meaning to situations since it evaluates intermediate data of attributes too. Same has been achieved in this study which covers crops yield with particular reference to yam using attributes like name, size, length, age and maturity date. Most importantly, a fuzzy classification using fuzzy SQL is achieved in this work. This informs the beauty of fuzzy logic which has general applications, as can be seen in the agricultural sector in this research. A fuzzy classifier is fundamentally considered here which categorizes agricultural produce into grades so that it can be used for the respective purpose it is meant, enhancing more credibility to crop yields post-harvest operations.

This research can further be enhanced by considering other attributes which may be in a different domain and using fuzzy SQL and fuzzy classification in different areas with full implementation. Also, classification using approaches like neural networks can also be considered in subsequent researches.

REFERENCES

[1] V. Malik, A. Gautam, A. Sahai, A. Jha, and A.R Singh. “Satellite Image Classification Using Fuzzy Logic”. *International Journal of Recent Technology and Engineering,* vol. 2(2), pp. 204-207, 2013.

[2] D. Ganeshkumar. “Fuzzy Logic Based Classification of Vibrations Using LabView”. Thesis. [Online]: shodhganga. inflibnet.ac.in>bitstream, 2009.

[3] P. Kudlacik. “Fuzzy SQL Queries in Standard SQL Database”. *Journal of Medical Informatics and Technologies,* vol 25, pp. 28-37, 2016.

[4] R. Jain. and A. Abraham. “A Comparative Study of Fuzzy Classification Methods on Breast Cancer Data”, presented at 7th International Work Conference on Artificial Natural Neural Networks, IWANN’03, Spain, 2003.

[5] L.I. Kundcheva. (2008). "Fuzzy Classifier". *Scholarpedia,* 3(1), 2925. Available: http://www.scholarpedia.org/w/index.php?title=Fuzzy_classifiers&action=cite&rev=133818. [Dec, 12, 2020].

[6] Balasubramanian, A. (2014). Classification of Agricultural Crops. *Technical Report, Center for Advanced Studies in Earth Science, University of Mysore, Mysore.*

[7] M. Hude and M. Vujosevic. (2009). "Selection and Classification of Statistical Data Using Fuzzy Logic".[Online]: https://www.researchgate.net/publication/240615894_Selection_and_Classification_of_Statistical_Data_Using_Fuzzy_Logic.

[8] M. Hudec and M. Vujosevic. "Fuzzy systems and Neuro-fuzzy systems for the Municipalities Classification", in Eurofuse anniversary workshop on "Fuzzy for Better", M. Pupin Institute, pp. 101-110, 2005.

[9] A. Amo, J. Montero, G. Biging, and V. Cuttelo. "Fuzzy Classification Systems". *European Journal of Operational Research*, vol. 156, pp. 495-507, 2004.

[10] C. Werro. "Fuzzy Classification of Online Customers". PhD Thesis, University of Fribourg, Switzerland, 2008.

[11] J. Galindo, A. Urrutia, and M. Piattini. *Fuzzy Databases: Modeling, Design and Implementation.* Idea Group Publishing, London, 2006.

[12] C. Nnabuife. "NAQS Releases More Conditions for Yam Export". *Nigerian Tribune* (Dec. 5, 2017).

[13] E. Kozlowska. "Basic Principles of Fuzzy Logic". Internet: http://access.feld.cvut.cz/view.php?cisloclanku=2012080002. Jan. 20, 2022. [2012].