# Sentence Recognition Using Hopfield Neural Network

**Bipul Pandey[1], Sushil Ranjan[2] , Anupam Shukla[3] ,and Ritu Tiwari[4]**

**[1]Soft Computing and Expert System Laboratory, Indian Institute of Information Technology and Management Gwalior, Gwalior, Madhya Pradesh ,India**

**[2]Soft Computing and Expert System Laboratory, Indian Institute of Information Technology and Management Gwalior, Gwalior, Madhya Pradesh ,India**

**[3]Soft Computing and Expert System Laboratory, Indian Institute of Information Technology and Management Gwalior, Gwalior, Madhya Pradesh ,India**

**[4]Soft Computing and Expert System Laboratory, Indian Institute of Information Technology and Management Gwalior, Gwalior, Madhya Pradesh ,India**

**Abstract**

Communication in natural languages between computational systems and humans is an area that has attracted researchers for long. This type of communication can have wide ramification as such a system could find wide usage in several areas. Web-Browsing via input given as textual commands/sentences in natural languages is one such area. However, the enormous amount of input that could be given in natural languages present a huge challenge for machine learning of systems which are required to recognize sentences having similar meaning but different lexico-grammatical structures. In this paper, we describe how a binary recurring neural network can be used to sufficiently solve this problem for English. The system uses the Hopfield Neural Network to recognize the meaning of text using training files with limited dictionary. Detailed analysis and evaluation show that the system correctly recognizes/classifies approximately 92.2% of the input sentences according to their meaning.

***Keywords:*** *Artificial Neural Network, Expert Systems, Machine learning, Natural Language, Sentence Recognition.*

## 1. Introduction

Communicating with computer-based systems via natural language has long been considered the future of human-computer system interaction. Though the variety of such systems and their different requirements has slowed down the progress in this field, the foremost problem is how to perform knowledge Discovery from natural language based data sources, which specifically proves to be much more difficult and prone to errors. The ability of Natural Languages to create expressions with different arrangements of same or different words and phrases gives rise to large number of lexico-grammatically diverse sentences which may yet mean similar. Hence system discovering knowledge from such sources require such machine learning that takes care of the complex and immensely important task of natural language text processing in an optimally sufficient way. The branch of Machine Learning is concerned with design and development of such procedures that make it possible for systems to cope with diverse data patterns and change themselves accordingly. It is focused to make the systems to automatically adapt themselves to recognize complex patterns and make decisions accordingly.

The basic problem to be solved is to successfully classify and recognize commands in the form of sentences in natural languages (English in our system). As such, our problem could also be seen in view of the much wider and immensely researched field of Pattern Recognition and Classification. A few problems where this approach has been immensely successful are recognition and categorization of sounds, images, texts, features etc. [1].

A Pattern can be described as “opposite of a chaos; it is an entity, vaguely defined, that could be given a name” [2]. A classification system is one that actually maps input vectors to a specific class. Hence, classification is basically the job of learning the procedure that maps the input data [3]. A Pattern Recognition system basically classifies the input patterns. This may be supervised or unsupervised depending on the input data and technique involved. The classes are defined using prior knowledge about the data related domain [4]. The proposed system uses Hopfield Classifiers to classify sentences according to their meaning. A Hopfield Classifier is a special type of Artificial Neural Network which works recursively to settle its output on one of the design points. These Design points are generally those on which the network settles for the initial training vectors. The network undergoes an unsupervised learning procedure.

The System comprises of the data source, Coding module to create patterns of words and sentences according to the specified algorithm and Hopfield Networks for classifying words and sentences. The Coding Module takes input from the data source for training both the Word Recognition Module and the Sentence Recognition Module. The training input are in the form of sentences which are broken into separate words. The separated words are coded into matrix of size (N x M) each, where ‘N’ is the number of alphabets in the word and ‘M’ is total number of English alphabets. These coded patterns of words are fed as training data to the Word Recognition Module. For Sentence Recognition Module, the training data comprises of sentences taken from the data source whose individual words are separated and coded by the Coding Module in the same way. Then, the matrixes for each word per sentence are appended to form the sentence patterns in the form of a column matrix and fed as the training patterns for the Word Recognition Module. The remaining sentences in the data source are then taken as a testing set. These are separated and coded into words similarly by the coding module. The individual word patterns are classified by the Hopfield classifier in the Word Recognition Module to similar word training patterns. The word training patterns are then appended as per their position in the sentence to give the sentence pattern which is classified by the Hopfield classifier in the Sentence Recognition Module.

Several genres of techniques have been successfully used for the problem of Pattern classification. However, Artificial Neural networks have started to gain a foothold and are increasingly attracting researchers to use them to find novel yet efficient methods of Classification. The most popularly used ones are the Feed-Forward Networks like Multi-Layer Perceptrons. Another widely used network is Self-Organizing Map (SOM).

Researchers have been enthusiastically applying the existing techniques of Pattern Recognition and the Artificial Neural Network paradigm to cater the problems of character, and word sentence recognition in textual, image and sound data. In [5], an algorithm had been proposed for the recognition of isolated off-line words. The algorithm is based on segment string matching and could do with moderately noisy and error prone data set. In their paper, Jones et al [6] have talked about how Language Modeling can be used to solve the problem of sentence recognition. They have used probabilistic grammar along with a Hidden Markov Identifier for adequately completing this task. In [7], an algorithm had been proposed to describe a framework for classifier combination in grammar-guided sentence recognition. Hybrid Techniques have also been used for the aforesaid problem. In [8], Hidden Markov Model (HMM) and Neural Network (NN) Model have been combined for the solution. Here, Word Recognition had been using a Tree-Structured dictionary while Sentence Recognition is done using a word-predecessor conditioned beam search algorithm to segment into words and word recognition. In [9], sentence recognition has been achieved which uses a template based pattern recognition and represents words as a series of diphone-like segments. In [10], word co-occurrence probability has been used for sentence recognition. The incurred results were also compared with the method using the Context Free Grammar. Binary Neural Networks have also been successfully used in the task of pattern recognition. Binary Hamming Neural Network has been applied to recognize sentences and have been found to sufficiently successful in this regard. The system proposed also takes advantage of greater speed of the Binary Networks to provide a very efficient solution to the problem of sentence recognition [11]. David and Rajsekaran have talked about how Hopfield classifiers can be used as a tool in Pattern Classification [12]. A combination of Hopfield Neural Network and Back Propagation Approach has also been used to propose method of vehicle license character recognition supported by a study of the relations among the study rate, error precision and nodes of the hidden layer [13].

## 2. Method Adopted

The architecture of the proposed system is shown in Fig 1. The first module comprises of the selected database. This data base contains sentences in English extracted from any general text, for example, for an intelligent web-browsing system, it comprises of sentences/ commands used while browsing the World Wide Web (www) in text format. Hence, the database is domain specific, i.e. relevant to the sentences that will be classified using the system. First the sentences or commands are extracted from the text database.

The architecture of the proposed system is shown in Fig 1. The first module comprises of the selected database. This data base contains sentences in English extracted from any general text, for example, for an intelligent web-browsing system, it comprises of sentences/ commands used while browsing the World Wide Web (www) in text format.

Hence, the database is domain specific, i.e. relevant to the sentences that will be classified using the system. First the sentences or commands are extracted from the text database.

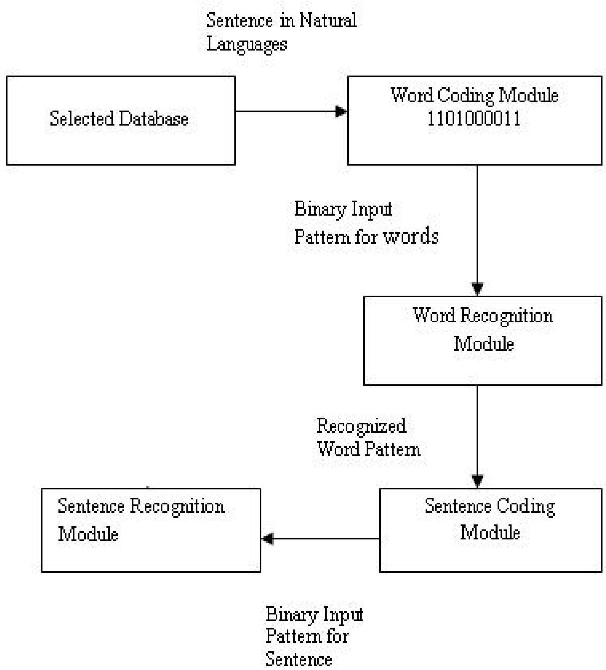


Fig. 1. Architecture of the System

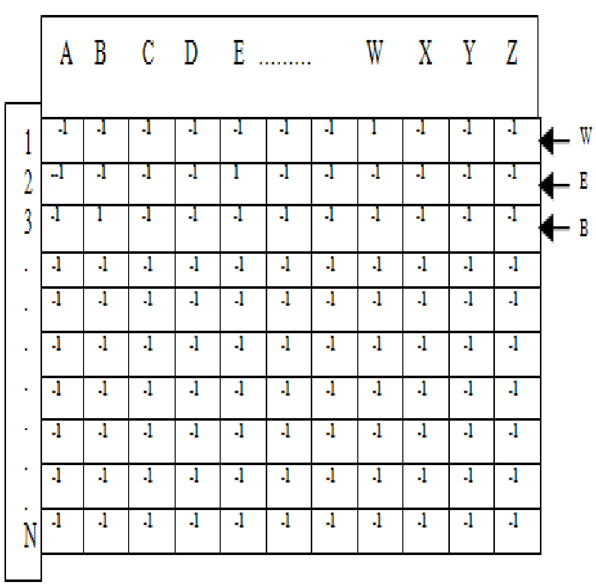


Fig. 2. Format for Binary Coding of Isolated Words

The extracted sentences are then passed on to the Word Isolation and Coding Module. This module separates the different words present in each sentence and codes them in the form of a matrix whose format is shown in the Fig 2.The words are coded in a binary manner (using 1 & -1). This is done because binary inputs work more efficiently in the recurring neural network being used. Also, use of non-binary inputs may lead to undesirable, spurious stable points in the network. The words are coded keeping in mind the relative position of the alphabets in the word, denoted by the row number in which they are present. What alphabet is present is represented by the column number in which they are present. Each column represents an alphabet of English depending such that the alphabet represented by a column has the same position in the English alphabets as the column number of that column like the 3$^{rd}$ column would represent 'c' alphabet of English, where 'c' is at 3$^{rd}$ alphabet in English alphabet. For Example, the word "web" is coded as shown in the Fig 2. The first 'w' is represented by the binary input 1 present in (1, 23) position in the matrix. Rest all positions are filled with the binary input -1 for the first row. Hence, for nth row all positions are -1 except where an alphabet is present

Considered one of the historical Neural Networks, the Hopfield Neural Networks are a type of Bidirectional Associative Memory. The network used here is the discrete version, i.e. the unit outputs are a discrete functions of the inputs.

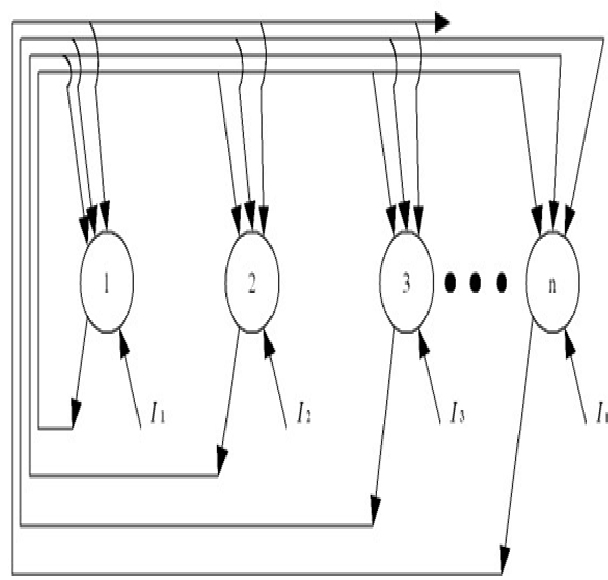


Fig. 3. Architecture for Hopfield Neural Network

Being recurrent means the output of each node is drawn back to all other nodes via weights. The Hopfield network is created by supplying input data vectors, or pattern vectors, corresponding to the different classes. These patterns are called exemplar class patterns. In an *n*-dimensional data space the class patterns should have *n* binary components {1,-1}; that is, each class pattern corresponds to a corner of a cube in an *n*-dimensional space. The network is then used to classify distorted patterns into these classes. When a distorted pattern is presented to the network, then it is associated with another pattern. If the network works properly, this associated pattern is one of the class patterns. The matrix of weight 'W' of a discrete time Hopfield classifier is defined:

$$W = 1/n\left(\sum_{i=1}^{D} \zeta_i^{T} \xi_i\right) \qquad (1)$$

where D is the number of exemplar class patterns $\{\xi_1, \xi_2, ..., \xi_D\}$ vectors consisting of +/-1 elements, to be stored in the network, and n is the number of components, the dimension, of the class pattern vectors.

Discrete-time Hopfield networks have the following dynamics:

$$x(t+1) = sign(W.x(t)) \qquad (2)$$

which is applied to one state x(t) at a time. At each iteration, the state to be updated is chosen randomly. The update continues till the network converges. A distorted pattern, x (0), is used as initial state which is updated till convergence of the network to give the output which is the associated pattern.

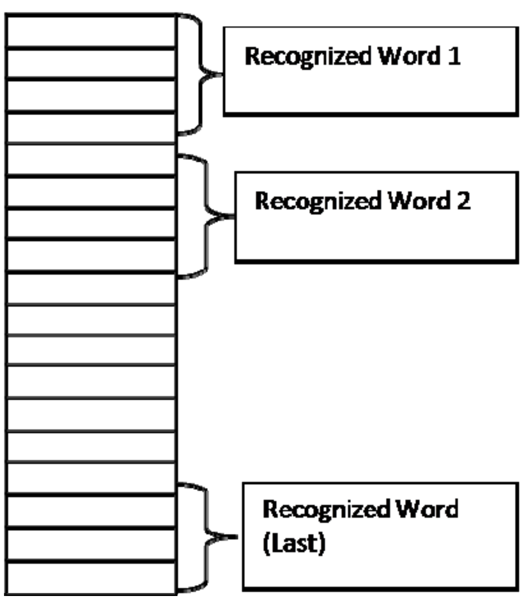


Fig. 4. Format for Binary Coding for Sentences using recognized word pattern in column matrix form.

After this, the coded words are passed on to the word recognition module. This module comprises of a single layer recurring Hopfield Neural Network trained with certain chosen words from the database. The Hopfield Classifier present classifies the input word patterns and gives as its output the training word pattern most similar to that input word. The different words recognized for the input sentence pattern are now combined to give a pattern in the sentence coding module which is the input for the sentence recognition module (Fig 4). This is done by appending the output patterns according to the position of the words in the sentence to give the sentence pattern in the column matrix form. This module also comprises of a single layered recurrent Hopfield Neural network trained with chosen sentence patterns. The network iterates and converges to give output which is the training sentence pattern most similar to input sentence pattern. This output is then matched with the original sentence to check whether the output conveys the same meaning or not.

## 3. Experimental Results

The Database, used for constructing the proposed system, is comprised of 500 sentences, with 50 sentences to be used as training data for the network in the sentence recognition module. The words of the 50 training sentences along with 100 more words were used for training the network in word recognition module. The Database here was prepared by taking 500 sentences from a general text regarding common web usage. These sentences were taken as a paragraph in the form of a text file. The different sentences separated by full stops were segregated using the Word Isolation and Coding Module. This module is built using Java Coding. Each sentence had its end marked by a full stop. The code takes advantage of this and separates sentences which start after one full stop and end before the next one. The module then isolates each word per sentence and forms word pattern for each using the matrix format described and outputs each in the form of a text file.

Table I. Details of Hopfield Classifiers Used

| ***Model*** | ***No. of Layers (word recognition module, sentence recognition module)*** | ***No. of Neurons (word recognition module, sentence recognition module)*** | Transfer Function |
|---|---|---|---|
| Hopfield Classifier | (single, single) | (393,500) | Satlins |

The 50 sentences chosen as the training data for the Hopfield Classifier in the Sentence Recognition Module have their words along with 100 more words taken as the training set for the Word Recognition Module. The details of the Hopfield Classifiers used in each of the module are shown in Table I.

The Sentence pattern of these 50 sentences is formed by appending the word patterns of each word of a sentence as per their position in the sentence to give the sentence pattern of that particular sentence. These patterns are then used as the training set for the Hopfield Classifier of the Sentence Recognition Module. The remaining sentences of the database are chosen as the testing data for the system. The words of each of the testing sentence are fed to the Word Recognition Module which gives the training word pattern most similar to that word as its output. The output patterns of each word are then appended as per its position to give the sentence pattern in the column matrix form as described earlier. The patterns of each of the sentence are then given to the Sentence Recognition Module where the Hopfield Classifier classifies and gives the most similar training sentence pattern as its output. The output of each testing sentence is then checked with the original testing sentence to find out where the output is similar in meaning it or not.

Using the proposed expert system, we could recognize/classify approximately 92.2% of the input sentences according to their meaning.

Table II. Result Obtained

| ***Technique Used*** | ***Total No. of Sentence Pattern In Database*** | ***No. of Training Sentence Patterns*** | ***No. of Training Word Pattern*** | Performance |
|---|---|---|---|---|
| Hopfield | 500 | 50 | 393 | 415/450 |

## 4. Conclusion and Future Works

The results show that the proposed system using binary recurrent Hopfield classifier has been promisingly successful to undertake sentence recognition. Though, these sentences are in different lexical and grammatical format, yet the system has been able to successfully process the text in order to produce the desired results.

The type of model described in the paper could be extended to developing other such exciting expert systems like an Intelligent Medical System which could take text input from patients and gives adequate prescriptions.

One of the main reasons for erroneous output in the proposed system turned out to be the mismatch in the effective part of a word pattern and that of the training words used. The effective part is the part where the binary values turn out to be 1. This means that that output of the word recognition is affected by the difference in the number of alphabets present in the input word and those in the training word patterns. This sometimes grows over to give incorrect sentence mapping. Hence, this is an area where the proposed system could be worked upon.

## References

[1] Paulus, D., and Hornegger, J., Applied Pattern Recognition, 2nd edition, Morgan Kaufmann Publishers, 1998.

[2] Watanabe, S., Pattern Recognition: Human and Mechanical, Wiley, 1985.

[3] Weiss, S. I., and Kulikowski, C., Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Networks, Machine Learning and Expert Systems, Morgan Kaufmann Publishers, 1991.

[4] Han, J. and Kamber M., Data Mining : Concepts and Techniques, 2nd edition, Morgan Kaufmann Publishers, 2008.

[5] Favata, J. T., “General Word Recognition: Using Approximate Segment-String Matching”, in 4th International Conference on Document Analysis and Recognition, 1997, pp. 92-96.

[6] Jones, G. J. F., Wright, J. H.,Wrigley, E. N., and Carey, M. J., “Isolated Word Sentence Recognition Using Probabilistic Context- Free Grammar”, IEEE Colloquium on Systems and Applications of Man Machine Interaction Using Speech I/O, 1991, pp. 13/1-13/5.

[7] Jiang, X., Yu, K., and Bunke, H., “Classifier combination for grammar-guided sentence recognition”, in Multiple Classifier Systems, First International Workshop, 2000, pp. 383–392.

[8] Marukatat, S., Artikres, T., Gallinari, P., and Dorizzi, B., “Sentence Recognition through Hybrid Neuro-Markovian Modeling”, in Sixth International Conference on Document Analysis and Recognition, 2001.

[9] Rosenberg, A. E., “Connected Sentence Recognition Using Diphone-like Templates”, in International Conference on Acoustics, Speech, and Signal Processing, 1988, vol.1, pp. 473-476.

[10] Murase, I., and Nakagawa, S., “Sentence Recognition Method Using Word Co-occurrence Probability and Its Evaluation”, in International Conference on Spoken Language Processing, 1990, pp. 1217–1220.

[11] Majewski, M., and Zurada, J. M., “ Sentence Recognition Using Artificial Neural Network”, Knowledge-Based Systems, vol. 27, issue 1 , 2008.

[12] David, V. K., and Rajasekaran, S., Pattern Recognition Using Neural and Functional Networks, Springer, 2008.

[13] Qiu, M., Mai H., and Liao X., “Application of neural network in vehicle license character recognition”, Jisuanji Gongcheng yu Sheji (Computer Engineering and Design). Vol. 29, no. 8, 2008, pp. 2041-2043.

**B. Pandey** is a student of final year of 5-year Integrated Post Graduate Course (BTech + MTech in IT) in Indian Institute of Information Technology and Management, Gwalior, India. His areas of research are soft computing, artificial intelligence, hybrid system design and Biometrics.

**S. Ranjan** is a student of final year of 5-year Integrated Post Graduate Course (BTech + MTech in IT) in Indian Institute of Information Technology and Management, Gwalior, India. field of research includes soft computing and Bioinformatics.

**A. Shukla** is an Associate Professor in the ICT Department of the Indian Institute of Information Technology and Management Gwalior. He completed his PhD degree from NIT-Raipur, India in 2002. He did his post graduation from Jadavpur University, India. He has 22 years of teaching experience. His research interest includes Speech processing, Artificial Intelligence, Soft Computing and Bioinformatics. He has published around 120 papers in various national and international journals/conferences. He is referee for 4 international journals and he is in the Editorial board of International Journal of AI and Soft Computing. He received Young Scientist Award from Madhya Pradesh Government and Gold Medal from Jadavpur University.

**R. Tiwari** is an Assistant Professor in the ICT Department of Indian Institute of Information Technology and Management, Gwalior. Her field of research includes Biometrics, Artificial Neural Networks, Speech Signal Processing, Robotics and Soft Computing. She has published around 60 papers in various national and international journals/conferences. She has received Young Scientist Award from Chhattisgarh Council of Science & Technology and also received Gold Medal in her post graduation.