Intelligent Internetworking Tutoring System by Using Adaline Neuron and Mel Cepstral Coefficients

Miguel A. Ruiz-Jaimes[1], Jorge A. Ruiz-Vanoye[2], Yadira Toledo-Navarro[1], Diego Márquez-Salazar[1], Rodrigo Beltran-Labra[1]

[1] Polytechnic University of the State of Morelos, Jiutepec, Morelos,
Mexico

[2] Autonomous University of Hidalgo, Hidalgo,
Mexico

{mruiz, ytnavarro, mado140055, blro130524}@upemor.edu.mx,
jorge@ruizvanoye.com

Abstract. We present a speech recognition prototype system in Spanish language to contribute to the study of the communication networks branch, alternative and innovative techniques of human computer interaction and the recognition with learning of the use and details about internetworking terms, in addition to the instruction on remote networks to carry out the search of all the possible routes, so subsequently choose the most efficient routes for data exchange among themselves, using dynamic routing by using one or more routing protocols such as: RIP, IPV 2, IGRP, EIGRP, OSPF, to later identify your metric either by jump count, bandwidth, delay and costs which the algorithm tells the protocol values to determine its metric value according to the lowest result you get a best route. The recognition of voice and taking of actions through the same is done using the artificial neuron (ADAptive LInear NEuron) and cepstrales coefficients of Mel as a method of extracting information from the audio files. This prototype, due to its basic neuron nature, was implemented and proved only in an controlled environment, the word's recognition rate it's between the 40% and 80% percent. The recognition eperimentation results are exposed in addition to the design of the intelligent tutor.

Keywords: artificial intelligence, intelligent tutor, speech recognition, adaline, cepstral coefficients, internetworking.

1 Introduction

The problem of learning subjects with high theoretical content is of great impact in various sectors of the population, for example, in energy sector people in charge of maintaining equipment must read a great manual to understand the use and management of each of the devices, this makes people as they read lose interest in not having interactive content that arouses their interest to continue learning. The same thing

happens in the educational sector, for this work we focus on the study of the subject of internetworking where concepts are used that, if they do not know how to use them correctly would endanger very important information of a complete network.

To maintain constant communication from one network to another requires routing protocols, which are responsible for finding the best route between the nodes (routers), to facilitate communication across the network [1]. The most commonly used dynamic routing protocols are: RIP (routing Information Protocol), EIGRP (Enhanced Interior Gateway Routing protocol), OSPF (Open Shortest Path First), is-is (intermediate system to intermediate System), and BGP (Border Gateway Protocol) [2].

The development of an intelligent tutor that provides the student with information about the aforementioned protocols through interactive content such as augmented Reality (RA), and voice recognition, would provide comfort and versatility for the process of Learning, in addition, speech recognition offers the possibility for students to interact with an electronic team [3].

Within the various applications that Artificial intelligence has we can find the recognition of voice of which there are innumerable uses, the functions and characteristics of the speech recognition technology are very wide, for this project we conceptualize the process of interpreting a spoken word to a pre-programmed instruction within a system, this is achieved through the training of an Adaline neuron using five samples of each word to recognize by the system.

2   Related Work

There are several works on the learning of different academic subjects such as STAAM [4] the system of mathematics that focuses only on solving mathematical problems of primary education, the system is based on the Web designed to help elementary students. Barron et al. make use of gamification elements as main components of the EMATIC system (Mathematics Education through ICT) [5], the system is oriented to mobile devices for the teaching of basic mathematical operations. On the other hand, Zatarain et al. present an intelligent learning system that integrates a social network for the learning of mathematics and where it also presents an architecture for the design of the system [6], 3 emphasizing the recognition of the user's emotions.

Thinking about the implementation of speech recognition, Perez and Caballero developed a voice-based emotion recognition system [7], the work considers 640 sentences of eight users, and the following basic emotions were considered: Anger, Happiness, Neutral and Sadness. The percentage of overall effectiveness of said system is 87%. Hernández and Lemus developed an application for training in TOEFL certification through voice recognition [8], the application has a database where it protects synthesized voices for recognition through SAPI (Set of tools for voice recognition created by Microsoft).

Using the Adeline Neuron, Medina et al. developed a biomedical signal processing system [9], the system supports the doctor in the diagnosis of mental illnesses, presents the structure of the neuron and the hardware used to detect eye movements.
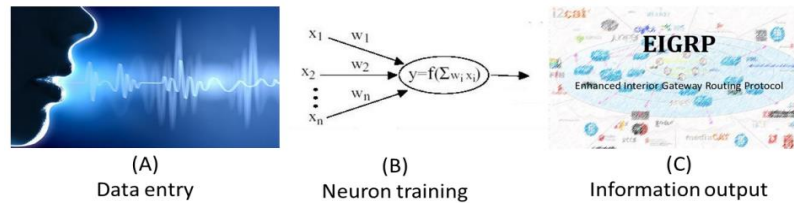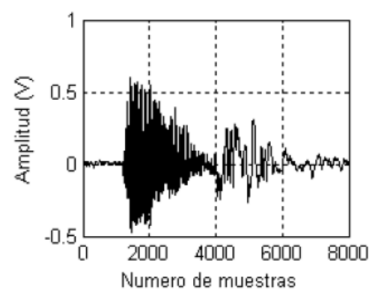
Fig. 1. System design.

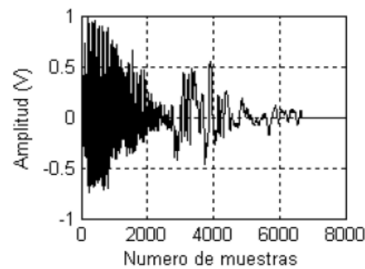**Fig. 2.** Voice signal generated by Matlab (Taken from [10]).

Fig. 3. Filtered and normalized voice signal (Taken from [10]).

3 Materials and Methods

For the development of this research was used a computer with a Operating System Windows 10 with an Intel Core i5 processor (6MB in Caché and 3.20 GHz.), a graphics card NVIDIA GeForce GTX 1070 with 8GB of RAM, integrated microphone and Matlab software. The design of the system developed in this work is shown in Figure:

A. Data entry: In this first stage the word of the Protocol is recorded to be recognized, then it is codified and finally normalized the voice, it is noteworthy that the recording of the voice must be under a controlled environment without ambient noise.

```
function newRecord = norm(record)
    maxi = max(abs(record));
    n = length(record);
    newRecord = zeros(1,n);
     for i=1:1:n
        newRecord(i)=record(i)/maxi;
     end
end
```
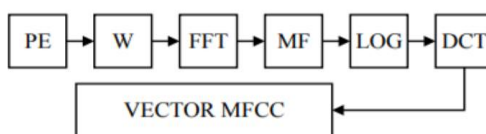
Fig. 4. Signal normalization function.

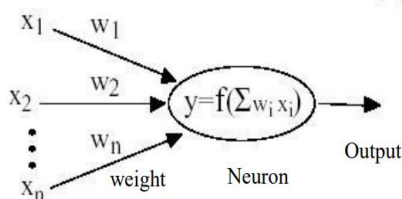**Fig. 5.** Scheme for obtaining the MFCC (Take from [13]).

Fig. 6. Neuron Adaline model (Take from [14]).

B. Neuron training: Once the voice is properly processed, the neuron is trained to recognize the previously recorded word.

C. Information output: Finally when the neuron recognized the type of voicerecorded protocol, the system displays the protocol information on the screen.

The type of audio files that are handled in the system is WAV, in mono system. When the system records the voice, you get a signal very similar to Figure 2, however, when the voice is filtered and normalized there is a signal with a little more amplitude as seen in Figure 3.

Voice normalization is performed through the wave-amplitude standardization function; Figure 4 shows the function that normalizes the audio signal:

MFCC (Cepstrales coefficients in the frequency of Mel), are coefficients for speech representation based on human auditory perception, their frequency bands are located logarithmically [11], this makes the human auditory response modeled more appropriately than linearly spaced bands [12]. Figure 5, contains a diagram for obtaining the FC, where:

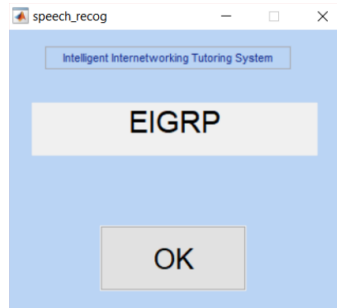- PE: pre emphasis filtering,
- W: Windows,

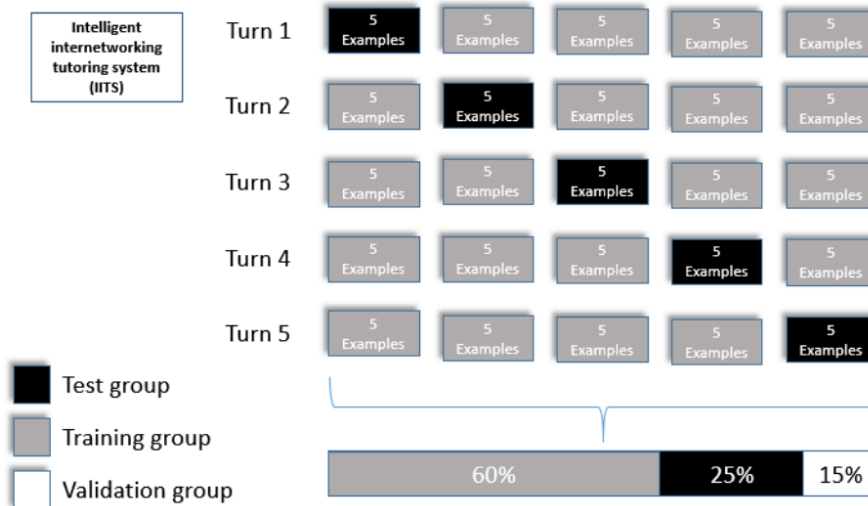**Fig. 7.** Developed software user interface.

Fig. 8. Proportion of used audio scheme.

- FFT: Fast Fourier transformed,
- LOG: Logarithmic scale,
- DCT: Discrete cosine transformed [13].

The training of the system is carried out using the samples corresponding to each word, they establish random weights in a beginning, for later, having the neuron with the corresponding training, they obtain the appropriate values for Classify the words obtained by means of the recording of signal in mono frequency to 44100Khz. Figure 6 refers to the neuron model used in this research work.

Characteristics extraction and neuron trainment was made by using 12 different values taken from the MFCC model, so as it is showed in Figure 6, neuron has X_{12} entries that will define word's singularities to identify each word used in the system trainment.

Table 1. Words used in the recognition prototype implementation and its recognition rate at final test, words are routing protocols.

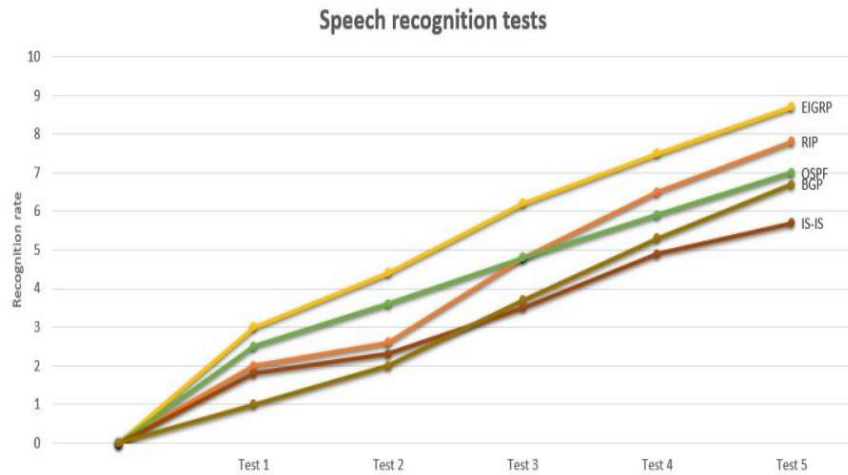| Words	Recognition rate				
Test	1	2	3	4	5
EIGRP	30%	42%	60.5%	74%	88%
RIP	26%	36%	48%	65%	79%
OSPF	20%	26%	49%	59%	70%
BGP	19%	23%	32%	52%	68%
IS-IS	10%	20%	33%	49%	57%

Fig. 9. Graph of speech recognition rate.

4   Problem Statement

Voice recognition plays an important role in networks and telecommunications, which is why it was thought to develop this system that will allow a better understanding and interaction of the cisco network operating systems which will allow the students who have the curricula of CCNA of Cisco [15], to have a better understanding and understanding of each of the network concepts (technologies, standards and protocols) and commands that make up each of the configuration modules of the network devices (routers and switches).

This system will allow through voice recognition that the student has a more dynamic, flexible and understandable knowledge that allows them to have a clearer knowledge of the concepts with the intention of motivating students to accredit each of the modules so that later present your certification exam.

Selected words are the more relevant protocols that CCNA of cisco [15], currently teach in bachelor's degree, due to it is a project prototype, methods that are implemented are just in order to discover and test the system functionality in a real context.

5   Analysis of Results

As a result of the tutored system, the software prototype shown in Figure 7 was obtained, which is able to perform the voice recognition of the five different dynamic routing protocols with which the project is being worked on, the process software operation is to start the software, then record the word or in this case the desired protocol, then the system will make the selection of the appropriate term based in its previous training, after that the system shows on screen information of the selected protocol in addition to the recognition of the same.

Figure 8, shows the selection and proportion of words or recordings used for the training and subsequent operation of the system, in addition to the validation group of the same, in the system were used 5 groups of recordings per word, four are for training and one test, validation was performed later when the system was put into practice with different users.

Figure 9, shows the graph of the training process and the rate of recognition of the words (protocols) used as samples in the project, it is remarkable how the improvement through the training process is notorious and can be concluded from this means that a greater number of training cycles, the better the recognition of the words in the system.

The words used for the recognition and operation of the system were the five dynamic routing protocols previously named (eigrp, ospf, rip, bgp and is-is), based on the recognition of each one of them is displayed relevant information and it is sought that the user out in the same, as well as in related topics where the main objective is to describe the operation of the protocols of dynamic routing as base in the internetworking.

For now, the interaction with the system by the user is through voice commands and selective information elements, which are accessed through the use of the mouse.

Word's recognition rate and words used into system training and implementation divided by tests, are shown in table 1.

Results in table 1 were graphed in figure 9, all tests including final result in test 5. We tried to continue testing after test 5 but results were not improved anymore.

Due to it is a controlled environment, all testings including final test were made with only one student and in the same environment sound conditions.

Selected words (protocols), were used due to are the most important protocols currently teached in Cisco Networking Academy course in informatics bachelors degree in the Politechnical University of State of Morelos and it actually solve the main fact of lack of networking hardware devices and because that is the main porpouse of this research, as a future work it is planning to make a more complex system, using a bigger artificial neural network and implementing the multi-user recognition function, even in an uncontrolled sound environment.

6   Conclusions and Future Research

A system was obtained with a trained neuron that is able to recognize with 70% effectiveness the words established in the program in a controlled environment. The system is able to interpret the voice message recorded by the user and then show on the screen the response obtained as well as information about each protocol.

As future work is intended to implement the use of augmented reality in the intelligent tutoring system, with which the user obtains a greater interactivity with the information, as well as with the physical objects in the system, situation that represents relevance because the direct interaction with the physical components of a training network always improves performance at the time of taking the practice to the implementation of the case in real life.

In addition, you want to expand the catalog of available terms for your search within the system, thereby expanding the possibilities of information and add interaction between related terms.

Also, as part of the improvement and scalability of the project, is intended to make an improvement of the network of neurons for the recognition, thereby expanding also the possibilities of use in different controlled environments or semi controlled, where the user can perform searches with greater freedom and in different sound environmental circumstances.

References

1. Enciso, L., Quezada, P., Fernández, J., Figueroa, B.: Análisis deparámetros de desempeño de los protocolos de enrutamiento Ad Hoc aplicado a un entorno urbano. TELECOM UNI, pp. 1–7 (2015)
2. Paciello, G.: Routing, Fundamental concepts. Journal Cisco Support, pp. 03–47 (2017)
3. Torres, C., Perpiñan, G., Mattos, L., Castro, J.: Automatic speech recognition using fourier transform and neural network. Colombian Journal of Avanced Tecnologies, pp. 116–120 (2005)
4. Zatarain, R., Barrón, M., García, J.: Sistema tutor afectivo para el aprendizaje de las matemáticas usando técnicas de gamificación. Research in Computing Science 111, pp. 83–96 (2016)
5. Barrón, M., Zatarain, R., Aranda, M., Gómez, H., Mejía, J.: Un tutor inteligente, afectivo y configurable para el aprendizaje de números naturales de 3er grado. Research in Computing Science 77, pp. 45–54 (2014)
6. Zatarain, R., Barrón, M., González, F., Oramas, R.: Diseño e implementación de un sistema de aprendizaje inteligente y afectivo para la enseñanza de las matemáticas del segundo año de primaria. Research in Computing Science 77, pp. 31–44 (2014)
7. Pérez, L., Caballero, S., Trujillo, F.: Integración de optimización evolutiva para el reconocimiento de emociones en voz. Research in Computing Science 93, pp. 9–21 (2015)
8. Hernández, M., Lemuz, R.: Multimedia Application for Training in TOEFL Certification Using Speech Recognition. Research in Computing Science 128, pp. 67–75 (2016)
9. Medina, A., Gómez, L.G., Cisneros, A., Melgar, E., Nango, G., Morales, N., Villegas, J.M., García, N.R., Bermudez, J.I.: Processing of Biomedical Signal with Neural Network Adaline. In: Proceedings of the 12th Iberian Conference on Information Systems and Technologies, pp. 2256–2261 (2017)
10. Barrero, V., León, J.P., Sánchez, C., Rosas, F.: Determinación de la arquitectura óptima de una red neuronal tipo perceptrón multicapa para la extracción de características relevantes y

reconocimiento de señales mioeléctricas. IEEE Transactions on Biomedical Engineering, pp. 11–16 (2014)

11. Mantilla-Caeiros, A.V.: Análisis, reconocimiento y síntesis de voz esofática: Instituto Politécnico Nacional (2007)

12. Del Pino, P., Granadillo, I., Miranda, M., Jimenes, C., Díaz, J.A.: Diseño de un sistema de medición de parámetros característicos y de calidad de señales de voz. Journal Ingeniería UC, pp.13–20 (2008)

13. Matínez-Mascorro, G.A., Aguilar-Torres, G.: Sistema para identificación de hablantes robusto a cambios en la voz. INGENIUS Journal of Science and Technology, pp. 45–53 Judith Shamian (2014). The power of unity. In: International Council of Nurses, pp. 301, International Nursing Review (2012)

14. Hong-wu, Y., Ya-Li, L., De-Zhi, H.: Speaker Recognition Based on Weighted Mel-cepstrum. Journal Computer Sciences and Convergence Information Technology, pp. 124–1136 (2009)

15. Ownn, D., Larson, W.: Development of a Genetic Linkage Map for Cisco (Coregonus artedi), to Facilitate Integrated Studies of Adaptive Diversity. Research Gate (2017)