



Performance Analysis of Data Mining, Neural networks &SVM and Fuzzy Algebra in Decision Making Process for Machine Learning Dataset

D Sukheja¹, Kriti Ohri, V Baby¹, Tranjeet Sood²

¹VNRVJIET, Hyderabad.
²MB Khalsa College, Indore

Abstract

The way towards mining to information and knowledge from the gigantic information has been composed by a few research analysts as a core research territory in database frameworks, information warehouse and mining, big data and machine learning. Furthermore, the procedure of information and knowledge mining has been used by various sorts of associations with an opportunity to generate the better revenues and expansions of their business by anticipating the future scenario. In present situation, knowledge mining process is a subset of machine learning, in the view of the fuzzy interval approach. Interval Eigen problem of interval fuzzy matrices in max-min algebra are explored. The portrayal of interval eigenvectors which has been exhibited in (Gavalec, Plávka, & Tomášková, in print) is utilized in decision making. Decision making is vital role in terms of machine learning or imparting artificial intelligence into machines which work upon the traditional logic hypothesis. It is a progression which assists a machine to imagine like a human being and for a human being to facilitate his/her distinctive decision making process. This paper describes the knowledge based decision making process and demonstrates a decision support system using fuzzy interval approach.

Keywords: Data, data collection techniques, data mining techniques and knowledge mining process, machine learning, eigen vectors, interval valued fuzzy matrix, decision making.

1. Introduction

In recent developments, the growth and availability of raw data is to take place at an unpredictable rate. This has created a huge opportunity for knowledge discovery and data science plays an important role in a decision making or decision support applications for the various types of organization from micro scale data analysis to macro scale knowledge discovery. Data, Data Collection and Data Mining are the fundamental terms of data science and used in analytical research, Machine Learning and research in decision making process. The data attributes viz. type of data, quality of data and consistency of data are common among this area therefore large numbers of organizations are taking the interest in data science process. Data science process is able to solve the problem and answer the questions of an organization in very effective manner. Data science process is described in fig 1.



Figure 1: Data science process

1.1. Data and Information

As a matter of fact, the term data conveys the meaning of a simple entity. But realistically it is very difficult and complex to define the term data. Data constitutes distinct pieces of information, usually formatted in a special way, it could be referred to as an entity that is recorded and stored it's also defined in terms of fact, text, graphics, image, audio and video that has meaning in the user's environment.

In terms of mathematics, data is only a figure which is required to generate a mathematical model or generated through mathematical model. According to [1], the definition of data in research perspective is sub divided into two levels; Primary Data and Secondary Data. Primary data inculcate the data originated by the researcher for the first time through direct efforts and experience, specifically oriented towards addressing the concerned research problem. This is also known as the first hand or raw data. Secondary data gives an indication of second-hand information which is already collected and recorded by any person other than the user for a purpose, not relating to the currently addressed research problem. In other words, it is generally concluded that Primary data is a real-time data whereas secondary data is the one which relates to the past. Based on characteristics of the data, it could be categorised through various ways, the hierarchy of data is shown in figure 2.

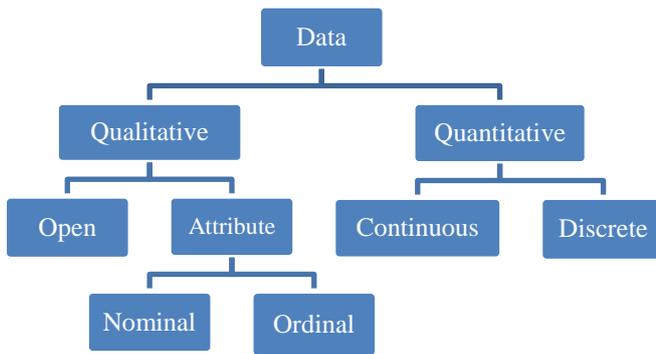


Figure 2: Hierarchy of Information

1.2. Data and Information Collection Techniques

Information accumulation is a fundamental useful part of research in all fields of concentrate not just in software engineering also in data science. As mentioned in figure 2, there are two fundamental sorts of data that clients find themselves working with – and collecting Quantitative and Qualitative data, in this way arrangement or order of information accumulation techniques again depends on qualitative (subjective) information (data) or quantitative information (data).

A. Qualitative Data Collection Methods

For the most part, subjective information accumulation strategies are tedious and expensive to gather the information. The normal and most solid subjective information gathering techniques may be Face-to-Face Personal Interviews, Qualitative Surveys, Web-based questionnaires, Focus Groups, Documental Revision, Observation and Case Studies. Every technique has its own advantage and disadvantage.

Face to Face Personal Interview technique is most frequent information accumulation mechanism to gather qualitative data for exploring reason as a result of its own approach. The questioner will gather information specifically from the objective/subject/protest (the interviewee), on a one-on-one and face-to-face interaction. This is ultimate when information to be acquired must be exceptionally customized. The information gathered through this strategy might be informal, unstructured conversational. In focus group data accumulation system is similar to that of the face to face interview; the distinction is the interview has been conducted in the form of group discussions/debate. To accumulate the versatile subjective research type of information focus group approach is highly recommended, but gathered information is again unstructured and very costly. Qualitative surveys are an exceptionally regular and valuable strategy to accumulate the semi-structured and structured information. The concept utilized in this technique is literature reviews, questionnaires and web-based questionnaires. The quality and reliability of the data is additionally relying upon the prepared questionnaires. Another most reliable and easy data or information technique is Documental Revision. This technique utilizes existing and reliable documents along with other reliable sources of information. It is utilized in a new exploration or investigation viz. literature review or literature survey. The quality of information depends on knowledge of extractors. Observation is another type strategy to gathering the data or information. Yet it is very sensitive and specific relevant to the theme. This technique relies on the estimation procedure or observed behaviour with respect to provide contribution of an instrument/ experiment. The fundamental hindrance of this technique is; sometimes the cost of information is high or potentially sometime it is extremely tedious for the observers to accumulate the information in light of the fact. Under the observation technique information can be interpreted differently using the accompanying mechanisms:

1. Descriptive observations: Jotting down the perceptions straight way from the observations.

2. Inferential observations: Writing down the perception indirectly e.g. physical behaviour of the subject.

3. Evaluative observation: Assuming and subsequently be judgmental from the behaviour. Additionally, ensuring the replication of these discoveries.

All this techniques are defined in [2], [3] and[4]. Naturally, the quantitative data is analysed at first and after those qualitative methodologies are utilized to investigate the trends distinguished in the numerical information.

As per the above discussion, there is various information accumulating techniques accessible. The essential capacity of information gathering technique is accumulation of primary or secondary information in terms of the quantitative or qualitative so that it tends to be bifurcated.

1.3. Data Identification and Separation

The data identification and partition idea is essential and utilize exceptionally significant part in information mining, machine learning algorithm or to outline artificial neural network model or Fuzzy system(Fuzzy Logic). On behave of partition of labelled or unlabelled information from accumulated dataset, distinctive information mining techniques or fuzzy logic/ Fuzzy algebra can be applied. Typically we apply unsupervised learning algorithm for unlabelled information and supervised learning algorithm for labelled information or decision making using fuzzy interval approach. There are numerous situations where unlabelled information is teeming and effortlessly obtained but labelled information often requires an expertization or arithmetically formulated to interpret. Let we assume a chance to expect a case in which a man needs to settle on a choice identified with buying a car with specific highlights that he essentially needs among the sort of numerous things with discernable however similarly contending highlights and thus he/she faces a troublesomely in the decision making that which one should he/she really purchase. To decision and develop an effective decision making process system the necessary steps are

1.3.1. Data Collection

Each car has several fetatures. These fetatures has been broadly classified into three categories such as,

- 1) Basic,
- 2) Average, and
- 3) Extraordinary.

The basic feature set consists of different features namely, fuel efficiency, cost efficiency, pick-up, suspension, air bags, centralized locking, seasonal offers, insurance, top- speed. Then the average set contains the features such as, Servicing, SUV, Seating capacity, Anti-lock braking, Exteriors. Finally the extraordinary set comprises the remaining features such as, Body style, Interiors, Keyless entry, Bluetooth, GPS, Maps, Electronic stability-skid control, Hatchback and Sedan. The survey has been conducted on the three broad features which are shown in the table 1. The model of cars considered for this survey are Maruti 800, Maruti Alto K10, Wagon R, Hyundai Santro, Chevrolet Spark, Tata Indica, Ford Figo, Maruti Swift, Mahindra Scorpio, Honda Brio, Honda Amaze, Chevrolet Beat, Volkswagen Polo, Mahindra KUV 100, Mahindra SUV 100, Hyundai i10, Hyundai i20, Nexa Belano, Nexa Ignis, Ford Fiesta, Maruti Ertiga, Tata Tiago, Toyota Etios, Skoda Laura, Hyundai Verna shown in table 2. Depending upon how much preference is to be given to each asset, in a percentage score from 0 to 100 (basically in a range).

Table 1: Set of Features used for survey

	FEATURES
F ₂ F ₁	Basic (Fuel efficiency, Cost efficiency, Pick-up, Suspension, Air bags, Centralized locking, Seasonal offers, Insurance, Top speed)
F ₂	Average (Servicing, SUV, Seating capacity, Anti-lock braking, Exteriors)

F ₃	Extraordinary (Body style, Interiors, Keyless entry, Bluetooth, GPS, Maps, Electronic stability-skid control, Hatchback, Sedan)
----------------	---

Table 2: Set of Cars used for survey

CARS	
C ₁	Maruti 800
C ₂	Maruti Alto K10
C ₃	Wagon R
C ₄	Hyundai Santro
C ₅	Chevrolet Spark
C ₆	Tata Indica
C ₇	Ford Figo
C ₈	Maruti Swift
C ₉	Mahindra Scorpio
C ₁₀	Honda Brio
C ₁₁	Honda Amaze
C ₁₂	Chevrolet Beat
C ₁₃	Volkswagen Polo
C ₁₄	Mahindra KUV 100
C ₁₅	Mahindra SUV 100
C ₁₆	Hyundai i10
C ₁₇	Hyundai i20
C ₁₈	Nexa Belano
C ₁₉	Nexa Ignis
C ₂₀	Ford Fiesta
C ₂₁	Maruti Ertiga
C ₂₂	Tata Tiago
C ₂₃	Toyota Etios
C ₂₄	Skoda Laura
C ₂₅	Hyundai Verna

Questionnaire used to conduct the survey; designed a questionnaire (shown in table 3) and used to conduct survey among buyers before decision making and the same used to conduct a survey among the reviewers.

Table 3: Model questionnaire

Car Owner Name	
Mobile Number	
Car Number	
Address	
Car Model	Maruti 800 Maruti Alto K10 Wagon R Hyundai Santro Chevrolet Spark Tata Indica Ford Figo Maruti Swift Mahindra Scorpio Honda Brio Honda Amaze Chevrolet Beat Volkswagen Polo Mahindra KUV 100 Mahindra SUV 100 Hyundai i10 Hyundai i20 Nexa Belano Nexa Ignis Ford Fiesta Maruti Ertiga Tata Tiago Toyota Etios Skoda Laura Hyundai Verna
On Road Price	A) Less than 4L B) 4 to 7 C) 7 to 10 D) 10 to 13 E) grater than 13.

Fuel efficiency	A) 8 – 12 B) 12 – 15 C) 15 – 18 D) 18 – 21 E) 21 or more.
Basic (Pick-up, Suspension, Air bags, Centralized locking, Seasonal offers, Insurance, Top speed)	Rate in the given range A) 0 – 20% B) 20 – 40% C) 40 – 60% D) 60 – 80% E) 80 - 100%
Average (Servicing, SUV, Seating capacity, Anti-lock braking, Exteriors)	Rate in the given range A) 0 – 20% B) 20 – 40% C) 40 – 60% D) 60 – 80% E) 80 - 100%
Extraordinary (Body style, Interiors, Keyless entry, Bluetooth, GPS, Maps, Electronic stability-skid control, Hatchback, Sedan)	Rate in the given range A) 0 – 20% B) 20 – 40% C) 40 – 60% D) 60 – 80% E) 80 - 100%

1.3.2. Data Pre-processing

Data pre-processing is the substantial action in data mining, machine learning, and decision making process or in any data science activities. The most part of data pre-processing, entails missing worth substitution, transformation, standardization (normalization), and discretization. The consequence of information pre-processing is the final training set.

A. Missing value replacement

There are numerous originators of missing quality. It is necessary to supplant the missing quality generally the survey could prompt a pointless end result.

B. Normalization

Normalization is also essential task in data pre-processing to diminish redundant variety either inside or between exhibits. Typically, normalization should be done on information with three ways, for example, Z-Score, by decimal scaling or min-max normalization. There are primary two types of normalization anchored in a) distance and b) proportion. Distance based normalization embraces vector construct that is with respect to Euclidian distance and linear based normalization which can correct skewness in information. The Proportion normalization embraces no monotonic normalization which remains on Z-Score. The normalization property necessitates that the scope of an equality or separation measure exists in a settled range.

C. Transformation

Transformation is also an important advance in information pre-processing. Transformation nearly packs the most extreme information. Transformation mainly entails smoothing, aggregation, speculation and discretization. A. Kusiak et al. introduced new transformation method named feature bundling [5]. At the point when this transformation technique applied to a training dataset it adorns classification accuracy of the decision rules generated from this set. Despite the fact that packaging is bound for the integer, categorical and regularizing highlights, it tends to be proceeded with esteem, for instance by utilizing relapse work.

D. Discretization

Discretization is a process of transforming continuous attribute esteem into finite set of intervals to generate attributes with a smaller quantity of distinct esteems. There are numerous categories of discretization methods such as Direct vs. Incremental, Single vs. Multi attribute, supervised vs. Unsupervised, Bottom up vs. Top down.

1.3.3. Data Exploration

Data Exploration is about portraying the information by methods for statistical and visualization techniques. We explore information to bring imperative aspects of that information into focus for promoting examination. Data exploration is of accompanying into two types first is Univariate analysis and second is Bivariate analysis. In Univariate Analysis the variables are explored one by one amid univariate analysis it might be

continuous or categorical based on the type of the variables utilized like Continuous Variables; The central tendency and spread of the variable signifies its nature, Categorical Variables; The distribution of each variable is represented by frequency table. The percentage of values under each category is estimated utilizing two metrics explicitly count and count%. In Bivariate Analysis, It illustrates an association between two variables. This also reveals the association and disassociation between variables at a pre-defined level. This might be viewed as a combination of categorical & categorical, continuous& continuous and categorical & continuous. Continuous & Continuous; The relation between two continues variables might be linear or non-linear. This should be done by assessing the correlation between the variables and can be communicated as:

$$\text{Correlation} = \text{Cov}(X, Y) / \sqrt{(\text{Var}(X) * \text{Var}(Y))} \quad (1)$$

The scope of correlation always ranges between [-1, 1]. Categorical & Categorical, The relation between two categorical variables might be analysed either by a two-way table, stacked column chart or by chi-square test. Categorical & Continuous, The relation between categorical and continuous variables will not demonstrate the statistical significance if the levels for each category is little. It tends to be finished by performing Z-test and ANOVA.

2. Data Mining Techniques for Decision Making Process

Data Mining is an investigative process intended to explore information i.e. data mining techniques are to expel the required information or select explicit information related with the task at that point to design and imagine the patterns for the decision support structure. Various data mining techniques are accessible to implement decision support structure. These data mining techniques are described at two levels: predictive and descriptive. Predictive technique and descriptive techniques also categorized at different subsequent levels. Predictive techniques have classified as (i) Classification, (ii) Regression, (iii) Time series analysis and (iv) Prediction level. The descriptive techniques have classified as (i) Clustering, (ii) Summarization, (iii) Association rules and (iv) Sequential Discovery.

2.1 Predictive data mining Technique

2.1.1. Classification

Classification technique comprises of anticipating a specific result in view of a given contribution. Classification techniques in data mining can process a large quantity of information. It can be worn to envisage categorical class labels, classifies data based on training set and class labels and it might be used for classifying newly accessible information. This technique has been implemented in two steps; During initial step, the model is created by applying classification algorithm on training dataset at that moment in the second step the extracted model is tested against a predefined test dataset to assess the model trained performance and accuracy. So, classification is the process to allocate class label from dataset whose class label is mysterious. K-Nearest Neighbour (KNN), Naïve Bayes and Decision Tree are common classification techniques. KNN is the sort of supervised learning method. The working procedure of KNN classifier is characterized below:

1. Calculate the distance between the attributes of training and test datasets.
2. Sort all the training information in light of the distance esteems.
3. Determine the neighbours (k) which are near to the test information.

4. Assign the majority class of training information for the test information.

The Euclidean distance between the training dataset as well as the test dataset can be derived as,

$$D(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (2)$$

$$D(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (3)$$

Where, $P = \{p_1, p_2, \dots, p_n\}$ is the set of training dataset,

$Q = \{q_1, q_2, \dots, q_n\}$ is the set of test dataset and D is the Euclidean distance.

According to [6], In KNN classification, the class enrolment of the test data sets is as identical as the class of training data sets which are closer to the test information. The steps involve in KNN Algorithm are mentioned

Step 1: Set training data = $TT(p)$,
 Step 2: Set testing data = $TS(q)$.
 Step 3: For all $TT(p)$ and for all $TS(q)$,
 Step 4: Compute the distance D by utilizing the Eqn. (2).
 Step 5: Find the smallest distance k .
 Step 6: Locate the corresponding information
 $TS(q_1), \dots, TS(q_k)$.
 Step 7: Assign $TT(p) = C_i$ //class with majority.

Algorithm 1: KNN

In Naïve Bayes classification technique the Bayesian classifier is constructed in view of the Bayesian network [7]. Naive Bayes classifier is a linear classifier, in which the attributes are considered as independent and have equal weight. In Bayes theorem, let E be the evidence and H be the hypothesis and $E = \{e_1, e_2, \dots, e_n\}$ be the set of samples with n attributes. $P(H/E)$ is the probability that the hypothesis H holds the given evidence E . $P(H/E)$ is the a posterior probability of H adapted on E and $P(E/H)$ is the a posterior probability of E adapted on H . $P(H)$ is the a priori probability of H , and $P(E)$ is the a priori probability of E [8]. It can be communicated as,

$$P(H/E) = \frac{P(E/H) P(H)}{P(E)} \quad (4)$$

Step 1: Locate $S =$ set of training data,
 $Y = \{y_1, y_2, \dots, y_n\} =$ set of sample data,
 $C = \{c_1, c_2, \dots, c_m\} =$ set of class.
 Step 2: If $P(C_u/Y) > P(C_v/Y)$ for $1 \leq v \leq m, u \neq v$.
 Step 3: Then $Y \in C$.
 Step 4: Maximize posterior hypothesis.
 Step 5: By using Bayes theorem,

$$P(C_u/Y) = \frac{P(Y/C_u) P(C_u)}{P(Y)}$$

Step 6: Maximize $P(Y/C_u) P(C_u)$ and $P(Y)$ as consistent.

Step 7: Calculate $P(Y/C_u) = \prod_{m=1}^n P(y_m/C_u)$ and

$$P(C_u) = \frac{\text{freq}(C_u / S)}{|S|}$$

Where

$$P(Y / C_u) = P(y_1 / C_u) \times P(y_2 / C_u) \times \dots \times P(y_n / C_u)$$

Algorithm 2: Naïve Bayes

Decision tree (DT) algorithm is an imperative approach for data mining methods. It is used for both classification and prediction. The decision tree is the diagram like structure that isolates the set of pertinent information into an officially characterized class [9]. Consider a training set $S = \{(a_1, b_1), \dots, (a_n, b_n)\}$, where $\{a_1, \dots, a_n\}$ are the set of feature vectors and $\{b_1, \dots, b_n\}$ are the set of labels. This process is recursive. The nodes will stop growing until the point when they accomplish the stopping criteria (SC). *BestSplit* restores the best split point and *FindSplit* splits the information according to the *BestSplit* point [10].

Step 1: Initialize training sample S , attribute list L .

Step 2: Create node N .

Step 3: If $S(N) \in C$, return N as leaf node.

Step 4: If ($L = \text{null}$),

return N as leaf node, and mark the node as belongs to large number category.

Step 5: Select $A \in I$ (largest information gain) and set $N = A$.

Step 6: For each $a_i \in A$, partition N .

Step 7: If $a_i = A$, generate branch from N and set $S_i =$ obtained sample set.

Step 8: If ($S_i = \text{null}$), mark leaf node with the most number of sample types

Step 9: Else mark it with return value $T(S_i, L - A)$

Step 10: Output the decision tree T .

Algorithm 3: DT

2.1.2. Regression

Regression is a data mining technique which gauges a scope of numeric qualities (alluded to as *continuous esteems*), for an agreed dataset. Regression is an information mining (machine learning) technique which fits an equation to a dataset. Regression investigation is a form of predictive modelling technique which considers the relationship between a dependent (target) and independent variable(s) (predictor). This technique is utilized for the prediction. Illustration: regression is used to predict the cost of an item or administration, given different variables. Regression is utilized across multiple industries for business and marketing planning, financial forecasting, environmental modelling and analysis of trends.

There are various kinds of regression techniques accessible to make predictions. These techniques are mostly driven by three metrics (number of independent variables, type of dependent variables and shape of regression line).

A. Linear Regression

It is deemed as one of the key modelling technique. In this strategy, continuous attribute is affirmed to the reliant variable, while independent variable(s) can be consistent or discrete. The regression line is linear. It establishes a relationship between dependent variable (Y) and one or more independent variables (X) by considering a best fit straight line i.e. regression line. The relation can be spoken as,

$$Y_i = a + b X_i + e_i \quad (5)$$

where a is intercept, b is the slope of the regression line, e_i is the random error term and $i = 1, 2, \dots, N$.

Step 1: Set Class models $C_i \in R^{y \times x_i}$, where $i = 1, 2, \dots, N$ and Vector $v \in R^{y \times 1}$

Step 2: Evaluate $\hat{\beta}_i \in R^{x_i \times 1}$ against each model where,

$$\hat{\beta}_i = \frac{C_i^T v}{(C_i^T C_i)}, \quad i = 1, 2, \dots, N$$

Step 3: \hat{v}_i is compared for each $\hat{\beta}_i$, such as $\hat{v}_i = C_i \hat{\beta}_i$, $i = 1, 2, \dots, N$.

Step 4: Calculate the distance between the original and predicted response variables using the following expression,

$$d_i(v) = \|v - \hat{v}_i\|_2, \quad i = 1, 2, \dots, N$$

Step 5: Select the class with the minimum distance $d_i(v)$.

Step 6: Output class v .

Algorithm 6: Linear Regression

B. Logistic Regression

Logistic regression is used to discover the probability of achievement occasion and disappointment occasion. We should utilize logistic regression when the reliant variable is binary (0/ 1, True/ False, Yes/ No) in nature. The probability can be ascertained as,

$$P(c = \pm 1 | d, a) = \frac{1}{1 + \exp(-c(a^T d + b))} \quad (6)$$

Where data d , weights (a, b) and class label c , d_i is a training instance, $i = 1, 2, \dots, l$ and $c_i \in \{1, -1\}$.

a. Polynomial Regression:

A regression equation is a polynomial regression equation if the power of independent variable is more than 1. In this regression technique, the best fit line is not a straight line. It is rather a curve that fits into the information focuses.

Step 1: Set data = d and weights (a, b) .

Step 2: Set class label = c .

Step 3: Calculate the probability

Step 4: Initialize training instance d_i and labels $c_i \in \{1, -1\}$,

where $i = 1, 2, \dots, l$.

Step 5: Calculate (a, b) , by minimizing the negative log-likelihood,

Step 6: To obtain a simple derivation assume $d_i^T \leftarrow [d_i^T, 1]$

Step 6: Calculate regularization abilities by adding regularization term $\frac{1}{2} a^T a$

Step 7: Output the regularized logistic regression,

$$\min_a f(a) \equiv \frac{1}{2} a^T a + z \sum_{i=1}^l \log(1 + e^{-c_i a^T d_i})$$

Where $z > 0$ is a parameter chosen by the users.

Algorithm 7: Logistic Regression

2.2. Descriptive techniques

Another approach of information mining to extricate the knowledge from huge datasets is known as Descriptive data mining. It is normally used to generate correlation, frequency, cross tabulation, etc. This Descriptive method can be characterized as to discover regularities in the information and to uncover patterns.

i. Clustering

Clustering is a separation of data into groups of similar objects, disparate object into another cluster. It is a way for discovering similarities between information according to their quality. This technique based on the unsupervised learning. It is also categorized at accompanying various levels;

a. Partitioned Clustering: A splitting up of information objects into non-overlapping subsets (clusters) with the end goal that every object is the subset.

b. Hierarchical clustering: An arrangement of nested clusters composed as a hierarchical tree.

The distinctive Clustering Algorithms are K-Means, single linkage algorithms and simulated annealing (SA) based clustering technique

ii. Summarization

Summarization is the process of diminishing the huge capacities of information in a significant and shrewd design with vital and applicable highlights. Summarization procedures like a classification of the mean and the standard deviations are regularly inferred for examinations and imagine the information and to create the report consequently.

iii. Association rules

Association rule mining unwrap the pattern that happens as often as possible among the dataset. It focuses in extracting affiliations, correlations, frequent sequence, and frequent item set and frequent patterns with intriguing quality among the dataset in the record repositories. The association can be communicated as $X \rightarrow Y$.

Support

Support determines how often a rule is applicable to a certain dataset, i.e., the sustainability of the rule is the percentage of exchanges that contain both X and Y among all transactions in the input dataset. Support can be computed as probability of the union set X & Y.

$$\text{Support}(X \rightarrow Y) = P(X \cup Y) = \frac{n(X \cup Y)}{N} \quad (7)$$

Confidence

Confidence determines how frequently the items in Y appear in transactions that enclose X, i.e., the confidence of the rule is the conditional probability of transactions that enclose Y among transactions that enclose X. The conditional probability also can be computed through proportion of support.

$$\text{Confidence}(X \rightarrow Y) = P(X/Y) = \frac{n(X \cup Y)}{n(X)} \quad (8)$$

2.3. Data mining based decision making process

After the deliberately documental update of a wide range of information, information accumulation, and data mining systems which are specified in the past areas of this paper. There are different approaches to discover the information from the gathered information. Information might be as far as content, pattern, trends and so on. This relies upon the mix of information, information gathering strategy and information mining system/machine learning algorithms. The proposed basic incremental learning mining process is characterized in figure 2.

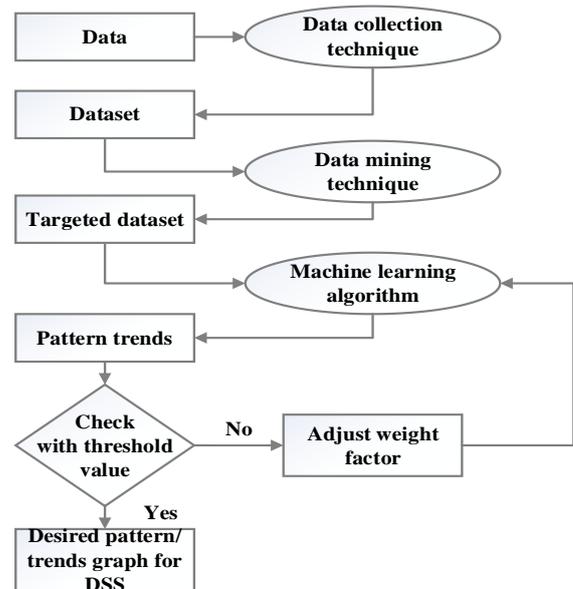


Figure 3: Incremental knowledge mining process

The mentioned process will work as follows:

1. All sorts of information is a contribution to information gathering strategy. According to the sort of information, the structure of information and its attributes, the information accumulation procedure will have used to gather the information. At long last, the information gathering method gathers the information in a particular configuration and returns an informational collection which holds the labelled and/or unlabelled information with different characteristics in single format.
2. With the assistance of expertization and numerical plan, data identification and partition will have applied to differentiate the labelled and unlabelled data construct two datasets one for labelled information and second for unlabelled information.
3. The distinctive re-processing technique can apply on labelled and unlabelled dataset to prepared targeted dataset. On enormous dataset, this is returned by data collection technique and returned targeted dataset after filter to given colossal dataset.
4. This predetermined triggered dataset exchange to the machine learning algorithm or machine learning process which gives the trends and patterns.
5. The scope of trends and patterns are standardized between [0-1] by:

$$N(\text{pattern}) = \frac{\max(\text{pattern}) - \text{current}(\text{pattern})}{\max(\text{pattern})} \quad (9)$$

Where $\max(\text{pattern})$ is the upper limit pattern esteem, $N(\text{pattern})$ is the normalised pattern esteem and $\text{current}(\text{pattern})$ is the current pattern esteem.

6. This trends and patterns will be contrasted with the threshold value δ , if threshold estimation will meet then pattern or trends will be acknowledged by decision support system utilizing of triggered dataset transferred to the machine learning algorithm / machine learning process through supervised or unsupervised learning law in the wake of applying weight adjustment factors, which can be communicated as:

$$\text{Output} = \begin{cases} \text{Accept} & , \text{if } \text{pattern} \geq \delta \\ \text{Weight adjustment} & , \text{otherwise} \end{cases} \quad (10)$$

7. The process will also terminate output are same in last three iterations or to meet threshold estimations.

2.4. Evaluation of Data mining techniques for decision making process

The evaluation of the decision making process using the different data mining techniques has been figured out in this section. Based on the above techniques it has been concluded that suppose that 70% of the buyers who buy product any one of the car by credit card, also buy another one car and, in addition, the car is never sold independently. Therefore, it is quite easy to build a model to predict the purchase of a car with a 70% probability of occurrence using the different data mining techniques.

3. Decision Making Process Using Neural Networks and Svm

3.1. Neural Network

Neural Network (NN) is a numerical (or) computational model in light of biological neural networks. It is also defined as an impersonation of biological neural system. It is also known as Artificial Neural Network (ANN) (or) Simulated Neural Network (SNN) [11].

```

Step 1: Set weight = W
Training set = T
Step 2: Repeat
For ( e ∈ T )
assign O = output (network , e)
Forward pass
T = output for e .
Calculate Error = (T - O)
Calculate Δwi for all weights from hidden to output layer
Calculate Δwj for all weights from input to hidden layer
Update W .
Step 3: Until stopping criterion satisfied.
Step 4: Return network
    
```

Algorithm 4: NN

Neural Network based data mining comprise of three major phases:

- Network construction & training: constructs and trains a three-layer neural network.
- Network Pruning: Aims to eliminating redundant links & units without swelling the classification error rate.
- Rule Extraction: Extracts the classification rules from the pruned network [12].

The neural networks are used to display the complex relationships between inputs and outputs or to discover plans in the information.

3.2. Support Vector Machine

Support Vector Machine (SVM) is a computer algorithm which learns by the paradigm to allow labels to objects. Consider an issue of classifying m points into n-dimensional real space R^N which can be represented as $m \times n$ matrix [13]. Consider a set of input samples $(a_x, b_x), x=1, 2, \dots, N$, where N is the quantity of samples, $a_x \in R^N$ and $b_x = \{+1, -1\}$ has two classes such as, positive class and negative class, i.e. $b_x = 1$ is the positive class and $b_x = -1$ is the negative class [14]. The classification hyper plane in N - dimensional space is $\omega a + z = 0$. Consider a hyper plane $f(X)=0$, which isolates the information.

$$f(X) = \omega^T a + z = \sum_{y=1}^N \omega_y a_y + z = 0 \tag{11}$$

where ω is a vector on N dimensional space and z is a scalar.

$$b_x f(X_x) = b_x (\omega^T a_x + z) \geq 1 \tag{12}$$

where $x=1, 2, \dots, N$. The QP problem is articulated as [13],

$$\min \phi(\omega) = \frac{1}{2} \|\omega\|^2 \tag{13}$$

Where $b_x [\omega \cdot a_x + z] \geq 1, x=1, 2, \dots, N$ if the training data is not detached linearly, the formula must be modified to tolerate the classification violation samples as below:

$$\min \phi(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + C \left(\sum_{x=1}^N \xi_x \right) \tag{14}$$

Where $b_x [\omega \cdot a_x + z] \geq 1 - \xi, x=1, 2, 3, \dots, N, \xi_x \geq 0, x=1, 2, 3, \dots, N$. Introduce Lagrange multipliers, the dual formula for this dilemma can be written as,

$$\max W(\alpha) = \sum_{x=1}^N \alpha_x - \frac{1}{2} \sum_{x,y=1}^N \alpha_x \alpha_y b_x b_y (a_x \cdot a_y)$$

Where $0 \leq \alpha_x \leq C, x=1, 2, \dots, N$

$$\sum_{x=1}^N b_x \alpha_x = 0 \tag{15}$$

By taking care of the above dilemma, the classifier can be expressed as,

$$f(X) = \text{sign} \left(\sum_{x=1}^N \alpha_x b_x (a \cdot a_x) + y \right) \tag{16}$$

Where α_x is the solution of QP hitch [16].

```

Step 1: Set input = (a_x, b_x) where x=1, 2, ..., N, a_x ∈ R^n
and b_x = {+1, -1}
Step 2: f(X) = ω^T a + z = ∑_{y=1}^N ω_y a_y + z = 0
Step 3: Minimize the QP problem as,
min φ(ω, ξ) = 1/2 ||ω||^2 + C (∑_{x=1}^N ξ_x)
Step 4: Calculate the dual Lagrangian multipliers
min L(ω, y, α) = 1/2 ||ω||^2 - ∑_{x=1}^N α_x b_x (ω a_x + z) + ∑_{x=1}^N α_x
Step 5: Calculate the dual quadratic optimization (QP) problem
max W(α) = ∑_{x=1}^N α_x - 1/2 ∑_{x,y=1}^N α_x α_y b_x b_y (a_x, a_y)
Step 6: Solve the optimization problem ∑_{x=1}^N b_x α_x = 0
Step 7: Output as f(X) = sign (∑_{x=1}^N α_x b_x (a · a_x) + y)
    
```

Algorithm 5: SVM

3.3. Evaluation of ANN and SVM for decision making process

Decision making using SVM and ANN has been examined in this section.

3.3.1. ANN for decision making

The ANN classifier uses the back propagation algorithm to select the best choice of the car to the buyers. This type of supervised neural network allows the transfer of knowledge simply by the definition of examples, which provide the specific decision-making process with a very user-friendly approach. These improvements and alternatives will also need the same rigorous evaluation in managerial tasks such as forecasting and decision making.

3.3.2 SVM for decision making

In our learning model we use three different kernels. These kernels are linear, polynomial, and Gaussian. Using these kernels, we can create different classifiers and can improve the management process to select the best choice.

4. Decision Making and Support System by Fuzzy Interval Approach

4.1. preliminaries

In many real applications especially optimization tasks, is possible to use a mathematical structure known under the unifying name external algebras or max-algebras. Max-algebras are concerned with characterization of steady states of the system and with searching of the Eigen value and of the eigenvector, which is generally known as Eigen problem and consequently describing of the eigen space. The group of max-algebras includes a many various algebras. Max-plus algebra and max-min (fuzzy) algebra are the most widely known.

Max-min algebra, sometimes called as fuzzy algebra or Bottleneck algebra is exploited in many application as fuzzy control systems, stability of discrete event systems and so on. By max-min algebra, we comprehend a linear structure on a straightly ordered set with two binary operations maximum and minimum, used comparatively as addition and multiplication in the classical linear algebra. The operations max and min are stretched out to matrices and vectors in a natural way. Due to the increasing complexity of the financial condition and the unclearness of innate subjective nature of human reasoning, Intuition is tic Fuzzy Set (IFS) has received increasing consideration during the time spent for the decision making which is standout amongst the most complex administrative processes in management. The proposed decision making process based on fuzzy interval approach is explained as follows. In particular we have used 2 dimensional arrays or a matrix to authenticate the approach. On the based on designed questionnaire in table 3, Consider l buyers

$B = \{b_1, b_2, \dots, b_l\} (b > 2)$ has to choose from an arrangement of k alternatives $C = \{c_1, c_2, \dots, c_k\} (c \geq 2)$ and each alternative associated with a set of highlights characterized as $f_i = \{f_1, f_2, \dots, f_n\} (n = 8)$ where $f_i \in c_i$. The goal of this approach is to find the alternative which is the most acknowledged to the purchasers.

4.2 Construction of BF and FC matrices for Decision Making

4.2.1. The BF Matrix

We made inquiries to the purchaser about his decisions in terms of highlights, and converted the candidate's preference into interval esteems using fuzzy logic. At that point, we have organized the information gathered through the Questionnaire into interval valued matrix BF, comprising purchasers in rows and highlight choices in columns, which can be spoke to as:

$$B_{l \times F} = \begin{matrix} & F_1 & F_2 & F_3 & F_4 & F_5 & F_6 & F_7 & F_8 \\ B_1 & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] & [0.0,2] & [0.6,0.8] & [0.4,0.6] & [0.4,0.6] & [0.2,0.4] \\ B_2 & [0.4,0.6] & [0.6,0.8] & [0.6,0.8] & [0.0,2] & [0.6,0.8] & [0.2,0.4] & [0.4,0.6] & [0.2,0.4] \\ B_3 & [0.2,0.4] & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] & [0.6,0.8] & [0.8,1] & [0.2,0.4] & [0.6,0.8] \\ B_4 & [0.4,0.6] & [0.4,0.6] & [0.6,0.8] & [0.4,0.6] & [0.6,0.8] & [0.6,0.8] & [0.8,1] & [0.4,0.6] \\ B_5 & [0.4,0.6] & [0.4,0.6] & [0.6,0.8] & [0.6,0.8] & [0.8,1.0] & [0.8,1.0] & [0.4,0.6] & [0.8,1.0] \\ B_6 & [0.4,0.6] & [0.6,0.8] & [0.8,1.0] & [0.4,0.6] & [0.8,1.0] & [0.6,0.8] & [0.2,0.4] & [0.0,0.2] \\ B_7 & [0.4,0.6] & [0.4,0.6] & [0.8,1.0] & [0.4,0.6] & [0.8,1.0] & [0.8,1.0] & [0.4,0.6] & [0.2,0.4] \\ B_8 & [0.8,1.0] & [0.4,0.6] & [0.8,1.0] & [0.4,0.6] & [0.8,1.0] & [0.4,0.6] & [0.8,1.0] & [0.6,0.8] \\ B_9 & [0.6,0.8] & [0.2,0.4] & [0.4,0.6] & [0.2,0.4] & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] & [0.2,0.4] \\ B_{10} & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] & [0.2,0.4] & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] \end{matrix} \quad (17)$$

4.2.2. The FC Matrix

We have collected reviews about the most common reference cars through the assessment we had conducted, with the end goal we can rank each reference car in terms of its highlights in interval esteemed sets, thus arranging the FC matrix as reviewed features in rows and reference cars in columns, which can be spoke to as:

$$F_{l \times C} = \begin{matrix} & C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 & C_8 & C_9 & C_{10} \\ F_1 & [0.0,0.2] & [0.2,0.4] & [0.4,0.6] & [0.4,0.6] & [NI] & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] & [0.8,1] & [0.6,0.8] \\ F_2 & [0.4,0.6] & [0.6,0.8] & [0.6,0.8] & [0.4,0.6] & [NI] & [0.4,0.6] & [0.4,0.6] & [0.6,0.8] & [0.6,0.8] & [0.2,0.4] \\ F_3 & [0.4,0.6] & [0.4,0.6] & [0.6,0.8] & [0.6,0.8] & [NI] & [0.6,0.8] & [0.6,0.8] & [0.6,0.8] & [0.6,0.8] & [0.6,0.8] \\ F_4 & [0.2,0.4] & [0.2,0.4] & [0.4,0.6] & [0.4,0.6] & [NI] & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] & [0.6,0.8] & [0.8,1] \\ F_5 & [0.4,0.6] & [0.6,0.8] & [0.6,0.8] & [0.6,0.8] & [NI] & [0.6,0.8] & [0.6,0.8] & [0.6,0.8] & [0.6,0.8] & [0.6,0.8] \\ F_6 & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] & [0.2,0.4] & [NI] & [0.2,0.4] & [0.6,0.8] & [0.8,1] & [0.6,0.8] & [0.4,0.6] \\ F_7 & [0.2,0.4] & [0.2,0.4] & [0.2,0.4] & [0.2,0.4] & [NI] & [0.4,0.6] & [0.4,0.6] & [0.4,0.6] & [0.6,0.8] & [0.4,0.6] \\ F_8 & [0.6,0.8] & [0.4,0.6] & [0.6,0.8] & [0.4,0.6] & [NI] & [0.4,0.6] & [0.6,0.8] & [0.4,0.6] & [0.6,0.8] & [0.6,0.8] \end{matrix} \quad (18)$$

$$B_{l \times C} = \begin{matrix} & C_1 & C_2 & C_3 & C_4 & C_5 & C_6 & C_7 & C_8 & C_9 & C_{10} \\ B_1 & 1.8 & 2.0 & 2.0 & 2.0 & 0 & 2.6 & 2.6 & 2.0 & 0.8 & 2.0 \\ B_2 & 0 & 1.6 & 3.0 & 2.6 & 0 & 3.2 & 2.8 & 2.8 & 2.4 & 2.2 \\ B_3 & 2.4 & 1.8 & 2.6 & 2.4 & 0 & 2.0 & 2.8 & 2.4 & 1.6 & 1.6 \\ B_4 & 0.6 & 1.4 & 2.8 & 4.0 & 0 & 4.0 & 4.2 & 3.2 & 2.4 & 1.6 \\ B_5 & 0.6 & 0 & 1.4 & 2.0 & 0 & 2.6 & 2.6 & 3.0 & 1.6 & 1.4 \\ B_6 & 0.4 & 1.2 & 2.4 & 1.6 & 0 & 1.2 & 2.0 & 2.0 & 1.6 & 0 \\ B_7 & 0.6 & 0 & 1.2 & 1.8 & 0 & 1.8 & 2.4 & 2.8 & 0 & 0.6 \\ B_8 & 2.0 & 0.6 & 1.4 & 1.2 & 0 & 1.2 & 2.0 & 0.6 & 1.8 & 1.4 \\ B_9 & 2.2 & 1.6 & 0.6 & 0 & 0 & 0.6 & 0.6 & 0.6 & 0 & 2.4 \\ B_{10} & 2.8 & 2.8 & 1.2 & 1.8 & 0 & 2.4 & 1.8 & 1.8 & 0 & 1.2 \end{matrix} \quad B_{l \times C}$$

Note that the similar questionnaire was accommodated the review survey and in addition for the purchaser.

4.2.3. The BC matrix

We will follow the method for attaining two crisp sets by utilizing Intersection, Union operators as a replacement for usual expansion and multiplication task on an interval-valued matrix. For this reason, instead of taking direct intersection of 2 sets, we have taken a intersection of intervals. We have derived the accompanying relations among the purchaser and the item utilizing the general matrix properties.

Property 1: "Person X Feature Intersection Feature X Reference" gives a relation in a set form Person X Reference. , i.e. $(BF) \cap (FC) \Rightarrow (BC)$.

Property 2: "Person Intersection Feature X Reference" gives a relation in a set form Person X Feature Intersection Reference, i.e. $B \cap (FC) \Rightarrow (BF) \cap C$.

Property 3: " Increase/Decrease in Feature Intersection Reference " gives a relation in a set form "Increase/Decreased Feature Intersection Reference" which also relates " feature intersection increased/decreased reference" also relates a set of "feature intersection reference increase/decrease", i.e. $\alpha(F \cap C) = (\alpha F) \cap C = F \cap (\alpha C) = (F \cap C)\alpha$.

After the convergence of person-feature and feature-reference interval-valued matrices, we get a person-reference valued matrix. Now, we analyze the resultant matrix in which each row speaks to a person and every section is a reference, now the amplest person-reference intersection interval for every person will be the most suitable reference option for all other references. From this time, using such matrix logic we can decide on the most appropriate option in an option set containing all other references and hence ease the process of decision making.

4.3. Evaluation of Fuzzy logic for decision making process

From the above outcomes, we come to know that the noteworthy value of (B_i, C_i) is demonstrating that is the most reasonable car model and segment which full-fill the most of maximum possible requirement of buyer. For example C_6 is the best car for buyer B2, similarly buyer 1 prefers C_6 or C_7 as the best cars similarly buyer 3, 4, 5 prefers C_7, C_7 and C_8 respectively as the best than the others.

5. Comparison Analysis and Results

The objectives of the conducted experiments are to analyze the effect of applying various algorithms for decision making on cooperating with communities of machine learning algorithms. In our experiments, we compare the performance of different decision making algorithms that are data mining, ANN & SVM and fuzzy interval based approach. The performance evaluation is conducted on a real-world dataset obtained making survey from the human beings about their buying interest on the car models described in Section 1.3.1

5.1 Performance Evaluation

We take several performance measures related to the machine learning algorithms for instance, sensitivity, specificity, precision, accuracy, and finally F1 score and compare the decision making algorithms that are equipped with machine learning algorithms with the ones that are rational and take actions to enhance their overall performance.

Consider a classification problem that has only two classes such as, positive and negative. Each instance in the data has mapped to either a positive or a negative label. Given a classifier and an instance, there are four possible outcomes:

- True Positive (TP): If the instance is positive and it is classified as positive
- False Negative (FN): If the instance is positive but it is classified as negative
- True Negative (TN): If the instance is negative and it is classified as negative
- False Positive (FP): If the instance is negative but it is classified as positive

Precision

Precision (Positive Predictive Value (PPV)) is a measure of retrieved instances that are relevant. In other words,

$$Precision = \frac{TP}{TP + FP} \tag{20}$$

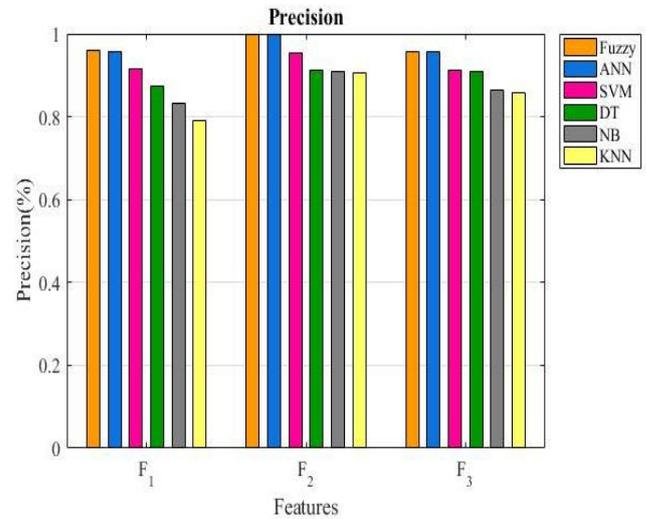


Figure 4: Precision value for different algorithms

Figure 4 shows the precision values of different decision making algorithms for the three feature set. It shows that the fuzzy interval based approach has high precision values compared with the other machine learning algorithms for all the three feature sets.

Recall

Sensitivity (recall or True Positive Rate (TPR)) is the proportion of actual positives which are correctly identified as positives by the classifier.

$$TPR = \frac{TP}{TP + FN} \tag{21}$$

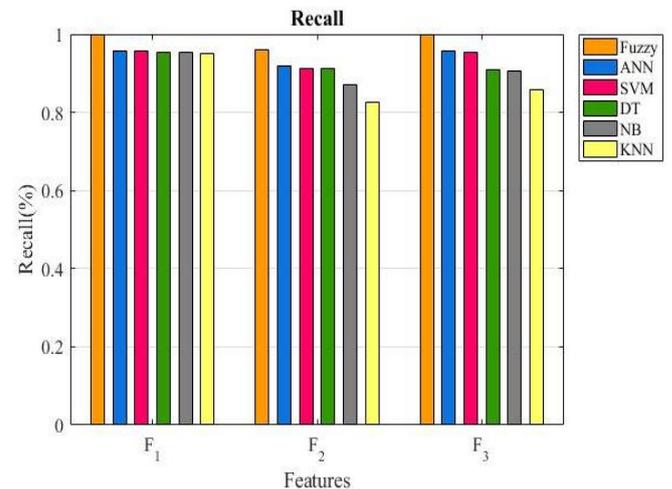


Figure 5: Recall values for different algorithms

Figure 5 illustrates the recall values of different decision making algorithms for the three feature sets. From the above figure we can conclude that the fuzzy performance has better recall when using the three (F₁, F₂, F₃) feature sets

Accuracy

The term accuracy is the simplest scoring measure. It calculates the proportion of correctly classified instances.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{22}$$

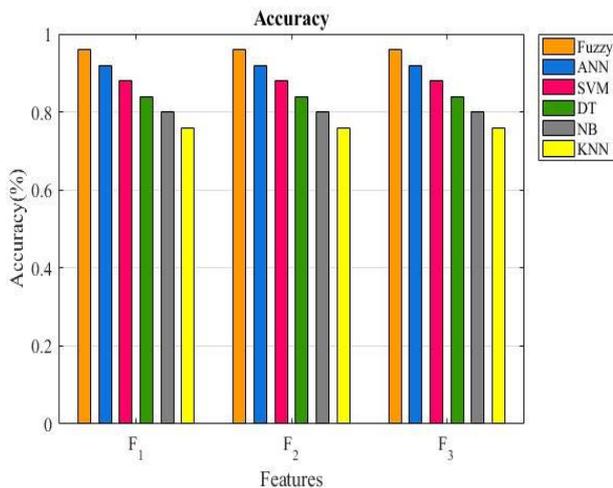


Figure 6: Accuracy values for different algorithms

Figure 6 demonstrates the accuracy of different decision making algorithms for the three feature sets. It shows that the fuzzy interval based approach has high accuracy compared with the other machine learning algorithms for all the three feature sets.

F₁ Score

F1 Score denotes the harmonic mean of sensitivity and precision.

$$F_1 = 2 \cdot \frac{PPV \cdot TPR}{PPV + TPR} \quad (23)$$

$$F_1 = \frac{2TP}{2TP + FP + FN} \quad (24)$$

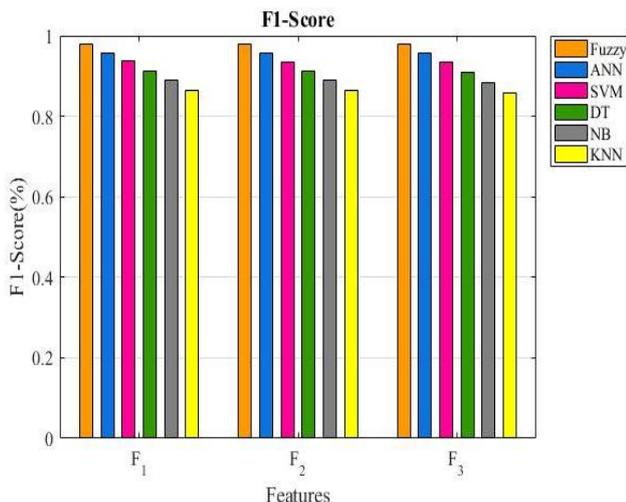


Figure 7: F1 core values for different algorithms

Figure 7 shows the values of F1 score for the six different decision making algorithms based on the three feature sets. It concludes that the fuzzy based approach has high F1 score compared with the other machine learning algorithms for all the three feature sets. Finally, we can draw the conclusion from the experiment results obtained. Thusly, compared with the other decision making algorithms fuzzy based approach has high precision rate, recall rate, accuracy, and high F1 score value. So that we can conclude that the fuzzy interval based approach is the best decision making algorithm compared with the other algorithms.

6. Conclusion

In this paper we have presented the performance analysis of the different machine learning algorithms such as data mining techniques (KNN, NB, DT), SVM & KNN in addition to the fuzzy

interval based approach for the purpose of decision making. We have considered an example of decision making process to buy a car among 25 cars and which are evaluated by three broadly classified unique feature (F₁, F₂, and F₃) sets. We have conducted a survey among the buyers and reviewers using a same questionnaire. For that purpose, here we have compared the different algorithms based on some performance measures. The comparison has been conducted between the six different algorithms based on the feature sets. The results obtained here are more promising and from the obtained results we can also observe the performance of all the selected cars based on its features. From that we can conclude that the fuzzy interval matrix based approach has high performance (in terms of precision, recall, accuracy, and F1-score) rather than the other algorithms. It makes the decision making process more precise and highly reliable compared with the other decision making algorithms.

Reference

- [1] <http://keydifferences.com/difference-between-qualitative-and-quantitative-data.html>
- [2] www.isixsigma.com/dictionary/continuous-data/
- [3] <https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/>
- [4] <https://alaworkshopdata.wordpress.com/data-collection-tools/>
- [5] A. Kusiak, Member, IEEE, "Feature Transformation Methods in Data Mining, IEEE transactions on electronics packaging manufacturing, vol. 24, no. 3, July 2001.
- [6] Adeniyi, D. A., Z. Wei, and Y. Yongquan. "Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method." Applied Computing and Informatics vol.12, no.1, pp. 90-108, 2016.
- [7] Choubey, Dilip Kumar, et al. "Classification of Pima indian diabetes dataset using naive bayes with genetic algorithm as an attribute selection." Communication and Computing Systems: Proceedings of the International Conference on Communication and Computing System (ICCCS 2016). 2017.
- [8] D'Agostini, Giulio. "A multidimensional unfolding method based on Bayes' theorem." Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment vol.362, no.2-3, pp.487-498, 1995.
- [9] Yu, Zhun, et al. "A decision tree method for building energy demand modeling." Energy and Buildings", vol.42, no.10, pp.1637-1646, 2010.
- [10] Meng, Qi, et al. "A communication-efficient parallel algorithm for decision tree." Advances in Neural Information Processing Systems. 2016.
- [11] Lindell, Yehuda, and Benny Pinkas. "Privacy preserving data mining." Advances in Cryptology—CRYPTO 2000. Springer Berlin/Heidelberg, 2000.
- [12] Saxena, Abhinav, and Ashraf Saad. "Evolving an artificial neural network classifier for condition monitoring of rotating mechanical systems." Applied Soft Computing vol.7, no.1, pp.441-454, 2007.
- [13] Pal, Mahesh, and P. M. Mather. "Support vector machines for classification in remote sensing." International Journal of Remote Sensing vol.26, no.5, pp.1007-1011, 2005.
- [14] Cao, Li-Juan, and Francis Eng Hock Tay. "Support vector machine with adaptive parameters in financial time series forecasting." IEEE Transactions on neural networks vol.14, no.6, pp.1506-1518, 2003.
- [15] Zeng, Zhi-Qiang, et al. "Fast training Support Vector Machines using parallel sequential minimal optimization." Intelligent System and Knowledge Engineering, 2008. ISKE 2008. 3rd International Conference on. Vol. 1. IEEE, 2008.
- [16] Widodo, Achmad, and Bo-Suk Yang. "Support vector machine in machine condition monitoring and fault diagnosis." Mechanical systems and signal processing, vol.21, no.6, pp.2560-2574, 2007.
- [17] <https://alaworkshopdata.wordpress.com/data-collection-tools/>
- [18] <https://www.scribd.com/document/273130375/Data-Collection>
- [19] <https://www.thoughtco.com/regression-1019655>
- [20] Saja H. Rasool1 , Faiq M.S.AI-Zwainy1 "Estimating Productivity of Brickwork item using Logistic and Multiple Regression Approaches" SJET, 2016; 4(5):234-243
- [21] Nikam, Orient. "A Comparative Study of Classification Techniques in Data Mining Algorithms" J. Comp. Sci. & Technol., Vol. 8(1), 13-19 (2015)