International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || Volume 03 - Issue 02 || February 2018 || PP. 26-34

Soft Computing Techniques for Gene Annotation

K. Lohitha Lakshmi[1], P.Bhargavi[2], S.Jyothi[3]
[1]Research Scholar, [2]Assistant professor,, [3]Professor
[1,2,3]Department of Computer Science
Sri Padmavati Mahila Visvavidyalayam, Tirupati-517 502

Abstract: Soft computing is a collection of computational techniques which are new and advanced. These computational techniques which are involved in soft computing reflect human brain in functionality. These techniques are more prominent to the area like bioinformatics in which implementation of mathematical models are difficult. Bioinformatics is a promising research area which includes intra disciplinary fields involving biology, computer science, mathematics and statistics to analyze biological sequential data which includes operations like gene identification, gene structure prediction and gene structure analysis. Soft computing techniques are prominent to solve such type of real world problems included in science and engineering.

Keywords: Bioinformatics (BI), Evolutionary Computing (EC), Fuzzy Logic (FL), Gene Annotation (GA) , Neural Network (NN),Probabilistic Reasoning (PR) , Soft Computing (SC).

1. Introduction

Soft computing refers to a collection of computational techniques which can be implemented in diverse areas like computer science, artificial intelligence and machine learning. Bio medical applications are included in Bio informatics. Soft computing techniques included in bio medical applications are aid in Gene identification and analysis which experiments are helpful in Gene Annotation (GA).

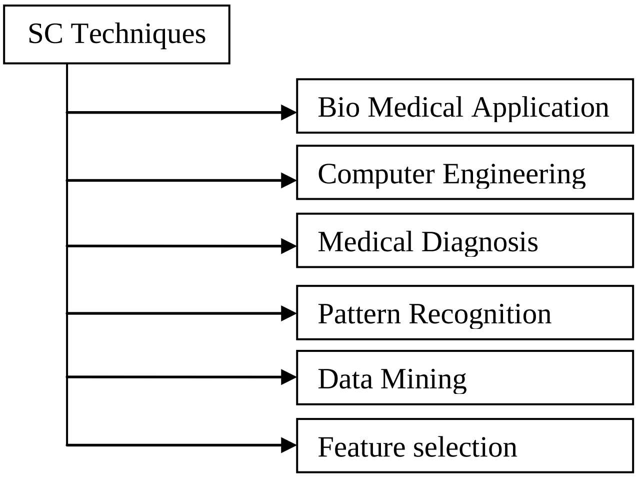

Fig.1 Soft computing techniques applied in different fields.

Gene Annotation is identified as a rising field for experimentation in gene identification, drug discovery and genetic disease prevention etc. The operations related to gene annotation includes vast amount of data. This leads to the combination of soft computing with bioinformatics. Soft computing techniques are applied in different fields related to Genetic data analysis. Soft computing techniques are applied on large data sets which include methods related to AI, ML and statistics. Feature selection is an important application area related to Bio informatics. [2].

2. Soft Computing Techniques

In this paper detailed classification of soft computing techniques has been discussed. Soft computing refers to a collection of computational techniques which can be implemented in diverse areas like computer science, artificial intelligence, machine learning. Technically soft computing comprised in to four disciplines based on knowledge and data driven approaches to find optimal solution. The Probabilistic Reasoning (PR) and Fuzzy Logic (FL) are used in knowledge driven reasoning and the two other Neural Network (NN) and Evolutionary Computing (EC) disciplines are subjected to data driven optimization search techniques. Nature involves large part of science. Soft computing techniques are inspired by nature and these are help ful in transforming knowledge from natural systems to computational system [3].In knowledge driven approach input is knowledge which is taken from experience. This approach emphasizes logic as a modelling tool. Frame works proposed from this point of view are probabilistic models and Fuzzy Logic [4]. In data driven search main source of input is data base i.e. large amount of data repositories.

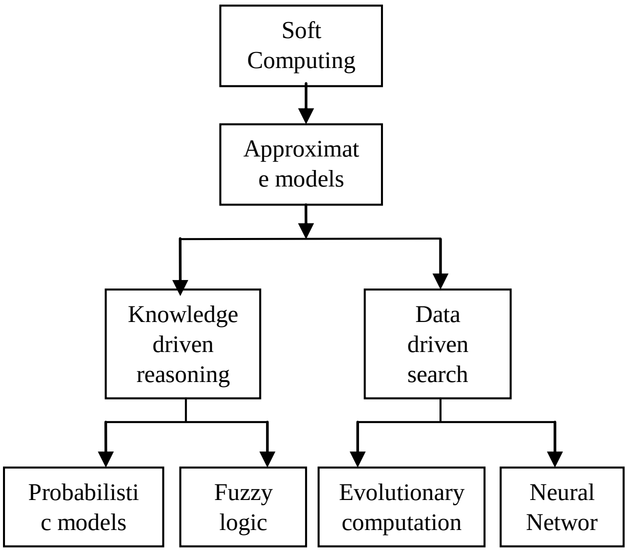

Fig. 2 Classification of soft computing techniques

2.1 Probabilistic Models

Probabilistic models are developed to evaluate the system which includes randomized and probabilistic data. Real world problems are ill defined due to the dynamic nature of variables in different instances. For such type of problems it is difficult to define precise model. Classification of probabilistic models are shown in the below diagram.

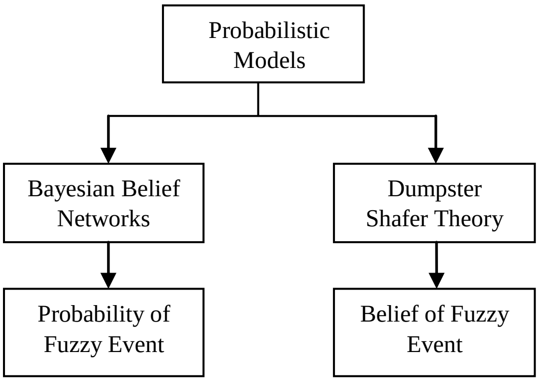

Fig.3 Classification of probabilistic models

2.1.1 Bayesian Belief Networks

Bayesian calculus is meant for conditional probability. Probabilistic reasoning maintains consistency [5].The value of certainty lies between 0 and 1 represent degree of belief. This conditional probability is linked to joint and conditional probability through Bayes theorem which represents Bayesian interpretation of probability. This can be represented as

$$: \quad \boldsymbol{P}\left(\frac{\boldsymbol{B}}{\boldsymbol{A}}\right) = \frac{\boldsymbol{P}\left(\frac{\boldsymbol{A}}{\boldsymbol{B}}\right)\boldsymbol{P}(\boldsymbol{B})}{\boldsymbol{P}(\boldsymbol{A})} = \frac{\mathbf{P}(\mathbf{B},\mathbf{A})}{\mathbf{P}(\mathbf{A})} \qquad (1)$$

P (A/B) also represents the conditional probability of event A with respect to event B [5]. A network model which represents uncertainty in knowledge (belief) is known as belief network. Network reflects the structure of relationships and probability measures the strength of those relationships. Bayesian belief networks provide an effective frame work to work with previous knowledge which can provide better scope to improve future performance.

By considering the problems in Bayesian networks and fuzzy logic a new inference mechanism is proposed called as Fuzzy Bayesian Networks. In this proposed mechanism probabilistic relations among fuzzy variables are modelled into a Bayesian network in which each fuzzy variable is represented as a node. By this phenomenon we can reduce the complexity of structure compared to ordinary Bayesian Networks [5].The starting point to probability of fuzzy sets can be defined in the following way when x is a continuous variable with defined by : $\mathbf{P}(\overline{\mathbf{X_i}}) = \int_{\mathbf{x}} \boldsymbol{\mu}_{\mathbf{i}}(\mathbf{x})\mathbf{P}(\mathbf{x})\boldsymbol{d_x}$ (2)

It is a derivation from Bayesian probability. $\bar{X_i}$ is a marginal probability.
The Bayesian rule can be generalized to fuzzy sets as [6]

$$\mathbf{P}\ (\bar{\boldsymbol{X}}/\bar{\boldsymbol{Y}}) = \frac{\boldsymbol{P}(\bar{\boldsymbol{X}},\bar{\boldsymbol{Y}})}{\boldsymbol{P}(\bar{\boldsymbol{Y}})} \qquad (3)$$

By this we can conclude the probability of fuzzy event in Bayesian network as “By fuzzyfying all fuzzy variables efficiency of data compression can be achieved in a given Bayesian network”.

2.1.2 Dempster Shafer Theory

Dempster Shafer theory also referred as “The theory of belief functions “. It’s a general frame work which works with uncertainty. Later it is termed as “Theory of Evidence” which combines evidence from all available sources of input and leads to get a state of degree of belief. DST is based on beliefs where as Bayesian belief network depends on degree of probability [7].DST provides clear distinction between epistemic uncertainty and physical uncertainty [8]. The original purpose of DST is to compute the degree of belief of statements from a subjective probability of different sources. The subset of space of prepositions ‘□’ lies between [0, 1] scale. The total uncertainty over the space is ‘1’.

Zadeh was the first to generalize the DS theory based on fuzzy event based on the concept of information granularity and possibility distribution. Then he defined the expected certainty, denoted by EC (B), and the expected *probability*, denoted by E∏ (B), as a generalization of D-S belief and fuzzy event [7].

$$\mathbf{E\pi\ (B)}=\textstyle\sum_i \boldsymbol{m}(\boldsymbol{A_i}\mathbf{sup}(\mathbf{B}\square\boldsymbol{A_i}\ \mathbf{)} \qquad (4)$$

$$\mathbf{EC(B)}=\textstyle\sum_i \boldsymbol{m}(\boldsymbol{A_i})\boldsymbol{inf}(\boldsymbol{A_i}\mathbf{=>\ B) = 1\text{–}\ E}\boldsymbol{\pi}(\boldsymbol{B^c}\ \mathbf{)} \qquad (5)$$

Where $\boldsymbol{A_i}$ denotes local fuzzy element induced from fuzzy event, **sup (B□$\boldsymbol{A_i}$)** measures that degree B intersects with A_i, *inf* (A_i=> B) measures the degree A_i is included in B. Form this it is generalized that in D-S theory to deal with fuzzy sets the belief and possibility functions are treated as lower and upper possibilities [9].

2.2 Fuzzy Logic

Fuzzy is a logical system represents to deal with inaccurate and imprecise data which has no particular boundaries. Fuzzy Logic is an approach to computer science which mimics the thinking way of human brain to solve real world problems. Fuzzy logic provides approximate reasoning which involves deductive inferences as well as interpolative inferences. Fuzzy system is designed to reduce the complexity [10].

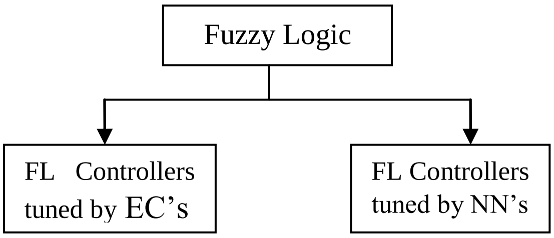

Fig.4 Fuzzy Logic Classification

2.2.1 Fuzzy Logic Controllers Generated and Tuned by EC’s

Several applications have been developed by combining fuzzy logic and evolutionary computation. Evolutionary computing techniques that are evolutionary algorithms uses fuzzy logic based techniques or fuzzy tools to improve the GA behaviour [11][12].

Fuzzy logic control systems like “fuzzy stop criteria” “representation tasks” etc are developed to control the systems performance by taking the consideration of set of parameters such as scaling factors of each variable and taking the fuzzy set which represents the meaning of linguistic values and IF_THEN rules[13]. These set of parameters are generated for monitoring each control system in different evolutionary computing systems.

2.2.2 Fuzzy Logic Controllers Tuned by Neural Networks

NN's are capable of learning, adaptation or parallel computation. These NN's lack flexibility, human interaction, interpretability or knowledge representation. These all capabilities are inversely followed by fuzzy logic. Fuzzy frame work is different [14]. Sufficient knowledge is required for the accurate tuning of fuzzy logic controller. NN is an effective way to implement tuning fuzzy logic controllers.FLC is one of the most intelligent controllers and reflects the human behaviour in process control.FLC provides better performance than conventional controllers. Neural network is one among the best tuning tool for FLC. FLC system comprises of mainly three units 1.Fuzzification unit-in which crisp values are converted to linguistic variables.
2. Fuzzy inference unit- heart of the FLC in which decision making process undergoes.
3. Defuzzification unit-obtained results are converted into crisp values.
FLC using NN as a tuning tool is known as neuron Fuzzy Controller [15]. A coupled tank system proposed by Saurabh and Susmitha Rajan proved in their work based on the simulation results that NFC is better than FLC.

2.3 Evolutionary Computing

Evolutionary computation is a randomized global search paradigm. The important main stream forms of EC's are: Genetic Algorithms (GA), Genetic Programming (GP), Evolutionary Programming (EP), and Evolutionary Strategies (ES).GA takes majority of contributions integrating with fuzzy logic related to natural evolution problems [16]. GA is being used instead if EC. Sometimes GP, ES or EP also referred [17].

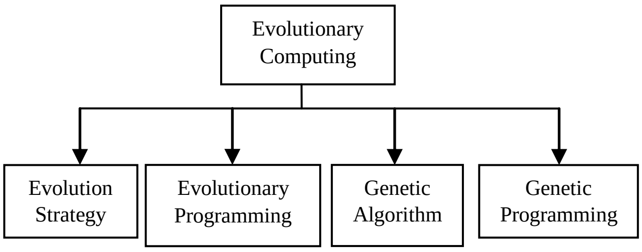

Fig .5 Evolutionary computing classifications

Evolutionary programs (EP's) generating finite state automata that describe strategies or behaviours.GA, ES, EP are similar in functionality in solving a problem [18]. Each are applied trail population and impose random changes to those population and incorporate selection process to determine which solution is more adaptable for future generation experiments. As stated by Fogel GA's [19] emphasis on different models of genetic operators which are observed in nature,

2.4 Neural Network's

Based on the network structure feed forward networks are divided into Single layer feed forward network, Multi layer feed-forward network. In single layer feed forward network, the inputs are directly connected to the outputs which consists of a series of weights. The synaptic links carrying weights directly connects from every input to outputs in forward single way direction. The sum of the products of the weights of the each neuron in the network is calculated, and if the value is above some pre set threshold value the neuron fires and adopts activated value or adopts deactivated value [20].

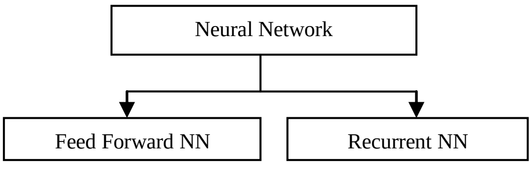

Fig.6 Neural Network Classification

Multi layer feed-forward network consists of multiple layers. Besides having input and output layers also have one or more intermediate layers called hidden layers. The relation of connectivity can be mentioned as (IL->HL1->HL2->OL).Recurrent network differs from other networks is that in this network output of at least one neuron is fed back into itself as input [21].

3. Gene Annotation:

SC techniques are able to deal with incomplete and noisy data in which conventional computing techniques are not. Huge volume of data exists in genetic related experiments. SC techniques are designed to deal with such type of long and complex data sets [22].Gene structure can be analyzed by extracting the data related to protein coding region, RNA coding region and regions of DNA in between genes that are antigenic regions. Gene structure can be stated as

$$(x_1,x_2\text{----------}x_n) \in \Sigma^x$$

$$\text{Where } \Sigma = \{A, C, G, T\} \quad (6)$$

x= sequence of DNA.

x_1,x_2-----x_n are individual elements of DNA[23].

Each SC techniques need two important aspects in every gene annotation program. One is the type of information used by the program and the other one is the appropriate algorithm employed to combine the information into consistent analysis technique.

Gene Annotation (GA) is a process which includes identification of coding regions in a genome and analysis of structure which defines the function of a gene. The basic unit of GA is description of individual gene and its protein product and the function of that particular gene.

4. Soft Computing Techniques in Gene Annotation

There are number of reasons why SC techniques are used in gene annotation.

1. SC techniques are used to extract knowledge from the experience by extracting the hidden facts.
2. SC techniques deal with the inputs which are unpredictable in biological experiments due to the dynamic nature and complexity involved in life sciences.
3. SC techniques are capable to adapt the changes that are generated day to day in biological research.

Soft computing techniques included in bio medical applications are aid in Gene Annotation which experiments are helpful in genetic disease analysis, feature selection and feature prediction.Every experiment related to Genetics includes vast amount of data to be managed. Feature selection is an important application area related to Bio informatics. In Machine Learning and statics variable selection and attribute selection is the process of selecting relevant features for a model [25][26].

4.1. Use of Bayesian Networks in Gene Annotation

Bayesian networks provide a comprehensive modular framework for representing complex systems. The algorithms available to learn structure of BN mainly classified into constraint based and score based methods. Constraint based methods mainly focus on identifying conditional independence relations among variables in the observed data [27].Score based methods consider a possible BN structures and assign a score to each measures in the observed data. Several studies have been proposed related to scientific issue of gene product functional annotations. King OD and Foulger RE propound that “decision trees and Bayesian networks are used to predict novel gene annotations by learning patterns from available annotation profiles”[28]. Tao and colleagues [29] proposed advancement to this work by using a k-nearest neighbour (k-NN) classifier, through which a gene inherits the annotations that are common among its nearest neighbour genes in a gene network. Belief networks use conditioning operation to update the probability values and many approximate methods have been proposed to perform inferences using belief functions [30].

4.2. Dempster Shafer Theory in Gene Annotation

A key requirement of all Bayesian techniques is that probability distributions must be specified in prior. DST overcomes the Bayesian constraint by assigning mass values may not satisfy the classical probability axioms. DS theory applied on these rule sets and gets optional rule set and afterwards applies fusion on these rule sets and generators confine classifier rule set as resultant which can be used for further implementations [31].DSTE has been facing challenges on particular aspects to overcome such type of problems. Some hybrid methods are proposed to get some more clarity [32].

4.3. Use of Fuzzy Logic in Gene Annotation

Historical data provides a view about future up to some extent. Biological data related to gene is intrinsically vast and Fuzzy. Fuzzy logic and fizzy set theory is ideal frame works for defining biological systems [33].

Many algorithms have been proposed to prove that Fuzzy set theory is a powerful computational tool for Gene structure prediction. Application of Fuzzy k- nearest neighbour algorithm is one of the best examples to find protein secondary structure prediction.

A gene annotation can be described by mainly two attributes. These are sequence of a nucleotide or amino acids and gene ontology [34]. Gene ontology is used to classify gene properties that are functions, cell localization and process.

4.3.1 Fuzzy Logic Controllers Tuned by EC's And NN's in Gene Annotation

Fuzzy logic controllers(FLC) is a knowledge base system with linguistic control rules and fuzzification interface which transforms the crisp date into fuzzy sets [35].A fuzzy control system follows if-then conditional statement. Fuzzy logic control systems, when applied to evolution programming techniques can solve many hard optimization problems. The parallel nature of evolutionary programming with the real world application leads to the rise of FLC techniques [36].These techniques are more effect in solving the travelling sales man problem and problems of scheduling, partitioning and control.

4.3.2 Role of Evolutionary Computation in Gene Annotation

Evolutionary computation algorithms consider each individual from each generation to identify the changes of each population. And from the generated individuals some population will be sort out for next generation. This way evolution of off springs will be carried out to next generations in gene annotation processes [37][38]. The aim of EAs implantation is to maximize the fitness function value of each proceeding generations of the population [39].

Many algorithms have been proposed some of them are distribution algorithm, differential evolution and some evolutionary behavioural algorithms [40].Among EC algorithms GA is mainly considered for evolution of biological sequences [41][42]. The total characteristic of each individual gene determines the properties of population. The best chromosome sequence will be obtained to pre determine the characteristics of a particular gene [43][44]. The main contributions this field are estimation of distribution algorithms and Differential Evolution. Evolutionary behavioural algorithms like Particle swarm optimization and Ant colony optimization algorithms also population based [45]. But according to some analogies these algorithms mainly concerned with the social behaviour of individuals rather than individuals biological behaviour. In this way EA s have great impact in Gene annotation [46].

4.4 Neural Networks in Gene Annotation

Genome annotation is a multistep complex [47].The steps involved in genome annotation is

- Nucleotide level which involves gene prediction or identification.
- Protein level which involves structure determination of proteins.
- Process level annotation which involves mechanisms of bio chemical reactions.

Among these three categories nucleotide level plays an essential role in molecular biology. Gene prediction and gene identification involves dynamic programming, decision trees and Hidden Markov Model (HMM) which involves mathematical approach to apply conventional techniques. Neural Networks plays a vital role in gene identification from large DNA sequence [48].

Because of the incomplete or inaccurate vast knowledge GRAIL and GRAIL2 uses NN tools in gene identification [49]. Homogenous Gene Mark (HGM) used for protein coding DNA and Homogenous Markov Model[50] used for non -coding DNA finding strategies[51]". Genome scan use for integrated approaches to find database similarities, MORGAN uses decision trees and dynamic programming. MORGAN is a tool to identify genes in the vertebrate's DNA sequences. GENSCAN and UNVIEL using HMM, GHMM and SPLICEVIEW, SPLICEPREDICTION, AAT and DAGGER like so many tools have been proposed to find regulatory sequences of gene and decoding of junk DNAs. GLIMMER-M is a well-known tool developed for gene finder specifically for eukaryotes .SVM (Support Vector Machine) is a classification tool which comes under supervised learning methods in NNs [52].A multi classifier system has been proposed based on neural network for gene identification.

Soft computing and Computational biology is an interdisciplinary program that brings together the domain specific knowledge of science and engineering with relevant areas of computing and bioinformatics. Soft computing has the potential to revolutionise healthcare, and respond to the increasing volume and complexity in biomedical and bioinformatics data [53].

5. Conclusion

In the above paper detailed classification of computational techniques in soft computing has been discussed. The heuristic models which are already proposed can't provide mathematical implementation to real world application problems of gene annotation. Soft computing techniques can provide mathematical implementation models for uncertain, imprecise and incomplete data. In my further research work these computational techniques will be implemented for gene sequence analysis which is the part of gene annotation.

References

[1]. Ian Dunham 1, Anshul Kundaje, "An Integrated Encyclopedia Of Dna Elements In The Human Genome",Hhs Public Access, . 2012 September 6;489(7414): 57–74. Doi:10.1038/Nature11247.

[2]. D.K. Chaturvedi, "Chapter 2 - Soft Computing Techniques And Their Applications", Springer Science Business Media Singapore 2015 A.H. Siddiqi Et Al. (Eds.), Mathematical Models, Methods And Applications, Industrial And Applied Mathematics, Doi 10.1007/978-981-287-973-8_2.

[3]. Nature Inspired Computation Techniques and Its Applications in Soft Computing: Survey K. Himabindu1 , S. Jyothi2 1, 2Department of Computer Science, Sri Padmavati Mahila Visvavidyalayam (Women's University), Tirupati, INDIA, International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887 Volume 5 Issue VII, July 2017- Available at www.ijraset.com

[4]. Nino Spataro1, Juan Antonio RodrIguez1, Arcadi Navarro1,2,3,4,† And Elena Bosch1," Properties Of Human Disease Genes And The Role Of Genes Linked To Mendelian Disorders In Complex Disease Aetiology", Human Molecular Genetics, 2017, Vol. 26, No. 3 489–500 Doi: 10.1093/Hmg/Ddw405 Advance Access Publication Date: 4 January 2017 Original Article.

[5]. Heping Pan, Daniel McMichael," Fuzzy Causal Probabilistic Networks - A New Ideal and Practical Inference Engine", May 22, 1998.

[6]. przemyslaw Grzegorzewski, "Conditional probability and independence of Instuitoinistic fuzzy events", Systems research institute, polish academy of sciences,ul. Newelska 6,01-447 warsaw, Poland, and F.Math. Inf. Sci., warsaw Univ. of Technology, Plac Politeechniki 1, 00-661 warsaw, Poland.

[7]. O.Cord On, F.Herreraand M.Lozano," On Thecombination of Fuzzylogic and evolutionary computation: Ashortreviewand Bibliography".

[8]. John Yen, "Generalizing The Dempster-Shafer Theory To Fuzzy Sets", This Article Is Partially Based On The Author's Ph.D. Thesis A.T The University Of California., Berkeley, Which Was Supported By National Science Foundation Grant Dcr-8513139.

[9]. Philippe Smets, "The degree of belief in a fuzzy event ", Information Sciences, Volume 25, Issue 1, October 1981, Pages 1-19.

[10]. P.Limbourg,"Multi-objective Optimization of Problems with Epistemic Uncertainty" proceedings of Evolutionary MultiCriterion Optimization, Third International Conference, EMO 2005, Guanajuato, Mexico, March 9-11,pp.413-427,2005.

[11]. F. Hoffmann," Evolutionary algorithms for fuzzy control system design ",Proceedings of the IEEE (Volume: 89, Issue: 9, Sep 2001).

[12]. PintuChandraShill , BishnuSarker and Kazuyuki Murase ,"Evolutionary Tuning Of Fuzzy Rule Base Systems For Nonlinear System Modelling And Control", International Journal of Computer Science & Information Technology (IJCSIT) Vol 4, No 5, October 2012.

[13]. nlm.nih.gov/pmc/articles/PMC2831005/#B34.

[14]. D. van Cleave ; K.S. Rattan," Tuning of fuzzy logic controller using neural network", Published in: National Aerospace and Electronics Conference, 2000. NAECON 2000. Proceedings of the IEEE 2000.

[15]. Susmitha Rajana, Saurabh Sahadevb," Performance Improvement of Fuzzy Logic Controller using Neural Network", International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015), Procedia Technology 24 (2016) 704 – 714.

[16]. BiaobiaoZhang, YueWu,JiabinLu, and K.L.Du, "Evolutionary Computation and Its in Neural and Fuzzy Systems", Accepted 4 August 2011,Miin-Shen Yang.

[17]. Ajith Abraham, Nadia Nedjah and Luizade Macedo, Mourelle, "Evolutionary Computation: from Genetic Algorithms to Genetic Programming", e (SCI) 13, 1–20 (2006) Springer-Verlag Berlin Heidelberg 2006 Ajith Abraham et al.

[18]. David B. Fogel Lawrence J. Fogel, "An introduction to evolutionary programming", European Conference on Artificial Evolution AE 1995: Artificial Evolution ,pp 21-33, book series (LNCS, volume 1063).

[19]. P.Limbourg,"Multi-objective Optimization of Problems with Epistemic Uncertainty" proceedings of Evolutionary MultiCriterion Optimization, Third International Conference, EMO 2005, Guanajuato, Mexico, March 9-11,pp.413-427,2005.

[20]. Lotfi A. ZadehUniv. of California, Berkeley, "Fuzzy logic, neural networks, and soft computing", Magazine Communications of the ACM CACM Homepage archive Volume 37 Issue3,March1994 ,Pages 77-84 .

[21]. Susmitha Rajana, Saurabh Sahadevb," Performance Improvement of Fuzzy Logic Controller using Neural Network", International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015), Procedia Technology 24 (2016) 704 – 714.

[22]. https://www.ncbi.nlm.nih.gov/books/NBK12983/

[23]. NeelamGoel, ShailendraSingh, andTrilokChandAseri, "A Review of Soft Computing Techniques for Gene Prediction", Hindawi Publishing Corporation, ISRN Genomics, Volume2013, ArticleID191206, 8pages, Received26December2012,

[24]. Neelam Goel ⇑ , Shailendra Singh, Trilok Chand Aseri, "A comparative analysis of soft computing techniques for gene prediction", Analytical Biochemistry 438 (2013) 14–21.

[25]. Abiola Oduguwa, Ashutosh Tiwari, Rajkumar Roy and Conrad Besant, "An Overview of Soft Computing Techniques Used in the Drug Discovery Process, Applied Soft Computing Technologies: The Challenge of Complexity", pp 465-480.

[26]. Chengwei Su,1 Angeline Andrew,2 Margaret R Karagas,2 and Mark E Borsuk1 "Using Bayesian networks to discover relations between genes, environment, and disease", BioData Min. 2013; 6: 6.Published online 2013 Mar 21. doi: 10.1186/1756-0381-6-6.

[27]. King OD, Foulger RE, Dwight SS, White JV, Roth FP. Predicting gene function from patterns of annotation. Genome Res. 2013;13(5):896–904. [PMC free article] [PubMed]

[28]. Tao Y, Sam L, Li J, Friedman C, Lussier YA. Information theory applied to the sparse gene ontology annotation network to predict novel genefunction. Bioinformatics. 2007;23(13):i529–i538.doi: 1093/bioinformatics/btm195. [PMC free article] [PubMed] [Cross Ref].

[29]. Yang Weng ; Yunmin Zhu ,"Combining Gene-Finding Programs by Using Dempster-Shafer Theory of Evidence for Gene Prediction", Accession no: 10.1109/ICCIAS.2006.294118 , Publisher: IEEE.

[30]. John Yen," Generalizing the Dempster-Shafer Theory to Fuzzy Sets", This article is partially based on the author's Ph.D. thesis a.t the University of California., Berkeley, which was supported by National Science Foundation Grant DCR-8513139. Part of the research described in the paper was done at USC/Information Sciences Institute and was supported by DARPA under contract No. MDA903-86-C-0178.

[31]. Alessandro Saffiotti, "Using Dempster-Shafer Theory in Knowledge Representation", Appears in Proceedings of the Sixth Conference on uncertainty in Artificial Intelligence (UAI1990).

[32]. Dong Xu, Member, IEEE, Rajkumar Bondugula, Mihail Popescu, Member, IEEE, and James Keller, Fellow, IEEE," Bioinformatics and Fuzzy Logic", supported by the National Library of Medicine Biomedical and Health Informatics Research Training grant 2-T15-LM07089-14.

[33]. Madara Gasparovica, Natalia Novoselova, Ludmila Aleksejeva, "Using Fuzzy Logic to Solve Bioinformatics Tasks", Scientific Journal of Riga Technical University Computer Science. Information Technology and Management Science 2010, Volume 44.

[34]. Susmitha Rajana, Saurabh Sahadevb, " Performance Improvement of Fuzzy Logic Controller using Neural Network", Science Direct 2212-0173 © 2016. Published by Elsevier Ltd. International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015).

[35]. Biaobiao Zhang,Yue Wu, Jiabin Lu, and K.- L. Du," Evolutionary Computation and Its Applications in Neural and Fuzzy Systems", Applied Computational Intelligence and Soft ComputingVolume 2011 (2011),Article ID 938240, 20 pages.

[36]. Franscesco Battaglia, Department of Statistics, Sapienza University, Rome, Italy, "Evolutionary Computation Methods and Their Applications in Statistics" statistica anno LXIX,n.2-3-2009.

[37]. S.K. Pal ; S. Bandyopadhyay ; S.S. Ray, "Evolutionary computation in bioinformatics: a review", IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) (Volume: 36, Issue: 5, Sept. 2006)

[38]. Aki Sorsa, Anssi Koskenniemi and Kauko Leiviskä, "Evolutionary algorithms in nonlinear model identification", university of OULU, Report A No 44, September 2010.

[39]. Alina Sirbu, Heater J Ruskin, Martin Crane "Comparison of evolutionary algorithms in gene regulatory network model inference ", BMC Bioinformatics. 2010; 11: 59.Published online 2010 Jan 27. doi: 10.1186/1471-2105-11-59.

[40]. Sung Hoon Jung, kwang hung cho, "Identification Of Gene Interaction Networks Based On Evolutionary Computation", Korean systems biology grant, MI0309000006-03B5000-00211.

[41]. JohnMcCall," Genetic algorithms for modelling and optimisation", Journal of Computational and Applied Mathematics Volume 184, Issue 1, 1 December 2005, Pages 205-222.

[42]. P. J. Fleming and R. C. Purshouse," Genetic Algorithms In Control Systems Engineering ",Annual Genetic Programming Conference (GP), (1996) – present, proceedings published by the Massachusetts Institute of Technology (MIT) Press. [An annual conference covering the field of genetic programming and now incorporated into the Genetic and Evolutionary Computation Conference (GECCO)].

[43]. Ajith Abraham1, Nadia Nedjah and Luiza de Macedo Mourelle," Evolutionary Computation: from Genetic Algorithms to Genetic Programming", A. Abraham et al.: Evolutionary Computation: from Genetic Algorithms to Genetic Programming, Studies in Computational Intelligence (SCI) 13, 1–20 (2006) www.springerlink.com c Springer-Verlag Berlin Heidelberg 2006.

[44]. Kun-Huang Chen, Kung-Jeng Wang, Min-Lung Tsai,Kung-Min Wang Angelia Melani Adrian,"Gene selection for cancer identification: a decision tree model empowered by particle swarm optimization algorithm", BMC Bioinformatics 201415:49, licensee Bio MedCentralLtd. 2014,Accepted: 7 February 2014,Published: 20 February 2014.

[45]. Robert Ekblom and Jochen B W Wolf," A field guide to whole-genome sequencing, assembly and annotation", Evol Appl. 2014 Nov; 7(9): 1026–1042.Published online 2014 Jun 24. doi: 10.1111/eva.12178.

[46]. S. Lakshmi, A. Shahin ,"A Survey on Gene Prediction Using Neural Network", International Journal of Computer & Organization Trends –Volume3 Issue4 – May 2013.

[47]. Eric E.Snyder and Gary D.Stormo, " Identification of coding regions in genomic DNA sequences: an application of dynamic programming and neural networks", Nucleic Acids Research, 1993, Vol. 21, No. 3 607-613.

[48]. v.bhaskara murthy ,G.Pardha Saradhi Varma, "Neural Networks-a case study in gene identification" S.C. Satapathy, S.K. Udgata, and B.N. Biswal (eds.), FICTA 2013, Advances in Intelligent Systems and Computing 247, DOI: 10.1007/978-3-319-02931-3_13, © Springer International Publishing Switzerland 2014.

[49]. Lorenzo Cerutti, Swiss Institute of Bioinformatics, "Gene finding and gene structure prediction", EMBNet course, March 2003.

[50]. Zhuo Wang, Yazhu Chen,and Yixue Li," A Brief Review of Computational Gene Prediction Methods",GenomicsProteomics Bioinformatics. 2004 Nov; 2(4): 216–221.Published online 2016 Nov 28. doi: 10.1016/S1672-0229(04)02028-5.

[51]. Duc-Hau Le, Nguyen Xuan Hoai, Yung-Keun Kwon, "A Comparative Study of Classification-Based Machine Learning Methods for Novel Disease Gene Prediction", © Springer International Publishing Switzerland 2015

[52]. Singaraju Jyothi, Peyakunta Bhargavi, (2018), Data Science and Computational Biology, in Deep Learning Innovations and Their Convergence With Big Data edited by S. Karthik, Anand Paul, N. Karthikeyan, IGI Global Publications, DOI: 10.4018/978-1-5225-3015-2.ch009.