I. INTRODUCTION

Data mining is an approach to extract data with efficient and reliable way from huge amount of data in database. It become difficult to analyze patterns for extract if volume is high. The mining means sorting the large dataset to identify the patterns and relationships in it. It can also help in problem-solving in the data analysis process. There are four stages in data mining. They are data sources, data exploration, Modeling, and deploying models. This technique is used in many fields such as mathematics, cybernetics, genetics, education, sports, medicine, marketing, etc.; web mining is a type of data mining used in customer relationship management. It also consist of parameters like association rule mining, sequence or path analysis, classification, clustering, and data forecasting.[2]

Fuzzy logic in data mining involves computation of data based on predictive models and its categories. In the traditional method this was done with “TRUE” or “FALSE”. Algorithms using Fuzzy logic have increased the power of decision making.[3]

II. WHAT IS ASSOCIATIVE RULE MINING?

Association Rule Mining :

Definition 1: Let PT be a set of patients, which contains different patients PT1, PT2, PT3,……PTn which may occur in different transactions, $PT = \{PT1, PT2, PT3,\ldots, PTn\}$.

Definition 2: Let ST sets of symptoms contains different symptoms $ST1, ST2, ST3\ldots STm$: $ST = \{ST1, ST2, ST3,\ldots, STm\}$ where $ST \subseteq PT$ in transactional data base DB.

Definition 3: An association rule represented in the form of an implication of $ST1 \rightarrow ST2$ where $ST1$, $ST2 \subseteq ST$, $ST1 \cap ST2 = \phi$, $ST1$ is called the antecedent and $ST2$ is called consequent. Support (ST) of an association rule is defined as the percentage of records that contain the total number of records in the database.[4]

$\text{Support, } ST(ST1 \rightarrow ST2) = \text{Supp}(ST1 \cup ST2)$
III. INFERENCE RULE MINING

The inference engine is a software tool that deduces new knowledge or fact by using domain fact knowledge to discover inference knowledge. Basically, it is one of the logic unit in the expert system.[5] It uses general rules of inference to reason from the knowledge base and draw conclusion which are not explicitly stated but can be inferred from the knowledge base. Let S be the set of symptoms contains different symptoms $S_1, S_2, S_3, S_4, S_5...,$ $S_n$ and DIS is the set of disease $D_1, D_2, D_3, D_4$. Suppose that we have following rules:

- **R1:** IF $S = (S_1, S_3, S_5)$ Then DIS=$D_1$
- **R2:** IF $S = (S_2, S_4, S_5)$ Then DIS=$D_2$
- **R3:** IF $S = (S_1, S_2, S_5)$ Then DIS=$D_3$
- **R4:** IF $S = (S_3, S_5)$ Then DIS=$D_4$


![Inference System Diagram]

IV. C-MEANS FUZZY LOGIC

The fuzzy c-means algorithm is exactly similar to the k-means algorithm:

1. There need to Choose a number of clusters.
2. Assign coefficients randomly to each data point for being in the clusters.
3. Repeat until the algorithm has converged.
   - Compute the centroid for each cluster (shown below).
   - For each data point, compute its coefficients of being in the clusters.[6]

V. PROPOSED WORK

In this paper, proposed an approach which helps in prediction of diseases symptoms through fuzzy associative rule mining.[7] It follows four phase procedure to achieve required inference from trainee dataset and the phases are as follows:

1. **Data Collection** – Required data needs to collect (Collect data of all suffering people). All Symptoms of disease needs to store into an excel file.
2. **Fuzzy Inference System** – Needs to consider four parts here - rule base, fuzzification, inference engine, and defuzzification. [8]

In rule base, decision will make based in the IF-THEN conditions. At the same time, in the fuzzification, the crisp input, which is temperature, pressure, etc., the Inference engine will trigger the condition by matching the input with rules. Defuzzification converts the rule into crisp input. There are three types of fuzzifier includes singleton, Gaussian, trapezoidal or triangular fuzzifier.[9]
Diagnosis of Disease:

VI. **FUTURE SCOPE**

This system will make the process easy to understand the exact disease according to symptoms which patient have. Prediction changes are high here and accurate to caught it. It is fast processing as well and result will be fast. It can be used in future for other dangerous diseases as well.

VII. **CONCLUSIONS**

From the proposed system, it has been concluded that the fuzzy association rule mining technique with the rule generation will provide the possibility of disease prediction was performed. In the future, it has been improved with more latest techniques.

ACKNOWLEDGMENT

It is a matter of great pleasure for me to submit this report on the Research Paper entitled “Enhancement in Data Mining using Fuzzy Associative Rule-Based Strategy”, as a part of curriculum for the award of “Master of Technology” in Computer Science & Engineering, Rao Pahlad Singh College of Engineering & Technology, Balana (Mahendragarh).

I would like to express my sincere gratitude to the faculty of Computer Science & Engineering Department, Rao Pahlad Singh College of Engineering & Technology, Balana (Mahendragarh) for the constant encouragement, expert advice, guidance, devotion and timely suggestions which helped me at every stage of this work. I would also like to thank to all my family members and friends who have helped me in one way or another.

**Author’s Contributions**

All authors equally contributed in this work.

**REFERENCES**


**Fig. 3 Fuzzy logic for prediction**