International Journal of Theoretical & Applied Sciences, 9(2): 260-268(2017)

ISSN No. (Print): 0975-1718
ISSN No. (Online): 2249-3247

Soft Computing- A Journey from Statistics

Shabia Shabir Khan
Department of Computer Science,
University of Kashmir, Srinagar, (Jammu & Kashmir), INDIA

(Corresponding author: Shabia Shabir Khan)
(Received 27 August, 2017 accepted 22 September, 2017)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Knowledge has always been the success factor for any organization (business or technical). Knowledge discovery and advantageous tools of artificial Intelligence have gained a lot of attention from the research and medical field. It can prove beneficial provided the technology is used in a proper way. This paper provides review of research work and contributions in the field of mathematical statistics, artificial intelligence, data mining, machine learning and soft computing. Literature relevant to the analytic aspects of the data has been discussed keeping in view various algorithms based on computational technique. Further, we shall look into the technical aspects, various concepts and issues that made us to move from mathematical field towards the process of soft computing.

Keywords: Computational technique, Statistics, Data Mining, Soft computing, Artificial Neural Network (ANN), Evolutionary computation, Genetic Algorithm (GA), Fuzzy Logic

I. INTRODUCTION

The ever increasing interest in the knowledge discovery made us to move from the statistical aspect towards designing creative algorithms in an attempt to find out the possible efficient solution so as to make strategic decisions as proper as possible. We often face a problem where we have a limited knowledge due to which it becomes impossible to exactly describe the existing state or predict the future events. This can be resolved by introducing more data (current as well as historical). Once we have an organized collection of data (information), typically in digital form, we try to move from operational aspect towards the analytical one. As far as the discovery of knowledge is concerned, the analytical users, being the actual Knowledge workers who take strategic decisions have a more important role to play as compared to the users who work on the operational aspect of a big system and take the tactical decisions. Now, as the application domain is shifted from operational data management to data management for strategic purposes, we need to change the data base design as well, keeping the dynamic nature into consideration. Moving Data from operational phase towards the analysis phase we have several techniques that help in strategic decision making and future prediction.

II. KNOWLEDGE DISCOVERY (*Moving Data from operational phase towards the analysis phase)*

The data from different operational data sources is heterogeneous, huge (with respect to dimension as well as size) and scattered all over the network. It is near to impossible for human intelligence to discover potentially useful information from such a large amount of data, so we need a system that would extract the nuggets of knowledge and help us in strategic decision making. Data Mining is used to extract the nuggets of knowledge or unknown interesting patterns from a dataset, database or data warehouse. It has taken inspiration from Machine learning, artificial intelligence and Statistics.

The three fields might thrive for the same thing but the approaches are different. Artificial Intelligence (AI) is concerned with creation of intelligent agents and usually the systems require the learning capability to explore the intelligence in them. Such systems need the machine learning techniques that deals with design and development of algorithms in order to induce the learning capability. So, Machine learning is considered to be an area within AI. Various Applications of AI lie in Robotics, Strategic planning and Scheduling, Manufacturing and Maintenance etc. Research workers from all the fields try to share the knowledge that have been gained so far in an attempt to provide new technologies and approaches for more understanding and extracting of hidden knowledge in the respective fields. So, one of the effective and important techniques for extracting knowledge from the operational or transactional data is the Data Mining (DM) technique which is considered to be the blend of Artificial Intelligence, Machine Learning and Statistics.

III. STATISTICS - QUANTIFIES NUMBERS

It is an important branch of science that is concerned with the collection and counting and description of the data. It uses the histograms for summarization of the data that provide the summary statistics which helps us in understanding the high level understanding of the data. Some other summary statistics / statistical measures often used are max (maximum among the given predictor values), min (minimum of among the given predictor values), mean (average of the predictor values), mode (most frequent occurring value for predictor), variance (measure of the dispersion of the predictor value from the mean) and median (predictor value that divides the dataset into two possible subsets with equal no. of records). These greatly help in the representation of the data. The word 'predictor' refers to one or more columns that help in predicting something and the 'prediction' refer to the particular column that is dependent on the predictors. Predictors can be continuous (infinite) or categorical (finite). Further the categorical predictor can be nominal (e.g. white, black), ordinal (e.g. adult, teen or infant, or interval [2,8].

IV. ARTIFICIAL INTELLIGENCE (AI) -- *(BEHAVES AND REASONS)*

An important basis for developing intelligent systems is Artificial Intelligence. AI is considered to be the field of Science and Technology that develop the computer programs in order to simulate the working of the biological brain. An important example of such a simulation is the program that plays chess. So, the major focus of AI is on the simulation of human intelligence comprising of reasoning, learning and problem solving. Turing Test is the one that is used to determine the extent to which the performance of the AI program in simulating the human intelligence. Robots like ZAR5 robot work on the concept of AI [3].
The methodological approach of intelligent system is to solve the complex problems efficiently. The intelligence is measured in terms of flexibility, its adaptability, management of uncertain and imprecise data, time complexity, storage, learning mechanism and reasoning [4]. The two main approaches of Artificial Intelligence are to focus on machine that would behave like humans in terms of thinking and acting and focusing on the system evaluation and emulation through computational processes wherein the systems are built on the basis of understanding of human behavior.
As far as working principle is concerned, AI makes much use of certain operators and quantifiers. Some of the important terms that are used in technical understanding of artificial intelligence are listed below [5]:
(i) Preposition Calculus and Predicate calculus
(ii) Quantifiers
(iii) Implication and Modal operators provide an extension to the basic logic.
The areas of Artificial intelligence have been categorized under 16 sub-areas and they include Reasoning, Data mining, Expert Systems, Fuzzy controllers, Machine Learning, Genetic Algorithms, Neural Networks, NLP (natural language processing), TOC (theory of computation) etc. [6, 7]. Some of the important paradigms are listed and explained below:

A. Expert System (ES)

Expert System is a field within AI and is defined as the set of programs/algorithms that provide the expert level functionality within a well-defined domain [1]. Expert system usually refers to the set of programs made up if IF-THEN rules wherein the association has been set up by the concerned human expert. The basic functionality has been presented by the Fig. 1.

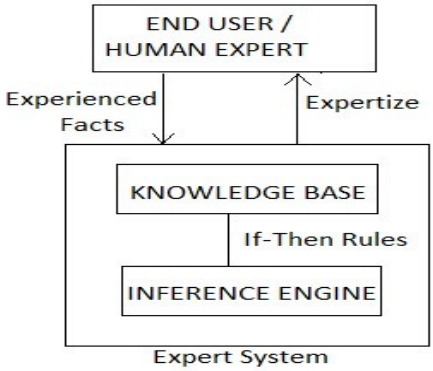

Fig. 1. Basic Architecture of an Expert System.

In the above case the inference engine works on the Inference rules:
(a) Modus Ponens (If P->Q is true and P is true then it implies Q is also true)
(b) Modus Tollens (If P->Q is true and ~P is true then it implies ~Q is also true)
(c) Demorgans Law (If ~(P^Q) is true then it implies (~P V ~Q) is also true and vice versa)
(d) Law of Conjunction (AND operation)
(e) Law of Disjunctions (OR operation)
Expert systems can be classified as under: [2]
(i) Forward chaining expert system e.g. XCON expert system that was built and used by DEC (Digital Equipment Corporation) in an attempt to automatically configure the components (storage space) in the VAX computer systems (large scale) instead of manual search by the user (expert).

(ii) Backward Chaining expert system e.g. MYCIN expert system developed by E.H. Shortliffe in 1976 at Stanford University in an attempt to perform the diagnostic work in the same way as the doctor diagnoses.

Certain drawbacks in expert system are the rule based complexities and the expert dependent knowledge that can lead to error. However it has shown unlimited successes in various fields. Further, in order to address some of the issues, modifications like Fuzzy logic has been introduced.

B. Fuzzy Logic (FL)

Fuzzy logic has the ability to make decisions in the same way as the human would make [9]. The deterministic computer logic has the ability to relate a particular situation to a finite set of states. For example, Color in a black and white image can be defined by the set {black, white} in case of the deterministic and rigorous computer logic. However Fuzzy Logic has been introduced in an attempt to offer a way in which a particular situation could be defined as a human logic would do e.g. a black and white image can also be defined by the grey shades in addition to the deterministic black and white color. This is possible using the degree of member ship, wherein each value in the fuzzy set has an associated degree (value) that will determine its strength or weakness towards a particular feature concerned. Tradeoff condition in FIS means more degrees of freedom and thus better approximation in intuitive meaning of fuzzy sets gets lost. The functional blocks of FIS include a fuzzification interface, rule base, database, decision making unit and a defuzzification interface (to calculate output actual).

Fuzzification interface is issued to calculate the fuzzy inputs. Rule base comprises of a number of fuzzy if-then rules using membership functions of the fuzzy sets that are defined in a database. Various inference operations on the rules are performed by decision making unit. Finally a defuzzification interface for calculating actual output is used to calculate the actual output [10]

Several operations can be performed on the Fuzzy sets like union (maximum), Intersection (minimum) or Inversion. Further the process of de-fuzzification is performed that helps in moving from a fuzzy set towards a point where a control decision could be made. Fuzzy logic has helped in simulating the biological intelligence. Its application has been presented by [2] in correcting the shortcomings of the rule-based systems).

This is an attempt to simulate the brain working with programming methods, for example building a program that plays chess. As far as the future prediction is concerned, we use the concept of Heuristics in rule based Intelligent Expert-System and for the 'almost natural' prediction we try to incorporate the fuzzy logic as well.

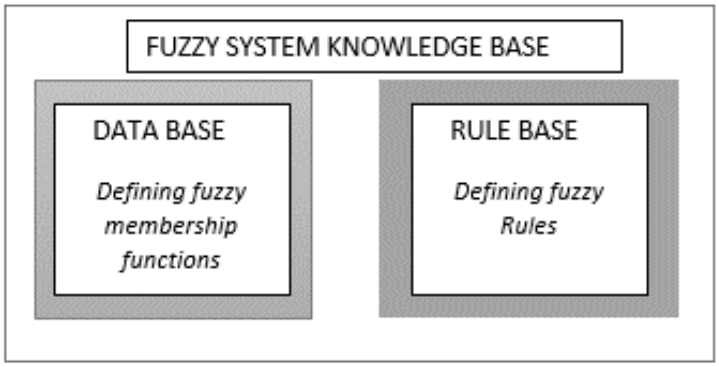

Fig. 2. Fuzzy System knowledge base.

Fuzzy TSK (Takagi-Sugeno) method has the advantage of using measured data for automating fuzzy modelling. The method takes input parameters each representing the membership functions followed by fuzzy (If –then) rules and results in the consequent parameters [11].

C. Genetic Algorithm (GA)

The artificial intelligent Genetic Algorithm has been inspired from the process of Natural evolution and helps in formulating the optimization strategies. It uses the selection operators, crossover and mutation in order to produce the offspring. It is based on the principle: "Survival of the fittest" [12].

In addition to this, we have optimization technique called Genetic algorithm that is based on the concept of natural evolution. GA is considered as the variant of stochastic beam search bearing some resemblance to process of natural selection where in successor state / offspring of the state are generated by combining the two parent states rather than modifying the single state.

The successor generated will populate the next generation depending upon its fitness value (fitness function). The GA works on the basic principle i.e. "survival of the fittest" which is based on following fours factors: Selection, fitness function, crossover and mutation that play important role in finding the hidden unknown solution of any problem e.g. It has been shown how 8-Queen problem can be easily solved using the Genetic Algorithm wherein the goal is to place eight objects/queens in such a way that no queen attack s any other queen diagonally/ horizontally/vertically [13].

This Optimization technique was first proposed by John Holland in 1975. It works on the natural evolution principle in order to solve the optimization problems in a wide variety of domains [14]. GA with the parameters (i.e. No. of input neurons, membership functions, learning and momentum rate) is used to get optimal structure of ANFIS [15,16]. Under such an optimization technique, the population of problem solutions (usually variables represented as binary strings) is maintained in the form of individuals/ chromosomes. These chromosomes consist of several genes which are actually the parameters of network. The problem solution strings are evaluated on the basis of an objective scoring function. At each step, following fitness evaluation, a new population of children chromosomes is produced for the next generation by applying rules/operations over the individuals/chromosomes from the current population/parents [17]

Gene->chromosomes->population->generation

The three basic rules/ operations to create next generation from the current population [18,19] have been listed out below:

(i) **Selection:** (Roulette wheel method) randomly selecting individuals/parents that are used to produce children fir next generation

(ii) **Crossover:** (single or multi-point method and uniform arithmetic method) combining chromosomes on the basis of crossover point to produce new offspring.

(iii) **Mutation:** (boundary method) randomly selecting chromosome string and a random position and negates/alters the bit values.

The steps are repeated for several generations so as to find the best solution and the population of the strings with initial random parameters is created as candidates of best solution.

Roulette Wheel method is used to determine the next chromosomes with randomly selected length wherein previous chromosomes are given chance to cooperate in the next generation so as to get stronger chromosomes [20]. This artificial intelligent technique with some enhancements has been used for optimization of ATM network model that helps in producing good network plans / scenarios under certain conditions [21].

V. DATA MINING – (*PREDICTS WITH MODELS AND EXPLAINS PATTERNS)*

Apart from this, above techniques (specially the statistical inference) can prove to be useful when we want an answer to "who", "what", "where" and "when". However, as far as strategic decision making is concerned, we are still left with some unanswered questions like "How" or "why".

The higher advancement in the field of analysis is the knowledge extraction. Actually, for many cases several concepts from Statistical Inference have been incorporated into data mining but the fundamental difference lies in the way they perform the function. In case of statistics, once a conceptual model is built (null hypothesis), we go for the validation of that hypothesis which leads us to the final acceptance or rejection of the null hypothesis. In contrast, data mining works almost in an opposite way wherein the first step does not start with the null hypothesis. Rather we just have a data set and we don't really know what and which pattern we are looking for. So, here we start by applying the interestingness criteria (notion) over the dataset in an attempt to get some interesting patterns forming the basis of the hypothesis thus the name "Hypothesis discovery" [22].

Inspired by Machine learning and Statistics, the process of data mining has been provided that extracts the nuggets of knowledge (potentially useful information) from the huge amount of complex data. DM helps in finding out the unknown patterns in the data set that help in predicting something that we don't know. Data mining, considered to have been originated from three branches of artificial intelligence --neural networks, machine-learning and genetic algorithms leads us to such advancement [23].

Data mining has been defined as the process of discovering hidden patterns and trends in data (in the DWH or any other historical data set) that are not immediately apparent by just summarizing or aggregating the data, thus involving the search for unknown useful information in a dataset [22, 24, 25]. It was defined as the exploration and analysis process through which meaningful patterns and rules are discovered by automatic and semiautomatic means. However, the same authors, considered "automatic and semi-automatic means" not an appropriate reference rather felt that the process of data mining had shortchanged the role of data analysis and exploration [26].

Data mining is controlled by Interestingness criteria (frequency, rarity, correlation, consistency, periodicity etc.) and we just apply it to the database to find out something or everything which we don't know about or which is interesting according to the specified criteria. Palace gave the fundamental example where data mining was used. It involves the analysis of the local buying patterns. Data Mining has helped a lot in providing the better understanding of the performance and execution of business organization [27]. By extracting such knowledge (useful information) the strategic decisions can be taken in much shorter time which is an important issue for the organizations in the competitive world now-a-days. Various technologies and computational tools have been developed and made available in an attempt to effectively collect and transform the voluminous complex data [28]. Data Mining models has been divided into two categories [29] i.e.

- Descriptive model - that discovers association rules and clusters
- Predictive models - that are generated using classification and regression algorithms.

Data Mining is used to predict the future (predictive analytics) in addition to explain the current or past situation (descriptive analytics). After the interpretation of the information, knowledge can be extracted by identifying relationships among patterns. This can provide the answer to "How". The principles of such relationships describe the patterns and can provide the answer to "why". Now following represent the overall steps for performing the process of data mining:

- Building a fresh Data Warehouse system using the process of ETL or ELT.
- Loading the Data Warehouse /Data base data into multidimensional databases.
- Data Analysis (using application software).
- Knowledge presentation (using graph, table, association rules etc.)

Bill Inmon in 1993 introduced DWH as an integrated collection of data which is time-variant, subject oriented and nonvolatile. DWH is used for retrospective and predictive. Architecture has been provided, to get the most recent updated data in data warehouse wherein the major focus is the utilization of fresh data in order to extract the real time knowledge [17,30]

For performing the extraction of knowledge, various kinds of data mining techniques like regression, classification, summarization and clusterization have been provided among which the two main prediction techniques are classification and regression [31]. Some of the classification algorithms are based on Artificial Neural Network, Naïve Bayesian and Decision Tree. However the process of regression involves the statistical prediction of the variable having numerical continuous attributes in the dataset.

Two main styles of Data Mining (DM) have been introduced by categorizing it into directed data mining (DDM) and undirected data mining (UDM) [26]. DDM is defined as one that uses available dataset to define the model or the predictive modeling using the techniques of classification, prediction and estimation in contrast to the UDM wherein no variable is produced as the target attribute using the techniques of clustering, association rules etc. Also, Data Mining was classified into two kinds of learning methods i.e. Supervised and Unsupervised [32]. Predictive modeling and supervised learning have the same principle of working. In the similar way, undirected DM, descriptive modeling and unsupervised technique represent the same technique. An overview has been presented by [33] to summarize the research (review and technical) on breast cancer diagnosis and prognosis using various supervised (classification) techniques, thus helping in the enhancement of prediction efficiency.

VI. COMPUTATIONAL INTELLIGENCE VS ARTIFICIAL INTELLIGENCE

Artificial Intelligence (e.g. heuristic expert systems) depends on knowledge whereas Computational Intelligence does not depends on knowledge but on the numerical data supplied. It has been explained by distinguishing AI systems from NN and considering FLC, NN, EC as its building blocks. Actual Computational Intelligence is said to have started from IEEE World Congress on Computational Intelligence in 1994. Further, Fogel reviewed the computational Intelligence in 1995 considering adaptation as a key feature of intelligence and inferring computational intelligence as a reason behind the intelligence in AI systems [34].Research has provided and accepted the soft computing techniques like Fuzzy controllers, neural networks , evolutionary algorithms considered to be computational Intelligence paradigms and their hybrids e.g., neuro fuzzy systems , adaptive neuro fuzzy inference system, coactive neuro fuzzy inference system , etc. so as to carry out the intelligent analysis. The two main approaches of Artificial Intelligence are to focus on machine that would behave like humans in terms of thinking and acting and focusing on the system evaluation and emulation through computational processes wherein the systems are built on the basis of understanding of human behavior. The later one is the rationalistic approach of intelligent systems concerning the automation of intelligent behavior and today it's the major focus of the research. Complex Problems can be solved using Intelligent Systems that combine knowledge, technologies and methodologies from various sources.

Research in Artificial intelligence involves the study of the mental abilities like understanding, reasoning, prediction, perception etc. However, in order to improve the understanding of intelligence and to find the faster and accurate solutions, we move from conventional methods to computational soft computing methods. A new approach for machine intelligence proposed by Zadeh was Soft computing which was differentiated by hard computing techniques in terms of Computational Intelligence [35].

VII. THE TECHNIQUE OF SOFT COMPUTING

Hard computing is concerned with analysis and design of physical processes and systems and is based on crisp data and systems, numerical analysis, binary logic, probability and approximation theory. On the other hand, soft computing is based on the analysis and design of intelligent systems using fuzzy data and systems, artificial neural networks, evolutionary computing, probabilistic reasoning and chaos theory in order to solve complex problems especially NP Complete problems wherein no specific algorithm has been given to solve the hard tasks. The principle constituents that lie behind the concept of soft computing are:

(i) FLC which are concerned with imprecision and approximate reasoning,
(ii) Neural networks with learning and curve fitting and
(iii) Evolutionary computation focusing on search and optimization
(iv) Probabilistic theory based on Bayesian concept

All of these are complementary to each other rather than competitive as in hard computing techniques. Artificial Intelligence (AI) is the science that tries to simulate human intelligence and comprises of various fields like Evolutionary algorithms, fuzzy logic, artificial neural network etc. These algorithms are the soft computing techniques that result in accurate predictions as compared to other statistical or non AI predictions [36].

Soft computing techniques have been used widely specially in the engineering and medical fields for the prediction purpose. Soft computing considers its role model as human. In biological world, some of the soft computing experiments that have been performed for molecular computing using DNA evolutionary computation, molecular recognition using DNA Neural Network Computation, membership function encoding in DNA based Fuzzy Systems and graphical representation of genetic programming over DNA molecules.

The start of Soft computing have many earlier influences linked to it like the concept of fuzzy logic and fuzzy sets and complex system analysis by Zadeh in 1973.

VIII. FUSION OF COMPLEMENTARY SOFT COMPUTING TECHNIQUES

One of the important research topic in soft computing is Combining Multiple estimators. Computational analytic tools help in solving or modelling complex or nonlinear problems that are difficult or impossible to solve such as hard problems [37].

(a) **Artificial Neural Network (ANN):** Artificial Neural Network helps in simulating high level structure of brain and solving complex problems. It is concerned with Learning and training mechanism and curve fitting. It has the capability of creating complicated models without having knowledge of the structure however gradually tries to adapt the model through training data.

(b) **Fuzzy Inference System (FIS):** Fuzzy Inference System helps in simulating low level stricture of brain and understanding the human uncertainty. Fuzzy logic expresses knowledge through linguistic terms and understandable rules ("computing with works"). It is concerned with imprecision and approximate reasoning.

(c) **Evolutionary computation (EC):** Evolutionary computation helps in simulating biological natural evolution and thus in achieving effective optimal solution. It is concerned with searching and optimization. Genetic Algorithm searches for the best parameters by applying the principle of natural evolution.

Neural Network has the exciting facility of training and learning mechanism. However in FIS we don't have such king of training algorithm .Although there are some genetic related techniques (linear /nonlinear optimization techniques) but we don't have training algorithm. Due to their individual limitations, these complementary techniques are combined with statistical analysis or with each other in order to solve challenging problems efficiently.

The three intelligent techniques *i.e.* neural nets, fuzzy logic controller and genetic algorithm can be used in intelligent system design for accurate prediction in various fields of work. Combining the advantages of non-linear techniques like FL and ANN proves to be a better intelligent hybrid tool wherein Fuzzy logic is capable of modelling the qualitative human aspects in the form of simple rules (qualification aspect) and ANN (often called black box) being highly powerful in computation has the learning capability (quantification aspect). The combination is of two extremes i.e. Fuzzy Inference System (a completely understandable system) and Neural Network (a black-box system)

The two basic reasons behind creating a hybrid neuro fuzzy model are:

- The learning ability of NN and interpretability (rule base system) in FIS.
- The lack of learning ability in FIS and the lack of interpretability in NN

IX. ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS)

Jang 1993, presented a powerful hybrid estimator tool called ANFIS by combining FIS and ANN [38] wherein fuzzy rules with appropriate MFs are being set by learning process of ANN. The hybrid of ANN and FIS wherein both are nonlinear estimators and model free can help us in dealing with uncertainties of complex data. ANN helps in defining the FL rules due to its learning power (training FIS). It is difficult to define the rules when there exist uncontrolled ambiguities in the data. The major concern in FIS is to determine, for each fuzzy variable, the shape and location of membership function. The efficiency of FIS is judged by the estimation of premise and consequent parts of the fuzzy rule.

So ANFIS uses both the capabilities of NN quantitative aspects and FIS quantitative aspects. The ANFIS technique is actually a function approximation technique that applies learning algorithm over the shape and location of the membership functions.

Here the major principle of combination is using the ability of ANN to train FL. Using the ability of ANN to train FL and using such a hybrid, one can use both capabilities of FL's qualification and ANN qualification aspects.

FIS is a system for modelling the qualitative aspects of human knowledge and reasoning with ability of using human language. Since the FIS cannot be trained, so we use ANN for its best learning capability (gradient descent). Meanwhile, neural networks are implicit and they are unable to use human language). To overcome the inabilities of both the techniques, hybrid model of ANFIS technique was introduced by Jang that integrates Fuzzy Inference System with Neural Network [38].

The summarized working steps in ANFIS technique as listed out by scholars [39] are:

- Defining the basic parameters of membership functions used, known as the initial or premise parameters.
- Determining the product of incoming signals representing the firing strength of a rule
- Normalization of firing strength
- Using Adaptive node function for calculating rule output comprising of consequent parameters

ANFIS focuses on determining nonlinear parameters at first level and linear parameters in the fourth layer in such a way that favorable output is obtained for each desired output. This is done by using the least square method and concept of gradient descent (36,40,41].

Three important techniques used in ANFIS are Takagi Sugeno FIS, Least Square Error and gradient descent training and Input Partitioning. Some of the basic techniques for data partitioning available for data identification used to identify the premise parameters membership functions are grid partitioning (GP), subtractive clustering method (SCM), fuzzy c-means clustering (FCM) [12].

Optimized ANFIS has been experimentally (MATLAB) used for prediction of UTS in FSW joints. The experiment outputs 1200 models with different number of membership functions, types and different combination of variables. The ANFIS model proved to be the best when its prediction results were compared to the prediction results of ANN [42]. Further experiments have been performed over various neuro fuzzy models and compared using UCI repository datasets [40,43].

X. APPLICATION POINT OF VIEW FOR SOFT COMPUTING

Research has been performed focusing on the challenge and that came up with applying artificial intelligent techniques especially fuzzy controllers and neural networks [45]. The greatest challenge in the medical field is diagnostics. Taking application into consideration many researchers have worked in finding its use in area of medicine or healthcare organizations. Research efforts have been made recently to apply the various softcomputing techniques over the medical diagnosis.

The efficient use of soft computing techniques can help in successful decision making that can improve health of the patients and hence the success of health care organizations [44,45].

Classifying patients according to disease subtype or on the basis of being healthy or diseased can be performed using various classification representations like decision tress etc. It is often the concern of physicians to classify patients according to the presence or absence of a disease. However, results from classification trees might not be accurate. To assure the accurate results we have certain machine-learning classification techniques. These include ensembles or bootstrap aggregation (bagging), boosting, random forests, neural networks and support vector machines [46].

Adaptive Neural Network models have been used in breast cancer detection. Breast cancer survivability prediction has been a challenging research problem for many researchers. Algorithms like artificial neural networks and decision trees have been commonly used along with statistical method like logistic regression for prediction especially in case of voluminous datasets. Apart from this a hybrid neuro-genetic approach that integrates Genetic algorithm and back propagation algorithm has been developed for clinical diagnosis and the same has been applied in a medical problem of predicting heart stroke disease for verification of the results.

Lung Cancer is one of the major cause of mortality in the western world as per statistical numbers provided by American Lung Cancer Society. Application of intelligent algorithms for diagnosis and thus earlier detection and treatment can help in predicting the survival rate for patients. Common classification methods such as SVMs & neural networks can be used in such kind of medical diagnostics [47].

Furthermore, Various useful applications of data mining exist in varying degrees and the benefit of this powerful technology has provided solutions to challenging problems apart from medical field as well [48,49].

XI. CONCLUSION

Combining the advantages of non-linear techniques proves to be a better intelligent tool such as using AI paradigm of fuzzy logic in the hybrid tool provides the capability of modelling the qualitative human aspects in the form of simple rules (qualification aspect) and artificial neural network being highly powerful in computation adds up the learning capability (quantification aspect) in the same hybrid tool. Various paradigms of Artificial Intelligence have been reviewed out in the paper. Evolution of soft computing have been chronologically detailed out to show the history pathway from simple mathematics towards the concept of soft computing through computational techniques. Apart from this, differences between statistics and various machine learning algorithms have also been reviewed out.

REFERENCES

[1]. Oke, S. A. (2008). "A literature review on artificial intelligence." *International Journal of Information and Management Sciences,* **19.4**: 535-570.

[2]. Berson, A., Smith, S. (2004). "Data Warehousing, Data Mining and OLAP", Tata McGraw-Hill Education.

[3]. Brunette, E. S., Flemmer, R. C., & Flemmer, C. L. (2009). A review of artificial intelligence. In Autonomous Robots and Agents, 2009. *ICARA 2009. 4th International Conference on (pp. 385-392). IEEE.*

[4]. Krishnakumar, K., (2003). "Intelligent systems for aerospace engineering – an overview," NASA Technical Report.

[5]. Rich, Elaine, and Kevin Knight (1991). "Artificial intelligence" McGraw-Hill, New.

[6]. Wang, Wen-Chuan, *et al*. (2009). "A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series." *Journal of hydrology,* **374**(3): 294-306.

[7]. Zhou, Zhi-Hua, and Min-Ling Zhang. (2007). "Solving multi-instance problems with classifier ensemble based on constructive clustering." *Knowledge and Information Systems,* **11**(2): 155-170.

[8]. Han, J., Kamber, M., and Pei, J., (2011). "Data Mining: Concepts and Techniques" 3rd edition, Morgan Kaufmann.

[9]. Gordon, J.L., (2005). From logic to Fuzzy Logic, Applied Knowledge Research & Innovation, AKRI Ltd.

[10]. Cornelis, Chris, Martine De Cock, and Etienne E. Kerre. (2003). "Intuitionistic fuzzy rough sets: at the crossroads of imperfect knowledge." *Expert systems,* **20.5**: 260-270.

[11]. Li, C., Zhou, J., An, X., He Y., He, H., (2008). "T–S fuzzy model identification based on a novel fuzzy c-regression model clustering algorithm", LNCS 5263, pp.786–795, Springer Verlag Berlin Heidelberg.

[12]. Fakhreddine O. Karray, Clarence De Silva, (2009). "Soft Computing and Intelligent Systems Design- Theory, Tools and Applications", Pearson Education Limited.

[13]. Wyld, David C.; Zizka, Jan; Nagamalai, Dhinaharan, (2012). Advances in Intelligent and Soft Computing, *Proceedings of the Second International Conference on Computer Science, Engineering and Applications (ICCSEA 2012)*, New Delhi, India, Volume **2**.

[14]. Goldberg D.E., (1989) "Genetic Algorithms in Search, Optimisation, and Machine Learning", Addison-Wesly, Reading.

[15]. Vafaei, Saeed, *et al.* (2015). "Enhancement of grid-connected photovoltaic system using ANFIS-GA under different circumstances." *Frontiers in Energy,* **9**(3): 322-334.

[16]. Kampouropoulos, Konstantinos, *et al.* (2014). "A combined methodology of adaptive neuro-fuzzy inference system and genetic algorithm for short-term energy forecasting. *Advances in Electrical and computer engineering".* Volume **14**, number 1.

[17]. Rainardi Vincent, (2008). Building A Data Warehouse, With examples in Sql Server, Apress Publication.

[18]. Hong, Tiancong, and Mrinal K. Sen. (2006). "Real-coded Multi-scale Genetic Algorithm for Pre-stack Waveform Inversion." 2006 SEG Annual Meeting. Society of Exploration Geophysicists.

[19]. Hartono, and Erianto Ongko, (2015). The Influence of Alpha Value as Multiplier Factor on Arithmetic Crossover in Genetic Algorithm. *International Journal of Computing and Technology (IJCAT),* Volume **2**, Issue 7.

[20]. Ferreira, C., (2006). Gene Expression Programming (Mathematical Modeling by an Artificial Intelligence). Springer-Verlag, Berlin Heidelberg : 55-56.

[21]. Routray, Susmi. (2011). "An Enhanced Genetic Algorithm Approach to ATM Network Design." *Bharati Vidyapeeth's Institute of Computer Applications and Management:* 312.

[22]. Fayyad, Usama, Gregory Piatetsky-Shapiro, and Padhraic Smyth. (1996) "The KDD process for extracting useful knowledge from volumes of data."*Communications of the ACM* **39**(11): 27-34.

[23]. Schumaker, R.P. *et al.,* (2010). Sports Data Mining Methodology, Sports Data Mining, Integrated Series In: Information Systems **26**, Springer Science+Business Media, LLC 2010.

[24]. Freitas, G., A. Laender, and M. Campos. (2002). "MD2-getting users involved in the development of data warehouse application." *Proc. of the 4th Int. Workshop on Design and Management of Data Warehouses*.

[25]. Brohman,K. and Boudreau, M.C., (2004). The Dance: Getting Managers and Miners on the floor together, ASAC, QUEBEC.

[26]. Berry, M.J.A., and Gordon Linoff, (2000) Mastering Data Mining, The Art and Science of Customer Relationship Management, John Wiley & Sons, New York, ISBN 0-471-33123-6.

[27]. Palace, Bill. (1996). "Data mining: What is data mining." online document.

[28]. Verhees, J., (2002). Enhance your Application – Simple Integration of Advanced Data Mining Functions.

[29]. Mailvaganam, H., (2007). Introduction to OLAP - Slice, Dice and Drill!", Data Warehousing Review.

[30]. Shabia. S.K, and Peer. M.A., (2012). "Expedition for the exploration of Apposite Knowledge." *IJCSIT-2012 (IJCSIT) International Journal of Computer Science and Information Technologies* **3**(5): 5164-5168.

[31]. Quinlan, J. R., (1996). Bagging, boosting, and C4.5. In Proc. Thirteenth National Conference on Artificial Intelligence, p. 725–730. AAAI Press.

[32]. Roiger, R. and Geatz, M., (2003). Data mining: a tutorial- based primer, Boston, Massachusetts, Addison Wesley.

[33]. Gupta, Shelly, Dharminder Kumar, and Anand Sharma. (2011). "Data mining classification techniques applied for breast cancer diagnosis and prognosis."*Indian Journal of Computer Science and Engineering (IJCSE)* **2**(2):188-195.

[34]. Fogel, David B. (2006). Evolutionary computation: toward a new philosophy of machine intelligence. Vol. **1**. John Wiley & Sons.

[35]. Rudas, Imre J. and János Fodor., (2008). Intelligent Systems, *Int. J. of Computers, Communications & Control*, Vol. **III**, Suppl. issue: *Proceedings of ICCCC 2008*, pp. 132-138.

[36]. Jang, J.S.R., Sun, C.T., Mizutani, E., (1997). Neuro-fuzzy and Soft Computing, a Computational Approach to Learning and Machine Intelligence, Prentice Hall,: 640 pages.

[37]. Sivanandam, S.N., Deepa, S.N. (2007). "Principles of Soft Computing", Wiley India Edition.

[38]. Shabia S.K, S.M.K. Quadri., (2014). Neuro-adaptive Intelligent Control Technique, *International Journal of Advance Foundation and Research in Science & Engineering (IJAFRSE),* Volume **1**, Issue 6, Pages: 54-62.

[39]. Jalalifar, H., Mojedifar, S., Sahebi, A.A., Nezamabadi-pour, H., (2011). Application of the adaptive neuro-fuzzy inference system for prediction of a rock engineering classification system. *Computers and Geotechnics,* **38**(6): 783-790.

[40]. Tharwat OS. Hanafy, (2010). "A modified algorithm to model highly nonlinear system. "*J Am Sci.* **6**(12): 747-759.

[41]. Sarkheyli, Arezoo, Azlan Mohd Zain, and Safian Sharif. (2015). "Robust optimization of ANFIS based on a new modified GA." *Neurocomputing,* **166**, pp. 357-366.

[42]. Mohammad W. Dewana, Daniel J. Huggetta, T. Warren Liaoa, Muhammad A. Wahaba, Ayman M. Okeilb, (2016). "Prediction of tensile strength of friction stir weld joints with adaptive neuro-fuzzy inference system (ANFIS) and neural network", Elsevier.

[43] Ghosh, Soumadip, *et al*. (2014). "A novel Neuro-fuzzy classification technique for data mining." *Egyptian Informatics Journal*, **15**(3): 129-147.

[44]. Kautz, Henry. (2013). "Data mining social media for public health applications." *23rd International Joint Conference on Artificial Intelligence (IJCAI 2013),* Beijing.

[45]. Kumar, AV Senthil. (2013). "Diagnosis of heart disease using Advanced Fuzzy resolution Mechanism." *International Journal of Science and Applied Information Technology (IJSAIT),* **2**(2): 22-30.

[46]. Austin, Peter C., Jack V. Tu, Jennifer E. Ho, Daniel Levy, and Douglas S. Lee. (2013). "Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes." *Journal of clinical epidemiology*, **66**(4): 398-407.

[47]. Victoria, López, *et al.* (2013). "An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics." *Information Sciences*, **250:** 113-141.

[48]. Jiban, Pal. K. (2011). "Usefulness and applications of data mining in extracting information from different perspectives." *Annals of Library and Information Studies*, **58**(1): 7-16.

[49]. Sohil, Pandya. D., and Paresh V. Virparia. (2013). "Comparing the application of classification and association rule mining techniques of data mining in an Indian university to uncover hidden patterns." *Intelligent Systems and Signal Processing (ISSP), 2013 International Conference on. IEEE.*