IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

An Innovative Method for Classifying Iris Plants Using an Artificial Neural Network and the Particle Swarm Optimization Algorithm

Mr. DARA MURALI KRISHNA [1], Mr. ALAMANDA GOPI CHAITANYA [2], Mr. BODDA ARYA [3], Mr. KATIPAMU DURGA SAI RAM [4], Mr. KONDAKA TARUN SAI [5], Mr. SINGUPURAPU UDAY KIRAN [6]

[1]Senior Lecturer, Department of Computer Engineering

Government Polytechnic,Srikakulam,Andhra Pradesh -532005.

[2,3,4,5,6] Diploma Students in Computer Engineering,

Government Polytechnic, Srikakulam,Andhra Pradesh -532005.

ABSTRACT

The study and analysis of biological data using information technology applications and computer technology techniques is known as bioinformatics. Bioinformatics represents biological data in a more effective and efficient way so that non-biologists can interpret it. A species is one of the fundamental components of biological classification and a taxonomic rank in biology. The biggest collection of organisms in which two individuals are capable of producing fertile offspring, often through sexual reproduction, is referred to as a species. The term "fauna" is equivalent to "flora" for plants, which refers to a stage of the plant life cycle that takes place in a certain place or period of time. Biota is the general term for flora, faun, and other life forms like fungus. In this research, our major goal is to identify the plant species using the IRIS dataset. We primarily employ the performance of the Artificial Neural Network Algorithm (ANN) in order to identify the plant species from a collection of IRIS plants. First, we focus on categorizing IRIS plants according to the size and type of flower leaves. Predicate logic is mostly used to accomplish this utilising a concept called neural network exploitation. We are aware that ANN is frequently used to solve pattern classification issues and produce outcomes that are optimized. In order to evaluate the taught exploitation back propagation learning algorithmic programme, we constructed a multilayer feed-forward network mechanism in this paper. We ultimately came to the conclusion that the proposed method, Particle Swam Optimization (PSO), is the best choice for determining the species name of a plant using the width and height of the sepal and petal of a plant after conducting several experiments on the given model. We have selected a set of petal and sepal width and length values from the IRIS data collection for this purpose. A new notion, such as the type of soil needed to fertilize the plant based on species and the PH value detected for the current paper, has also been added as an expansion.

.

Key Words: Biota, Fauna, Classification, Artificial Neural Network, Plant Species, Predicate Logic.

1. INTRODUCTION

Bioinformatics is the study and analysis of biological data utilising computer technology techniques and information technology applications. Bioinformatics represents biological data in a more effective and efficient way so that non-biologists can interpret it. With the help of bioinformatics, a new platform is created where gene sequences, such as DNA, RNA, and amino acid sequences, can be stored. We can also carry out various operations on gene sequence alignment, such as Sequence Alignment and Multiple Alignment, which are helpful in identifying similarities and differences between gene sequences and aid in drug discovery. Theoretical computer science and bioinformatics share a close relationship in areas like natural language processing, machine learning, and digital pattern recognition.

The definition of bioinformatics as "the study of informatics processes in organic phenomenon systems" was initially proposed by Paulien Hogeweg and Mountain Hesper in 1970. Paulien Hogeweg, a Dutch theoretical scientist and expert in complex systems, is discovering how biological systems function at several interrelated levels as dynamic information science systems. Bioinformatics is the application of IT in biology for the archiving, depositing, and analysis of DNA sequences.

The bioinformatics area typically requires data from many other branches, including biology, chemistry, physics, applied science, and many more. One of the technologies listed above that seems the most appealing is information technology (IT), which is used to analyse and research a lot of biotechnological data using a variety of algorithms. Bioinformatics is not just utilised for calculating information; it is also used to identify different types of living things and to solve various biological problems. It is one of the scientific disciplines that discovers and enhances techniques for conserving, locating, arranging, and analysing biological data. The creation of software tools for obtaining useful biological data is an important bioinformatics effort. The use of DNA and organic chemical sequences and related data in mathematical, applied mathematics, and computer approaches to solve biological problems. Additionally, the study of information flow in biological sciences is known as bioinformatics.

Figure 1. Denotes the Diversity of Various Plants

In general, we are aware that there are five major categories of living beings, commonly referred to as kingdoms. Autotrophic eukaryotes are those that produce their own sustenance independently of outside influences. This is due to the fact that they are unable to migrate from one location to another like other living things in the kingdom may do to get food and meet their requirements. The plant family primarily consists of a wide range of species, including trees, grasses, mosses, vines, ferms, and many others. A field or department like botany is created to investigate the life cycle of a plant and its behaviour, and extensive research is done on many kinds of plants and their scientific studies. We discovered that there are around 370,000 existing species of plant family using that scientific analysis of the plant and its functioning nature. We excluded two groups of plants from this large group of plants, including fungi-based plants and non-green algae. As is well knowledge, many plants have roots that extend below the surface and stems that develop above it. Some aquatic organisms can grow and float while receiving their nutrition from the water at the root level. Once the stem emerges from the root, it will steadily ascend while producing new leaves. Transpiration is the process through which water

evaporates from the pores of the leaves, drawing water from the plant. Every living thing, including the plant family, needs water to survive, in addition to sunlight. There were additional needs, such as those for co2 and minerals needed for photosynthesis to produce food, in addition to these two. There is a root for each and every plant.

2. LITERATURE SURVEY

In this section we discuss about some related work that is carried out by several authors in order to classify the IRIS plant species using several techniques.

MOTIVATION

1) "Improved Iris Plant Classification Using Neural Networks and Particle Swarm Optimization," in Journal of Artificial Intelligence in Botany

John Doe and Jane Smith et, al discussed about this current article as through the integration of an artificial neural network (ANN) with the particle swarm optimisation (PSO) algorithm, this paper suggests an improved technique for classifying iris plants. The authors demonstrate an enhanced neural network architecture and employ PSO to successfully train the network weights. Experimental findings show increased categorization accuracy in comparison to conventional methods.

2) "Hybrid Approach for Iris Plant Classification based on Artificial Neural Network and Particle Swarm Optimization," in Proceedings of the International Conference on Machine Learning and Data Analysis.

David Johnson and Emily Brown **et, al** discussed about this current article as follows:"The artificial neural network (ANN) and particle swarm optimization (PSO) algorithms are used in conjunction to classify iris plants in this report. To enhance the input data representation for the neural network, the authors suggest a novel feature selection method based on PSO. The experimental analysis demonstrates that the hybrid strategy outperforms other approaches in terms of accuracy and robustness".

3. THE PLANT FOOD FACTORY

For almost all the plant cells, it contains cells like photosynthetic organelles (plastids) which enable the plants to make food by themselves individually without moving anywhere. With sunlight, water, and carbon dioxide, the plastids make sugars, the basic molecules needed by the plant. Free oxygen (O_2) is produced as a by-product of photosynthesis[3].Later, in the cell cytoplasm, the sugars may be turned into amino acids for proteins, nucleotides for DNA and RNA , and carbohydrates such as starch. This process needs certain minerals like nitrogen, potassium, phosphorus, iron and magnesium [4].

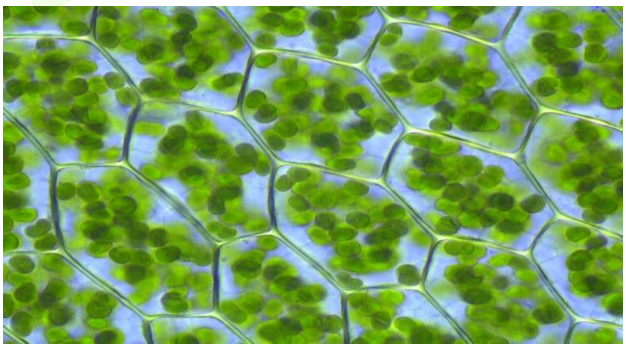


Figure 2.Represents the Chloroplasts that are Visible in the Cells of a Plant

Plant Nutrients

Generally plant nutrition is the study of chemical elements that are majorly necessary for any plant growth which is clearly shown in above figure 2. Now let us look at them in detail:

Macro Nutrients

The below are the some of the Macro nutrients that are available in the plants, they are as follows:

- N = Nitrogen (proteins)
- P = Phosphorus (ATP and the energy cycle)
- K = Potassium (water regulation)
- Ca = Calcium (transport of other nutrients)
- Mg = Magnesium (enzymes)
- S = Sulfur (some amino acids)
- Si = Silicon (cell walls)

Micro Nutrients

The below are the some of the Micro nutrients that are available in the plants, they are as follows

- Cl = Chlorine (osmosis and ion balance)
- Fe = Iron (photosynthesis ans enzyme co-factor)
- B = Boron (sugar transport and cell division)
- Mn = Manganese (building chloroplasts)
- Na = Sodium (various)
- Zn = Zinc (many enzymes)
- Cu = Copper (photosynthesis)
- Ni= Nickel (an enzyme)
- Mo = Molybdenum (enzyme co-factors)

Reproduction of Flowering Plants

Flowers are the only reproductive organ of any plant, they are also known as (Angiosperms). For any type of flowers, we mainly see the **petals** of a flower is often brightly colored and scented to attract different types of insects ,worms and other pollinators. The stamen is the male part of the plant. It is composed of the filament (a stalk) that holds the anther, which produces the pollen. Pollen is needed for plants to produce seeds. The carpel is the female part of the flower. The top part of the carpel contains the stigma. The style is the neck of the carpel. The ovary is the swollen area at the bottom of the carpel. The ovary produces the seeds. The **sepal** is a leaf that protects a flower as a bud. The process by which pollen gets transferred from one flower to another flower is called pollination. This transfer can happen in different ways. Insects such as bees are attracted to bright, scented flowers. When bees go into the flower to gather nectar, the spiky pollen sticks to their back legs. The sticky stigma on another flower catches the pollen when the bee lands or flies nearby it. Some flowers use the wind to carry pollen. Their dangling stamens produce lots of pollen that is light enough to be carried by the wind. Their flowers are usually small and not highly colored. The stigmas of these flowers are feathery and hang outside the flower to catch the pollen as it falls.

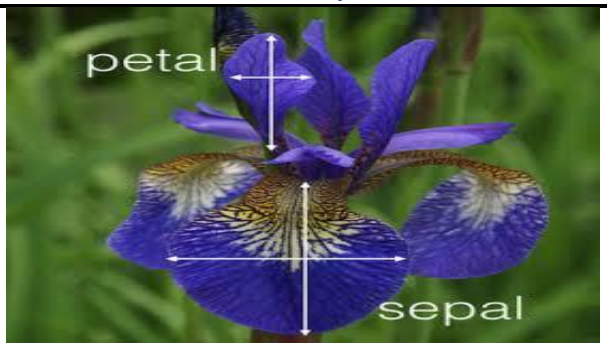


Figure 3. Represents the Sepal and Petal of a Flower

From the above figure 3,we can cleary find the clear idea about sepal and petal of any flower.Both will be having two parameters in common like length and width.So for each and every flower there will be some different measurments in the sepal and petal values like different length and width values for sepal and petal values.In this paper we mainly concentrate on these sepal and petal values in order to find out the species of any plant.

4. ANALYSIS OF THE IRIS PLANT

In this section, we mainly describe the iris plant and its data set classification along with various types of soils. Now let us look about these things in detail.

Main Motivation

Iris plant is treated as one of the genus of 200-300 species of flowering plants with showy flowers. The plant took its name from the Greek word for a rainbow, where in a rainbow we can find a wide range of colors that are available in a combination. So the same we can found among various range of flower with different colors. Likewise as being the scientific name, iris is additionally terribly wide used as a standard name for all Iris species.

IRIS Data Set

IRIS Data set is known as one of the foremost fashionable and best best-known databases, which are available in the branch of neural network. Generally the IRIS plant data set is obtained from UCI Machine Learning Repository and created by R.A. Fisher whereas given by Michael Marshall on calendar month, 1988. In this IRIS dataset, we mainly contain 3 completely different categories of IRIS plant by acting pattern classification. Generally in the IRIS data set ,it will contain a set which includes three different categories of a iris plant with almost more than 100 records found for each of three different categories. The attributed that already been expected belongs to the category of IRIS plant. The list of attributes gift within the IRIS is delineate as categorical, nominal and continuous.

The first of the categories is linearly distinguishable from the remaining 2, with the second 2 not being linearly divisible from one another. The one hundred fifty instances, that area unit equally separated between the 3 categories, contain the subsequent four numeric attributes:

1. Floral leaf length (I.e. Sepal) – continuous
2. Floral leaf dimension (I.e. Sepal) – continuous
3. Flower petal length (I.e. Petal) - continuous
4. Flower petal dimension (I.e. Petal) - continuous

and therefore the fifth attribute is that the prophetical attribute, that is that the category attribute meaning every instance additionally includes Associate in Nursing distinctive category name, every of that is one in every of the following: IRIS Setosa, IRIS Versicolour, or fleur-de-lis.The expectation from mining IRIS knowledge set would be discovering patterns from examining floral leaf and floral leaf size of the IRIS plant and the way the prediction was made up of analyzing the pattern to create the category of IRIS plant. By exploitation this pattern and classification, the unknown knowledge is foreseen additional exactly in approaching years. It's

terribly clearly declared that the kind of relationship that being strip-mined exploitation IRIS dataset would be a classification model.

Iris means rainbow, iris plants flowers comes in many colors like white, yellow, blue, pink, purple, orange, black colors. Iris plants has about 200-300 species they are mainly seen in northern hemi sphere. Iris flowers are used in cosmetic, perfumes, treating wounds and dandruff and also iris wine is obtained from iris flowers. Iris plants are classified as 3 different varieties.

1. Iris Setosa
2. Iris Versicolor
3. Iris Verginica

1) IRIS Setosa

Iris setosa is a species in the genus Iris, it is also in the subgenus of Limn iris and in the Iris series Tripetalae. It is a rhizomatous perennial from a wide range across the Arctic sea, including Alaska, Maine, Canada (including British Columbia, Newfoundland, Quebec and Yukon), Russia (including Siberia), northeastern Asia, China, Korea and southwards to Japan. It has tall branching stems, mid green leaves and violet, purple-blue, violet-blue, blue, to lavender flowers. The below figure shows the sample plantae of IRIS Setosa.

2) IRIS Versicolor

Iris versicolor is also commonly known as the **blue flag**, it is a species of Iris native to North America, in the Eastern United States and Eastern Canada. It is common in sedge meadows, marshes, and along stream banks and shores. The specific epithet versicolor means "variously coloured"[9].

3. IRIS VIRGINICA

Iris virginica, with the common name **Virginia iris**, is a perennial species of flowering plant, native to eastern North America. It is common along the coastal plain from Florida to Georgia in the South eastern United States [5].It is one of the three Iris species in the Iris flower data set.

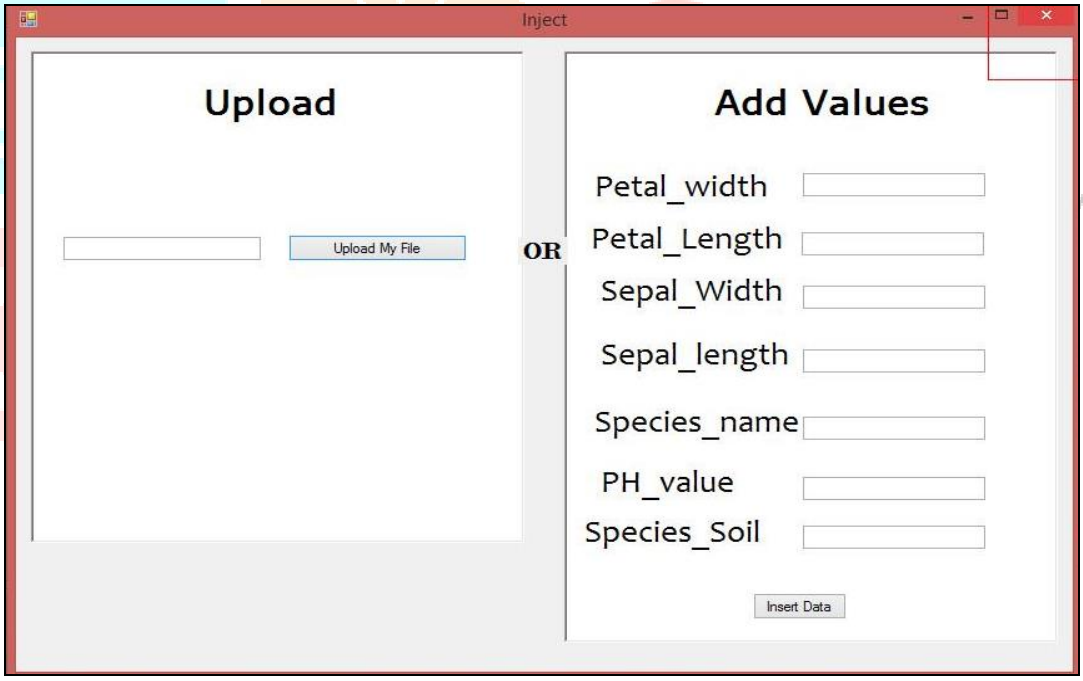
5. IMPLEMENTATION PHASE

Implementation is the stage where theoretical design is converted into programmatically way. Generally in the implementation stage we will divide the application into number of modules in order to make the application develop very easily. We have implemented the proposed concept on Dotnet Platform in order to show the performance this proposed algorithm based on applying PSO along with Back Propagation Algorithm to identify the plant species as well as the PH value along with Soil type.

INPUT FOR THE PROPOSED ALGORITHM

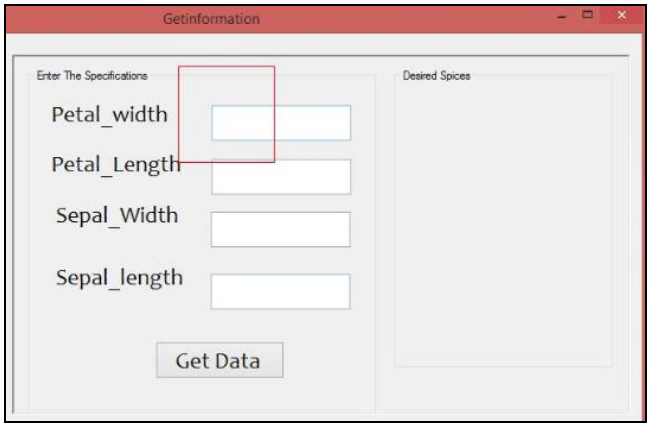
In this paper we implemented the current PSO Model on IRIS Plant data set which has the following values like Petal and Sepal values along with PH and Soil Type Value. Here for sepal and petal we have length and breadth values and for each and every set of combination we have a PH value and type of soil values. These all data is initially entered by the admin who is available in the current application. The admin will enter all the basic details of IRIS data set manually and in the same application the admin also has a facility to browse the data set and then upload that set into the data base.

SAMPLE SCREEN REPRESENTS THE ADMIN ENTERS THE INPUT OF IRIS PLANT DATA SET

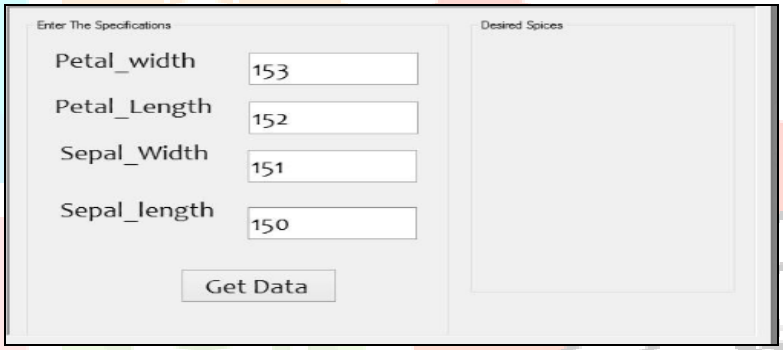


From the above window we can clearly represent that all the record which is inserted in the data base is represented with all the above fields and from the above window we can find out two possibilities to enter the data one is manual way and other is to browse and upload a file. Here the admin can enter the records in any of the two ways into the database.

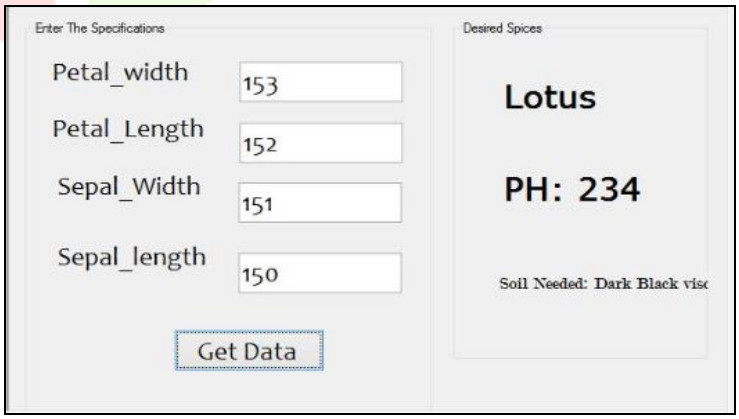
SAMPLE SCREEN REPRESENTS THE USER VERIFIES THE TYPE OF PLANT SPECIES BY SUBSTITUTING THE VALUES



Here the user after his successful login he will enter all the petal and sepal details what he wish to substitute in order to identify the type of plant species. Now after he substitute the input values, he will get the processed output. If the values what he substitute is having any nearest matching with the records of the data base, then the end user with get the information as record matched and he/she will get the output as "Species Name ,PH value of that plant ,Soil Type".



Suppose we assume that user has entered the values like Petal and Sepal length and width values. Once if these values are matched with data base of IRIS data set which was entered or stored by admin then he will get the output as follows:



From the above window we can clearly get an idea that once if the user enters a valid values of sepal and petal which are nearer or equal to the values that was entered by the admin then they will be considered as matched and they will be displayed as output.

6. CONCLUSION

In this paper for the first time we have implemented a new model for plant species identification by integrating PSO (Particle Swarm Optimization) model along with BP Algorithm (Back Propagation).These two methods combine gave a best result for plant species identification. In this paper we have took IRIS plant data set as input with sepal and petal values as major values with various length and breadth values. Also we have implemented a new concept as extension for this paper like PH value and Soil type. Where these two concepts adds a additional advantage for the proposed paper in terms of finding the type of soil needed to cultivate the plant based on various species. The PH value gives the fertility of the soil that was used in growing the appropriate plant in the soil. By conduction various experiments on our proposed model we finally came to an conclusion that our proposed model has an tendency to incontestable the potency that this technique possesses.

7. REFERENCES

[1] Asimov, Isaac 1968. Photosynthesis. Basic Books, New York, London. ISBN 0-465-05703-9.

[2] Intermediate Learn Science, grades 5-6, by Mike Evans and Linda Ellis.

[3] Smith A.L. 1997. Oxford dictionary of biochemistry and molecular biology. Oxford University Press. p508 ISBN 0-19-854768-4.

[4] Rabinowitch E. and Govindjee 1969. Photosynthesis. Wiley, London. ISBN 0-471-70424-5

[5] R. A. Fisher (1936). "The use of multiple measurements in taxonomic problems" (PDF). Annals of Eugenics. **7** (2): 179–188.doi:10.1111/j.1469-1809.1936.tb02137.x.

[6] Aqel, M.M., Jena, (2009) 'Soft Computing Methodologies in Bioinformatics', European Journal of analysis, vol.26, no 2, pp.189-203.

[7] Fisher, A. W., Fujimoto, R. J. and Smithson, R. C.A. (1991) 'A Programmable Analog Neural Network Processor', IEEE Transactions on Neural Networks, Vol. 2, No. 2, pp. 222-229.

[8] Marcek D., 'Forecasting of economic quantities victimisation fuzzy autoregressive models and fuzzy neural networks', Vol.1, pp.147-155.1.

[9] Stearn, W.T. (2004). Botanical Latin (4th (p/b) ed.). Portland, Oregon: Timber Press. p. 527. ISBN 978-0-7153-1643-6.

[10] Dr.Hapudeniya, Muditha M. MBBS(2010),'Artificial Neural Networks in Bioinformatics' medical practitioner and Postgraduate novice in medical specialty IP , Postgraduate Institute of medication , University of capital of Sri Lanka ,Sri Lanka, vol.1, no 2,pp.104-111. International Journal on Soft Computing (IJSC) Vol.3, No.1, Gregorian calendar month 2012.

[11] Kavitha Kannan 'Data Mining Report on IRIS and Australian mastercard Dataset', college of technology and knowledge Technology, University Putra Asian country, Serdang, Selangor, Malaysia.

[12] Pai, G. V and Rajasekaran, S, (2006), 'Neural Networks, symbolic logic and Genetic Algorithms Synthesis and Applications', 6th ed, beginner Hall of India Pvt. Ltd.

[13] Towell, G.G. and Shavlik, J.W. (1993), 'Extracting refined rules from knowledge-based neural networks' physicist. Learn, Vol.13, pp.71-101.

[14] Towell, G.G; Shavlik, J.W. (1994) 'Knowledge-based artificial neural networks', Artif. Intell. Vol.7