

UNIVERSIDAD NACIONAL DEL CALLAO

FACULTAD DE INGENIERÍA INDUSTRIAL Y DE SISTEMAS

UNIDAD DE INVESTIGACIÓN



INFORME FINAL DEL PROYECTO DE INVESTIGACIÓN

“ANALIZADOR SEMÁNTICO DEL LENGUAJE NATURAL EN IDIOMA CASTELLANO UTILIZANDO REDES NEURONALES ARTIFICIALES”

AUTOR: Sally Karina Torres Alvarado

PERIODO DE EJECUCIÓN: Del 01 de abril de 2021 al 31 de marzo de 2023

Resolución de aprobación N° 242-2021-R.- CALLAO, 27 DE ABRIL DE 2021)

Callao, 2023

DEDICATORIA

La presente investigación la dedico con todo mi amor a mi familia. Tu afecto, cariño y tiempo son los detonantes de mi felicidad y mis ganas de seguir avanzando y seguir buscando lo mejor para nuestra familia. Mis nenas por su paciencia y por ser mi motor en poder darles un buen ejemplo. Y a mi madre por ser mi fan número 1 que siempre me alienta a seguir esforzándome y a ver en lo más sencillo lo maravilloso.

AGRADECIMIENTO

Un agradecimiento en especial a mi querido esposo Mario Gastón Borja Borja por la ayuda y el asesoramiento profesional prestada en la presente investigación. A la Universidad Nacional del Callao por el apoyo brindado en la elaboración de Trabajos de investigación, especialmente a ICICYT por la ayuda prestada y a todos quienes han contribuido en la elaboración de esta investigación.

Índice general

RESUMEN.....	7
ABSTRACT	8
INTRODUCCIÓN	9
CAPITULO I: PLANTEAMIENTO DEL PROBLEMA	11
1.1. Descripción de la realidad problemática.....	11
1.2. Formulación del problema	13
1.2.1. Problema general	13
1.2.2. Problemas específicos	13
1.3. Objetivos.....	13
1.3.1. Objetivos generales	13
1.3.2. Objetivos específicos	13
1.4. Limitantes de la investigación	13
1.4.1. Teórica	14
1.4.2. Temporal.....	14
1.4.3. Espacial	14
CAPITULO II: MARCO TEÓRICO.....	15
2.1. Antecedentes.....	15
2.1.1. Nacional.....	15
2.1.2. Internacional.....	15
2.2. Marco.....	16
2.2.1. Teórico	16
2.2.1.1. Redes neuronales artificiales.....	16
2.2.2. Conceptual.....	21
2.2.2.1. Procesamiento del lenguaje natural(PLN).....	21
2.2.2.2. Comunicación	22
2.2.2.2.1. Comunicación y Comunicación Humana	22
2.2.2.2.2. Funciones básicas de la comunicación	23
2.3. Definición de términos básicos	24
CAPITULO III: HIPÓTESIS Y VARIABLES.....	25
3.1. Hipótesis	25

3.1.1. Hipótesis general	25
3.1.2. Hipótesis específicas.....	25
3.2. Definición conceptual de variables	25
3.3. Operacionalización de variables	25
CAPITULO IV: DISEÑO METODOLÓGICO.....	26
4.1. Tipo y diseño de la investigación	26
4.2. Método de investigación	26
4.3. Población y muestra	26
4.4. Lugar de estudio y periodo desarrollado.....	26
4.5. Técnicas e instrumentos para la recolección de la información	26
4.6. Análisis y procesamiento de datos	26
CAPITULO V: RESULTADOS	28
5.1. Diseño	28
5.1.1. Diseño de la red neuronal artificial	28
5.1.2. Aprendizaje de la red neuronal artificial	30
5.1.3. Operación de la red neuronal artificial.....	32
5.2. Implementación.....	33
5.2.1. Implementación de la red neuronal artificial	33
5.2.2. Implementación del algoritmo de aprendizaje	36
5.2.3. Interface de usuario.....	38
5.3. Pruebas de operación	40
CAPITULO VI: DISCUSIÓN DE RESULTADOS	57
6.1. Contrastación y demostración de la hipótesis con los resultados	57
6.2. Contrastación de los resultados con otros estudios similares	57
6.3. Responsabilidad ética	57
CONCLUSIONES.....	58
RECOMENDACIONES	58
REFERENCIAS BIBLIOGRÁFICAS.....	59
ANEXOS.....	60
Anexo A: Programa de generación de estructuras de las RNAs.....	61
Anexo B: Programa de generación de las RNAs	63
Anexo C: Programa de aprendizaje de las RNAs	67

Anexo C: Matriz de consistencia.....	74
--------------------------------------	----

Índice de figuras

Figura 1. Conexión de neuronas biológicas	16
Figura 2. Neurona biológica y su equivalente artificial	16
Figura 3. Representación del cálculo de una NA.	17
Figura 4. Función de transferencia tipo limitador fuerte	18
Figura 5. Función de transferencia lineal	19
Figura 6. Función de transferencia sigmoideal.....	20
Figura 7. Red neuronal artificial de 3 capas	21
Figura 8. Representacion simplificada del perceptron multicapa.....	21
Figura 9. Perceptron	22
Figura 10. Perceptron multicapa	38
Figura N 11. Estructura de las sub redes neuronales artificiales	39
Figura N 12. Algoritmo de aprendizaje de la RNAs	40
Figura N 13. Algoritmo de operación de la RNA.....	42
Figura 14. Base de datos.....	43
Figura 15. Estructura de la RNA.....	43
Figura 16. Datos de la RNA	44
Figura 17. Pesos de las entradas de las neuronas	44
Figura 18. Evolución de aprendizaje de la palabra ANIMAL.....	45
Figura 19. Evolución de aprendizaje de la palabra EL	46
Figura 20. Evolución de aprendizaje de la palabra CELULAR	46
Figura 21. Evolución de aprendizaje de la palabra COSA	46
Figura 22. Ejecución del programa de aprendizaje	47
Figura 23. Interface del usuario.....	48
Figura 24. Generación de diez neuronas adicionales.....	48
Figura 25. Interface de registro de oraciones	49
Figura 26. Generación de palabras.....	49
Figura 27. Evolución del aprendizaje de la palabra CUADERNO	53
Figura 28. Evolución del aprendizaje de la palabra GATO.....	54
Figura 29. Evolución del aprendizaje de la palabra BUS	54

Figura 30. Ejecución del aprendizaje de las oraciones “El raton es un animal” y “El gallo es un animal.....60
Figura 31. Evolución del aprendizaje de la oración “El raton es un animal”60
Figura 32. Evolución del aprendizaje de la oración “El gallo es un animal”61

Índice de tablas

Tabla 1. Oraciones para las pruebas.....	49
Tabla 2. Lista de palabras	51
Tabla 3. Resultado de reconocimiento de todas las palabras de prueba	54
Tabla 4. Prueba de la RNA 7, asignada a la palabra Loro	57
Tabla 5. Prueba de interpretación de las 10 primeras oraciones	61
Tabla 6. Prueba de interpretación de la sexta oración	63

RESUMEN

Las propuestas de desarrollo de interfaces hombre maquina en lenguaje natural tienen una capacidad de comunicación muy limitada ya que en general son desarrolladas en base a técnicas convencionales de programación utilizando generalmente conjuntos de preguntas – respuestas y no incluyen el proceso de comunicación que tiene el lenguaje natural, que es el análisis léxico, análisis sintáctico, análisis semántico y construcción de la respuesta.

En este trabajo de investigación se presenta una propuesta del análisis semántico automático para el idioma español utilizando redes neuronales artificiales. Para esto se diseñó e implemento la red neuronal artificial, se desarrolló el proceso de aprendizaje de la red neuronal artificial para representar en la computadora, el conocimiento expresado a través de oraciones gramaticales, y se probó la efectividad de la propuesta en la comprensión de las oraciones que corresponde al análisis semántico.

El análisis semántico implementado se realizó utilizando una red neuronal artificial que al poner en la entrada la oración, reconoce las palabras de la oración y posteriormente interpreta la oración con el conocimiento almacenado sobre la oración durante el proceso de aprendizaje.

Los resultados de pruebas se realizaron para un grupo de oraciones nominativas, para lo cual se realizó el proceso de aprendizaje de la red neuronal artificial con las palabras que contienen las oraciones en idioma español, seguido de proceso de aprendizaje de las oraciones para que la red neuronal artificial memorice el conocimiento transmitido por las oraciones y para probar si el sistema propuesto realiza el análisis semántico se ingresó las oraciones al sistema y se obtuvo una interpretación de 100% para la oraciones utilizadas en el proceso de aprendizaje.

Con los resultados obtenidos se concluye que se puede utilizar las redes neuronales artificiales para la implementación de análisis semántico innato que realizan las personas a las oraciones gramaticales.

Palabras clave: análisis semántico, idioma español, redes neuronales artificiales, procesamiento de lenguaje natural, inteligencia artificial

ABSTRACT

Proposals for the development of human-machine interfaces in natural language have a very limited communication capacity since they are generally developed based on conventional programming techniques, generally using sets of questions - answers and do not include the communication process that natural language has. , which is lexical analysis, syntactic analysis, semantic analysis and response construction.

This research work presents a proposal for automatic semantic analysis for the Spanish language using artificial neural networks. For this, the artificial neural network was designed and implemented, the learning process of the artificial neural network was developed to represent in the computer, the knowledge expressed through grammatical sentences, and the effectiveness of the proposal in understanding the sentences was tested. sentences corresponding to semantic analysis.

The implemented semantic analysis was carried out using an artificial neural network that, when inputting the sentence, recognizes the words of the sentence and later interprets the sentence with the knowledge stored about the sentence during the learning process.

The test results were carried out for a group of nominative sentences, for which the learning process of the artificial neural network was carried out with the words that contain the sentences in Spanish, followed by the learning process of the sentences so that the network artificial neural memorize the knowledge transmitted by the sentences and to test if the proposed system performs the semantic analysis, the sentences were entered into the system and an interpretation of 100% was obtained for the sentences used in the learning process.

With the results obtained, it is concluded that artificial neural networks can be used for the implementation of innate semantic analysis that people perform to grammatical sentences.

Keywords: semantic analysis, Spanish language, artificial neural networks, natural language processing, artificial intelligence

INTRODUCCIÓN

Lo ideal sería que las computadoras y los sistemas computacionales entiendan un lenguaje natural hablado, por ello para describir los problemas de la presente Investigación explicaremos la evolución del procesamiento en lenguaje natural, para luego analizar los avances que tienen los compiladores, y finalmente concluir en los problemas que involucra la creación de un analizador semántico de lenguaje natural para el idioma castellano.

El procesamiento del lenguaje natural (PLN), es un área de investigación en esta en continuo desarrollo, se aplica actualmente en diferentes actividades como pueden ser la traducción automática, elaboración automática de resúmenes, sistemas de recuperación de información, interfaces en lenguaje natural, etc. Si bien es cierto que en los últimos años se han realizado muy buenos avances, los fundamentos teóricos del PLN se encuentran todavía en estado de desarrollo, es claro que los obstáculos a superar en el estudio del tratamiento del lenguaje son considerables, los resultados obtenidos y la evolución en los últimos años ubican al Procesamiento del lenguaje natural en una posición para poder dirigir las aplicaciones informáticas del futuro: los medios de comunicación del usuario con la computadora pueden ser más flexibles y el acceso a la información almacenada más eficiente.

Como podemos ver una de las fases importantes en la interpretación de lenguajes, es el análisis semántico, el cual también es necesario realizar en el procesamiento del lenguaje natural, entonces desarrollar una forma eficiente de realizar el análisis semántico permitiría obtener una interface adecuada para las computadoras o sistemas computacionales.

Algunos de los métodos para resolver este problema, es utilizando teoría de lenguajes y compiladores, pero resultarían muy difíciles de implementar por la cantidad de palabras del idioma castellano y porque la semántica resulta muy compleja en comparación de los lenguajes de programación formales.

Por otro lado, las redes neuronales artificiales han tomado relevancia para la solución de problemas complejos.

Considerando la situación del avance en el procesamiento de lenguaje natural en la presente investigación se planteó el problema principal a resolver como: ¿Se puede

desarrollar un analizador semántico para el procesamiento del lenguaje natural para el idioma castellano?, con los problemas específicos: ¿Se puede desarrollar un software que comprenda las oraciones del lenguaje natural utilizando redes neuronales artificiales? y ¿Se puede representar el conocimiento transmitido a través de oraciones del lenguaje natural por medio de redes neuronales artificiales?

Para resolver el problema planteado se definió el objetivo general de “Desarrollar un analizador semántico de lenguaje natural para el idioma castellano”, alcanzando los objetivos específicos de “Desarrollar un software que comprenda las oraciones del lenguaje natural utilizando redes neuronales artificiales” y “Representar el conocimiento transmitido a través de oraciones dadas en lenguaje natural por medio de redes neuronales artificiales”

La investigación planteada tiene mucha importancia ya que permitiría avanzar en la implementación de interfaces hombre-máquina inteligentes que permita la comunicación a las personas con las maquinas en lenguaje natural y se justifica en el aspecto científico ya que la presente investigación hará un gran aporte al conocimiento científico y tecnológico ya que permitirá ser usado como base para sistemas que necesiten procesar y entender oraciones en lenguaje natural del idioma castellano, en el aspecto tecnológico se justifica ya que se desarrollará un prototipo para realizar las pruebas de análisis semántico del lenguaje natural en castellano y en el aspecto metodológico se justifica ya que se aplicara una metodología utilizada ampliamente en donde se diseña la propuesta de solución, se implementa un prototipo computacional y se realizan las pruebas que permitirá demostrar la funcionalidad de la propuesta, y en el aspecto social se justifica ya que se propondrá una solución que provee de una herramienta que ayuda a la interacción de las personas con la computadora y dispositivos móviles.

CAPITULO I: PLANTEAMIENTO DEL PROBLEMA

1.1. Descripción de la realidad problemática

Lo ideal sería que las computadoras y los sistemas computacionales entiendan un lenguaje natural hablado, por ello para describir los problemas de la presente Investigación explicaremos la evolución del procesamiento en lenguaje natural, para luego analizar los avances que tienen los compiladores, y finalmente concluir en los problemas que involucra la creación de un analizador semántico de lenguaje natural para el idioma castellano.

El procesamiento del lenguaje natural (PLN), es un área de investigación en esta en continuo desarrollo, se aplica actualmente en diferentes actividades como pueden ser la traducción automática, elaboración automática de resúmenes, sistemas de recuperación de información, interfaces en lenguaje natural, etc. Si bien es cierto que en los últimos años se han realizado muy buenos avances, los fundamentos teóricos del PLN se encuentran todavía en estado de desarrollo, es claro que los obstáculos a superar en el estudio del tratamiento del lenguaje son considerables, los resultados obtenidos y la evolución en los últimos años ubican al Procesamiento del lenguaje natural en una posición para poder dirigir las aplicaciones informáticas del futuro: los medios de comunicación del usuario con la computadora pueden ser más flexibles y el acceso a la información almacenada más eficiente.

Por ejemplo, con la creación de interfaces inteligentes el usuario tendría la facilidad para interactuar con la computadora en lenguaje natural. El uso de técnicas de PLN puede tener un alto impacto en la gestión documental y en los sistemas de traducción automática. Pero, por otro lado, la complejidad implícita en el tratamiento del lenguaje nos da limitaciones en los resultados y, por tanto, aplicaciones en áreas de conocimiento son muy concretas y con un uso restringido del lenguaje.

Los diversos idiomas que se originaron en el mundo en las diversas naciones, también conocidos como lenguaje natural, son complejos y uno de los acercamientos más difundidos a un lenguaje natural son los lenguajes de programación, los cuales se crearon con el objetivo de comunicar a las computadoras o sistemas computacionales el orden de ejecución de instrucciones de un programa. Estos lenguajes de programación,

para lograr entender el lenguaje formal consta de dos fases frontend (las fases de análisis) y back end (las fases de generación y optimización de código). Estas fases se comunican a través de una codificación intermedia (generada por el frontend), que puede ser una representación de la sintaxis del programa (un árbol sintáctico abstracto) o bien puede ser un programa en un lenguaje intermedio. El frontend depende del lenguaje fuente y casi siempre es independiente (o debe serlo) de la máquina objeto para la que se va a generar código; el back end depende del lenguaje objeto y debe ser independiente del lenguaje fuente (excepto quizá para algún tipo de optimización). El front end tiene tres etapas el análisis léxico, el análisis sintáctico y el análisis semántico. El back end tiene las siguientes etapas la generación del código intermedio, la optimización del código y por último la generación del código de máquina.

Como podemos ver una de las fases importantes en la interpretación de lenguajes, es el análisis semántico, el cual también es necesario realizar en el procesamiento del lenguaje natural, entonces desarrollar una forma eficiente de realizar el análisis semántico permitiría obtener una interface adecuada para las computadoras o sistemas computacionales (Baliri, 2014).

Una de las capacidades que tienen las personas para comunicarse, es el lenguaje el cual puede expresarse mediante signos orales o escritos.

El lenguaje natural es la forma que usamos las personas para comunicarnos con otras personas además el lenguaje puede realizarse mediante preguntas, expresar, emociones, describir hechos, etc, se trata entonces de un lenguaje aprendido como el español, el francés, el ruso, el inglés, etc. Realizar el análisis semántico para el lenguaje natural (castellano) es tedioso debido a los distintos casos gramaticales que se presentan y al número de palabras que tiene el lenguaje natural (castellano). También el proceso de realizar el análisis semántico requiere de la categorización de las estructuras gramaticales lo cual conlleva a un procesamiento muy complejo.

Algunos de los métodos para resolver este problema, es utilizando teoría de lenguajes y compiladores, pero resultarían muy difíciles de implementar por la cantidad de palabras del idioma castellano y porque la semántica resulta muy compleja en comparación de los lenguajes de programación.

1.2. Formulación del problema

1.2.1. Problema general

En base a lo expresado, el problema de la presente investigación se plantea de la siguiente forma:

¿Se puede desarrollar un analizador semántico para el procesamiento del lenguaje natural para el idioma castellano?

1.2.2. Problema específicos

- ¿Se puede desarrollar un software que comprenda las oraciones del lenguaje natural utilizando redes neuronales artificiales?
- ¿Se puede representar el conocimiento transmitido a través de oraciones del lenguaje natural por medio de redes neuronales artificiales?

1.3. Objetivos

1.3.1. Objetivos general

Desarrollar un analizador semántico de lenguaje natural para el idioma castellano utilizando redes neuronales artificiales.

1.3.2. Objetivos específicos

- Desarrollar un software que comprenda las oraciones del lenguaje natural utilizando redes neuronales artificiales.
- Representar el conocimiento transmitido a través de oraciones dadas en lenguaje natural por medio de redes neuronales artificiales.

1.4. Limitantes de la investigación (teórico, temporal, espacial)

Existen diversas formas de realizar análisis semántico, los cuales se utilizan en compiladores de lenguajes de programación sin embargo está demostrado que funcionan de manera eficiente en lenguajes formales, pero el uso de ellos en el procesamiento del lenguaje natural no ha tenido buenos resultados.

La presente Investigación estará limitada por un número fijo de oraciones en el Idioma castellano en formato escrito, que tendrán que ser analizadas por un prototipo de laboratorio de analizador semántico en base a redes neuronales artificiales, pero que permita realizar los experimentos correspondientes en su totalidad.

1.4.1. Teórica

Socialmente está delimitado a todas las personas de habla hispana de los cinco continentes, ya que en la actualidad el uso de computadoras y dispositivos móviles está generalizado, y como la presente Investigación pretende proporcionar una teoría que ayude a mejorar la interacción hombre máquina.

1.4.2. Temporal

La presente investigación corresponde a una investigación actual ya que el avance de las comunicaciones con las maquinas es algo inminente ahora usamos teclados, pero en un futuro próximo será hablado y por lo tanto esta investigación ayudará a mejorar esta comunicación hombre máquina.

1.4.3. Espacial

La presente investigación está dirigida a todos los países de habla castellana en América tenemos sur América y centro América, en Europa España, en África Guinea Ecuatorial y en países como Filipinas, Marruecos, etc. es uno de los dos o tres idiomas oficiales.

CAPITULO II: MARCO TEÓRICO

2.1. Antecedentes

2.1.1. Nacional

No se encontraron trabajos relevantes que correspondan a un nivel de investigación científica, sobre el análisis semántico automático aplicando técnicas computacionales, del idioma castellano o temas referentes al procesamiento de lenguaje natural.

2.1.2. Internacional

En *SemBRS: A Semantic Analysis based Book Retrieval Approach* (Xie, 2020) presenta un método para identificar libros con el contenido requerido por los usuarios que no tienen las palabras clave exactas, basado en análisis semántico de la descripción del usuario, sobre el contenido que necesitan.

En *Analysis of Semantic Comprehension Algorithms of Natural Language Based on Robot's Questions and Answers* (Li, 2020) presenta el análisis de algoritmos de comprensión semántica implementado con aprendizaje profundo, de preguntas y respuestas de lenguaje natural de robots.

En *English text quality analysis based on recurrent neural network and semantic segmentation* (Chen, 2020) presentan el análisis de la calidad del inglés escrito utilizando la segmentación semántica con redes neuronales artificiales recurrentes, que a diferencia de las redes neuronales tradicionales almacenan información.

En *Incorporating Context-Relevant Concepts into Convolutional Neural Networks for Short Text Classification* (Jingyun Xu, 2019) presentan la clasificación de texto corto mediante la incorporación de conceptos relevantes del contexto en una red neuronal artificial convolucional.

En *Semantic Holism and Word Representations in Artificial Neural Networks* (Musil, 2020) presenta un método para representar el holismo semántico y palabras con redes neuronales artificiales.

En *Thai Natural Language Processing* (Tapsai, 2020) presenta el análisis semántico del idioma tailandés.

2.2. Marco

2.2.1. Teórico

2.2.1.1. Redes neuronales artificiales

El cerebro está formado por millones de elementos altamente interconectados que se conocen como neuronas. Estas neuronas tienen tres componentes principales, las dendritas, el cuerpo de la célula o soma, y el axón. Las dendritas, son fibras nerviosas que reciben las señales eléctricas de otras neuronas y receptores y las transmiten al cuerpo de la neurona. El cuerpo de la célula, realiza la suma de esas señales de entrada y el axón es una fibra larga que lleva la señal desde el cuerpo de la célula hacia otras neuronas. El punto de conexión entre el axón de una célula y la dendrita de otra neurona se llama sinápsis, como se muestra en la figura 1.

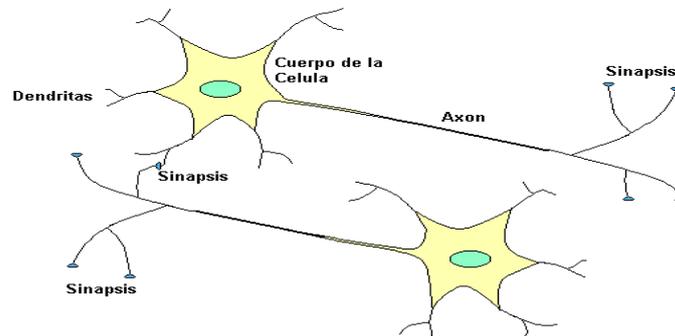


Figura 1. Conexión de neuronas biológicas.

Las RNA no son complejas como las redes biológicas del cerebro humano, pero se asemejan en que las dos son elementos computacionales sencillos, interconectados cuyas conexiones entre neuronas determinan la función de la red.

En la figura 2 muestra una neurona biológica y su equivalente artificial en donde las dendritas son las entradas en el vector X , los pesos equivalen a las sinapsis representados con el vector W , la sumatoria de las entradas por los pesos representa el umbral de activación del axón que es la salida J .

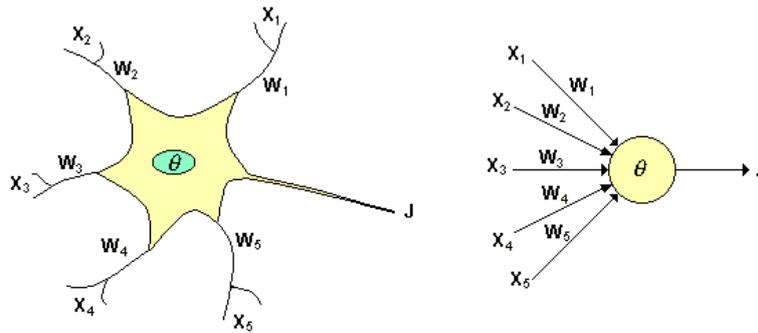


Figura 2. Neurona biológica y su equivalente artificial.

Las señales de entrada a una neurona artificial X_1, X_2, \dots, X_n son variables continuas en lugar de pulsos discretos, como se presentan en una neurona biológica. Cada señal de entrada pasa por una ganancia, conocido como peso sináptico, que tiene función análoga a la de la función sináptica de la neurona biológica. Los pesos de las entradas pueden ser positivos o negativos y el cuerpo de la neurona acumula las señales producto de las entradas por los pesos y la envía a la salida mediante la función de transferencia. El valor de la entrada neta se calcula con (1).

$$\theta = \sum_{i=1}^n W_i X_i \quad (1)$$

En la figura 3, se muestra el flujo de las señales de entrada hacia la salida, en donde, la suma del producto de las entradas por los pesos es un valor ponderado de la entrada que mediante la función de transferencia f se calcula con (2) el estado de la salida.

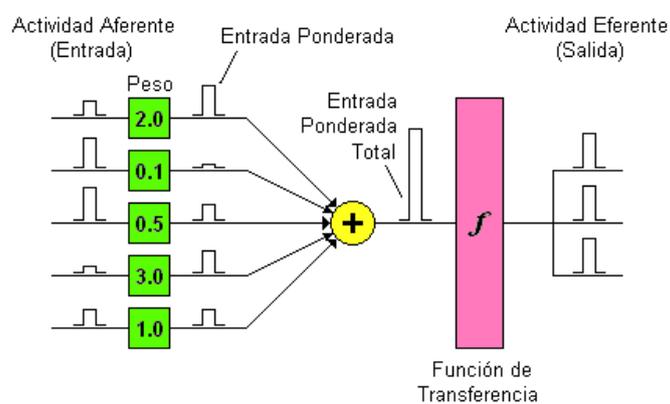


Figura 3. Representación del cálculo de una NA.

$$x_i = f(\text{neta}_i) \quad (2)$$

Donde, f_i es la función de activación para la neurona artificial, que convierte el valor neto de entrada en el valor de la salida.

Como función de transferencia se puede usar el limitador fuerte, que tiene dos estados en donde es 0 si el valor neto es menor que 0 y 1 si el valor neto es mayor que 1, como se representa en la figura 4, y se puede calcular con las formulas (3).

$$a = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \quad (3)$$

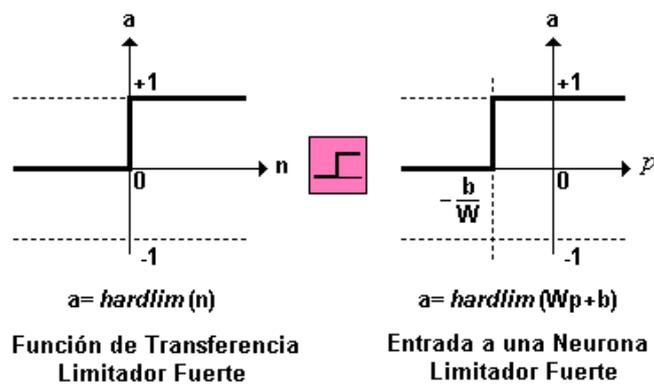


Figura 4. Función de transferencia tipo limitador fuerte.

También, se usa la función de transferencia lineal en donde la entrada es igual a la salida, como se presenta en la figura 5.

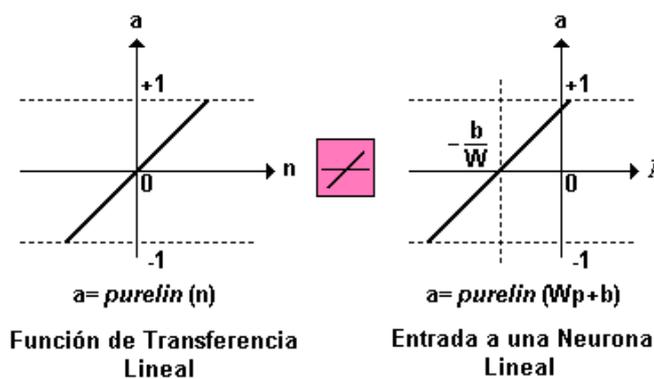


Figura 5. Función de transferencia lineal.

Otra función de transferencia es la sigmoideal, que introduce una activación no lineal diferenciable que se necesita en algunos métodos de aprendizaje como el *Backpropagation*, se calcula con la formulas (4) y se presenta en la figura 6.

$$a = \frac{1}{1 + e^{-n}} \quad (5)$$

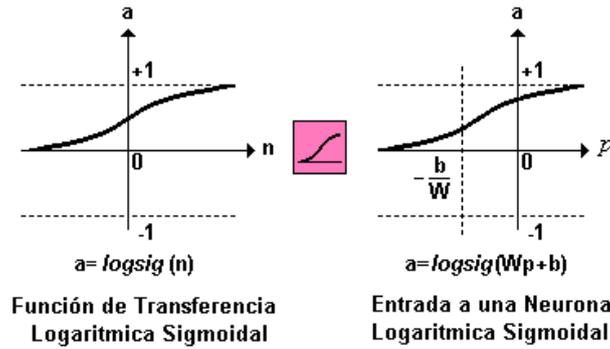


Figura 6. Función de transferencia sigmoidal.

Las redes neuronales artificiales (RNA) se conforman por la interconexión de varias neuronas artificiales, agrupadas por capas, que tienen varias neuronas; de acuerdo a la ubicación de la capa en la RNA, esta recibe diferentes nombres.

Capa de entrada: Recibe las señales de la entrada de la red, a veces no se considera el vector de entrada como una capa ya que no se realiza ningún proceso.

Capas ocultas: Estas capas son las que no tienen conexión con el exterior, sus elementos pueden tener diferentes conexiones y son estas las que determinan las diferentes topologías de la red.

Capa de salida: Recibe los valores de la última capa oculta y entrega la respuesta hacia el exterior.

En las RNA, todas las entradas se conectan a todas las neuronas de la primera capa, las salidas de las neuronas de la primera capa a las entradas de todas las neuronas de la segunda capa y así sucesivamente hasta llegar a la capa de salida que puede tener una o varias salidas como se puede ver en la figura 7.

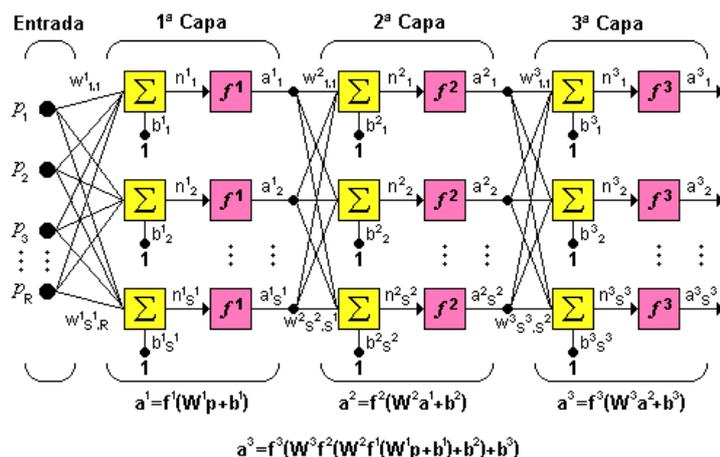


Figura 7. Red neuronal artificial de 3 capas.

Semejante a las redes neuronales biológicas, las redes neuronales artificiales almacenan el conocimiento en los pesos de las entradas de las neuronas, valores que deben ser modificados por un proceso de aprendizaje que puede ser supervisado y no supervisado.

El aprendizaje supervisado es cuando es conocido el vector de las entradas y el valor de la salida de la RNA y se ajustan los pesos con algoritmos iterativos que modifican los pesos hasta lograr el valor deseado de salida.

En el aprendizaje no supervisado solo tengo el vector de entrada y no conozco el valor de la salida y durante el proceso de aprendizaje se encuentran configuraciones que destacan ciertos resultados que dependen de los datos de la entrada.

La red Perceptrón fue propuesta por el psicólogo Frank Rosenblatt en el año 1957. Su objetivo fue representar algunas propiedades de los sistemas inteligentes, sin considerar los detalles de condiciones específicas y desconocidas para organismos biológicos específicos. Rosenblatt pensaba que la conexión en las redes es aleatoria, y se cuestionaba el análisis de McCulloch Pitts, quien empleaba lógica simbólica para analizar estructuras idealizadas. Rosenblatt consideraba que la herramienta de análisis más apropiada era la teoría de probabilidades, y esto lo llevó a una teoría de separabilidad estadística que utilizaba para caracterizar las propiedades más visibles de estas redes de interconexión ligeramente aleatorias.

La neurona de salida del Perceptrón que se muestra en la figura 6, realiza la suma ponderada de las entradas, y pasa el resultado a la función de transferencia de escalón y debe responder +1 si el patrón presentado pertenece a la clase A, o -1 si el patrón pertenece a la clase B.

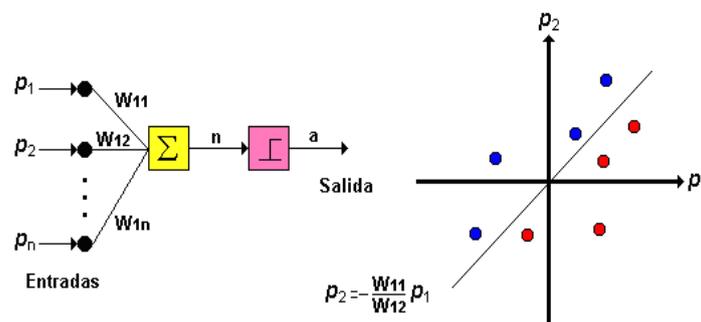


Figura 8. Perceptron.

La red neuronal artificial de Perceptrones tiene múltiples entradas y que incluya una entrada adicional con una ganancia b , como se puede ver en la figura 7 en donde se muestran las conexiones entre sus nodos de entrada y las neuronas de salida.

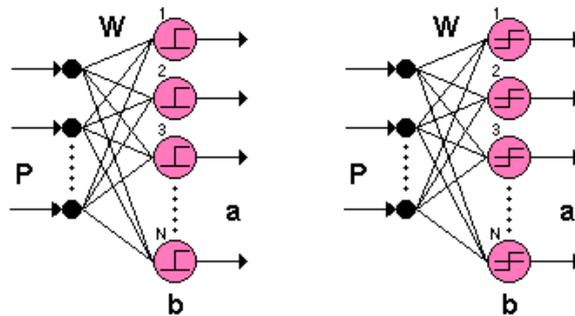


Figura 9. Perceptron multicapa.

2.2.2. Conceptual

2.2.2.1. Procesamiento del lenguaje natural (PLN)

El PLN es el desarrollo de técnicas para hacer software computacional que pueda comprender el conocimiento que transmite el lenguaje natural que hablan las personas. En este proceso intervienen las disciplinas como lingüística, ingeniería informática, filosofía, matemáticas y psicología. En vista que existen diversas áreas del conocimiento, a esta disciplina se conoce como ciencia cognitiva. (Cortez Vasquez, 2009)

Para el desarrollo de propuestas en PLN se están utilizando técnicas de inteligencia artificial, como:

- Redes neuronales artificiales,
- Lenguajes de programación declarativos,
- Algoritmos de búsqueda, y
- Lógica difusa.

Se realizan investigaciones para utilizar el lenguaje natural para hacer diversas tareas automáticamente y como representar el conocimiento en la computadora.

En general, las propuestas existentes que realizan el procesamiento de lenguaje natural tienen la siguiente estructura:

Primero, se representa el texto mediante una representación formal que contiene las características requeridas para el problema específico que se desea resolver.

Luego, el programa computacional utiliza esta representación de conocimiento, para hacer las búsquedas correspondientes.

Finalmente, la representación formal de lenguaje natural se convierte a expresiones en lenguaje natural para responder.

2.2.2.2. Comunicación

Lengua natural son una la variedad de lenguajes humanos que utilizan grupos de personas para comunicarse y tienen una sintaxis. El lenguaje natural es la lengua o idioma hablado y en su mayoría escritos por grupos humanos naciones para comunicarse. Los lenguajes naturales se generaron espontáneamente en los grupos de personas por la necesidad de comunicarse, a diferencia de otras lenguas artificiales, que fueron creadas con algún fin, como los lenguajes de programación, lenguajes formales basados en lógica matemática. Los lenguajes formales son los primeros intentos para mejorar la interacción hombre máquina.

2.2.2.2.1. Comunicación y Comunicación Humana

Según el American College Dictionary, comunicar es "formular o intercambiar pensamientos, opiniones o información, de manera oral, o por escrito o a mediante signos".

Newton, por su lado, postulo el principio de acción y reacción como un axioma fundamental de la mecánica: "las acciones mutuas de dos cuerpos son siempre iguales y dirigidas en sentidos contrarios". El esquema acción/reacción, estímulo/respuesta, que cominó durante dos siglos, redujo la comunicación a la dimensión física y mecanicista más elemental. Comunicación es el intercambio o interacción entre dos cuerpos. Si los cuerpos son organismos vivos, la interacción es más compleja. Estas propuestas clásicas son las bases del tratamiento de la comunicación en la actualidad que reflejan con gran precisión técnica la teoría de la información de Shannon y la Cibernética de Norbert Wiener.

Weaver realza el valor de la teoría matemática de la comunicación de la siguiente manera:

- Es una teoría generalizada que no es necesario definir qué tipo de símbolos se están considerando: si se trata de letras escritas o palabras, notas musicales, palabras habladas, música sinfónica, o cuadros.
- En el proceso de la comunicación, el mensaje puede cambiar de forma, y los cambios son realizados por los diversos transmisores. En el lenguaje oral, el cerebro es la fuente de conocimiento emisor, la voz es el mecanismo transmisor que usa el aire como medio de trasmisión de las ondas sonoras; pero en la comunicación telefónica es necesario convertir las ondas sonoras de la voz en una corriente eléctrica.
- El receptor es el mecanismo inverso del transmisor, ya que transforma o decodifica la señal transmitida por el medio, en un mensaje inteligible para el destinatario. Si la comunicación es cara a cara existe retroalimentación en donde el emisor puede reforzar el mensaje de acuerdo a la comprensión del destinatario. (Cortes, 1999)
- El principio de Shannon y Weaver se aplica a cualquier tipo de comunicación.

Existen cuatro planos de análisis en los sistemas de comunicación en diferentes áreas de investigación:

- Plano intraorgánico: procesos internos del organismo, como la recogida y elaboración de información biológica, psicológica, etc.
- Plano interorgánico: procesos lingüísticos o psicológicos entre organismos.
- Plano organizacional: procesos de comunicación institucionalizados, sistema de información del entorno de carácter más o menos interpersonal, etc.
- Plano tecnológico: equipo, aparatos y programas establecidos para generar, almacenar, elaborar, transmitir, distribuir datos.

2.2.2.2.2. Funciones básicas de la comunicación

Todos los organismos tienen dos funciones en la comunicación:

- Adaptación a los objetos del entorno, y
- Modificación de la conducta de otros organismos.

Así vez, cada una de estas funciones tiene dos a nivel intercomunicacional: Función informativa de los sistemas vivos tienen cierta viabilidad en el entorno si poseen medios adecuados para adquirir y procesar información sobre sí mismos y del entorno; función integrativa, acumulativa y autoorganizativa de los mensajes de los entornos

manteniendo el equilibrio y la estabilidad utilizadno la retroalimentación; funciones de mando e instrucción en las relaciones jerárquicas entre superior-subordinado; funciones de influencia y persuasión en las relaciones interpersonales de todo nivel, en la política, la publicidad o la oratoria.

En las funciones comunicativas predomina la dimensión pragmática. Las señales utilizadas en la mayoría de comunicaciones animales cumplen con estas funciones generales.

2.3. **Definición de términos básicos**

2.3.1. **Neurona artificial.** – La neurona artificial es un modelo simplificado de la neurona biológica.

2.3.2. **Red Neuronal artificial.** - Las redes neuronales artificiales son modelos simples del funcionamiento del sistema nervioso. Las unidades básicas son las neuronas artificiales, que generalmente se organizan en capas.

2.3.3. **Procesamiento del Lenguaje Natural.** - El Procesamiento del Lenguaje Natural es el campo de conocimiento de la Inteligencia Artificial que se ocupa de la investigar la manera de comunicar las máquinas con las personas mediante el uso de lenguas naturales, como el español, el inglés o el chino.

2.3.4. **Aprendizaje supervisado.** - Los algoritmos de aprendizaje supervisado basan su aprendizaje en un juego de datos de entrenamiento previamente etiquetados. Por etiquetado entendemos que para cada ocurrencia del juego de datos de entrenamiento conocemos el valor de su atributo objetivo. Esto le permitirá al algoritmo poder “aprender” una función capaz de predecir el atributo objetivo para un juego de datos nuevo.

CAPITULO III: HIPÓTESIS Y VARIABLES

3.1. Hipótesis

3.1.1. Hipótesis general

Las redes neuronales artificiales permitirán desarrollar un analizador semántico de lenguaje natural para el idioma castellano.

3.1.2. Hipótesis específicas

- Las redes neuronales artificiales permitirán desarrollar un software que comprenda las oraciones del lenguaje natural.
- Las redes neuronales artificiales permitirán representar el conocimiento transmitido a través de oraciones dadas en lenguaje natural.

3.2. Definición conceptual de variables

Variable independiente. - La variable independiente es: MÉTODOS Y TÉCNICAS DE APLICACIÓN DE LAS REDES NEURONALES ARTIFICIALES PARA REALIZAR UN ANALIZADOR SEMÁNTICO DE LENGUAJE NATURAL.

Variable dependiente. - La variable dependiente es: EFICIENCIA EN LA COMPRESIÓN DE ORACIONES EN IDIOMA CASTELLANO.

3.3. Operacionalización de variables

Indicadores de la hipótesis general:

Indicador : Complejidad del analizador semántico.

Índice : Número de neuronas en la red (0% - 100%).

Indicador de la hipótesis secundaria:

Indicador 1 : Cantidad de respuestas del sistema para hacer la comprensión.

Índice 1 : Porcentaje de respuestas correctas (0 – 100%).

Indicador 2 : Precisión en la comprensión de oraciones en idioma castellano.

Índice 2 : Porcentaje promedio de precisión (0 – 100%).

CAPITULO IV: DISEÑO METODOLÓGICO

4.1. Tipo y diseño de la investigación

La presente investigación según las variables que los representan es experimental y según el nivel de medición y el análisis de la información es de nivel IV porque es el aporte de un nuevo conocimiento destinado a procurar soluciones de problemas prácticos.

La presente investigación es de nivel Predictiva II por contener un diseño experimental el cual permitirá demostrar la hipótesis y luego de las formalizaciones de las estructuras sintácticas del Idioma castellano se logrará el procesamiento de dichas estructuras en un prototipo.

4.2. Método de investigación

La investigación utilizara el método científico.

4.3. Población y muestra

La población es un prototipo de software que se desarrolló para hacer las pruebas al análisis semántico de oraciones en formato escrito del idioma castellano.

La muestra es toda la población, es decir el prototipo de analizador semántico que se desarrolló.

4.4. Lugar de estudio y periodo desarrollado

La investigación se realizó en la Universidad Nacional del Callao Facultad de ingeniería industrial y de sistemas, escuela profesional de ingeniería de sistemas.

El periodo es de dos años.

4.5. Técnicas e instrumentos para la recolección de la información

La recolección de datos teóricos científicos se realizó mediante la exploración conceptual de materiales bibliográficos y para la demostración de la hipótesis se obtuvo datos experimentales mediante el prototipo tecnológico desarrollado.

4.6. Análisis y procesamiento de datos

Las variables de la investigación no son de carácter estocástico, por lo tanto, no requieren de técnicas estadísticas análisis y tampoco se realizan encuestas. Si no más

bien se aplican técnicas experimentales para obtener datos del prototipo de software desarrollado.

CAPITULO V: RESULTADOS

5.1. Diseño

5.1.1. Diseño de la red neuronal artificial

El análisis semántico consiste en identificar los componentes de una oración e interpretar en sentido de la secuencia de palabras, y al comprender el sentido de la oración se puede construir la respuesta.

En general las personas, en el proceso de comunicación tienen medios dos medios comúnmente usados para captar la oración, que puede ser a través de oído en forma oral y mediante la vista en forma escrita, acto seguido las personas realizan el análisis léxico, que consiste en la verificar si las palabras están en el vocabulario, si una palabra no está en el vocabulario entonces es detectado de manera inconsciente y puede ser que no se posible entender la oración si se trata de una palabro como el objeto sobre el cual se comunica o el verbo o el complemento que puede ser adjetivos entre otros.

Si todas las palabras fueron reconocidas, entonces las personas realizan el análisis sintáctico, que evalúa si la oración está correctamente estructurada, en caso que la estructura se incorrecta, es detectado inmediatamente y puede terminar el proceso si la alteración de la estructura hace incompresible la oración.

Si la estructura de la oración es correcta, entonces pasa al análisis semántico en donde comprende la oración y de acuerdo a los comprendido se da una respuesta, esta puede ser oral o escrita, o de otra índole como pasar objetos, realizar alguna acción más compleja.

El proceso normal en la comunicación es que las oraciones pasan el análisis léxico, el análisis sintáctico, el análisis semántico y recibe una respuesta el interlocutor.

Sin embargo, hay que tener en cuenta que el idioma castellano existió primero antes que existiera la teoría sobre las diferentes fases de la comunicación y la teoría es una descripción del proceso de comunicación. Por otro lado, el proceso de aprendizaje de las personas es primero las letras, luego aprenden las palabras formadas por estas letras, posteriormente aprenden la comprensión y formulación de oraciones formadas por palabras, para comunicar o recibir información de los objetos que nos rodean. En base a esto, este trabajo presenta una propuesta preliminar sobre la implementación artificial

para una computadora, de la comunicación escrita en idioma castellano iniciando por el aprendizaje de palabras, posteriormente compresión de las oraciones y una respuesta escrita utilizando redes neuronales artificiales, en contexto limitado a preguntas al sistema.

La RNA de la solución propuesta está conformada múltiples sub RNAs interconectadas, así, en un primer nivel memoriza las palabras y en un segundo nivel memoriza el conocimiento comunicado mediante las oraciones en idioma español, como se muestra en la figura 10.

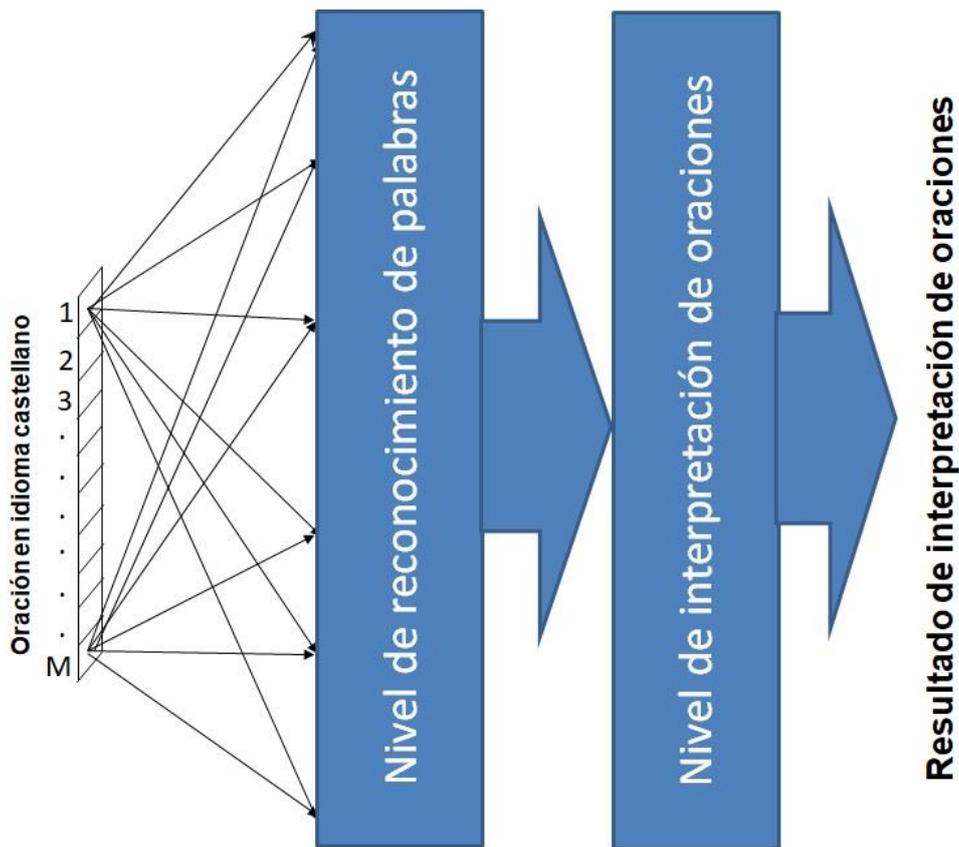


Figura 10. Estructura de la RNA del sistema

Cada uno de los niveles está conformado por sub RNAs interconectadas conformadas por neuronas artificiales organizadas en capas como se muestra en la figura 11. Las sub RNAs tienen L número de neuronas en la entrada, K número de neuronas en la segunda capa, P neuronas en la tercera y pueden existir N capas.

La cantidad exacta de sub RNA, de neuronas en las capas, de número de capas en la sub RNA se determina experimentalmente durante el proceso de aprendizaje.

En el nivel de reconocimiento de palabras se requiere una sub RNA por cada palabra que se use y en el nivel de interpretación de oraciones se requiere una sub RNA por oración.

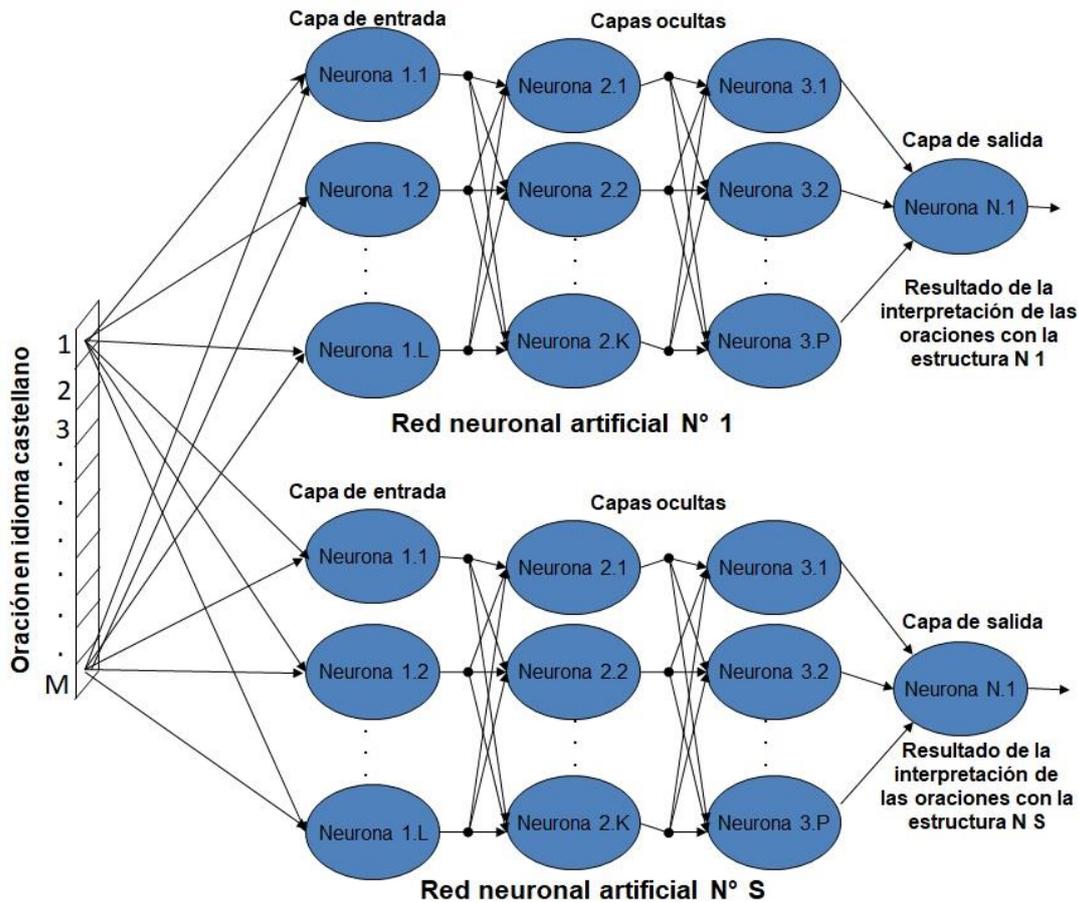


Figura N 11. Estructura de las sub redes neuronales artificiales.

5.1.2. Aprendizaje de la red neuronal artificial

Para que una persona pueda comprender una oración debe conocer las palabras que conforma la oración y además debe tener el conocimiento memorizado sobre lo que comunica la oración.

Por ejemplo, para la oración ¿El perro es un animal?, primero debe conocer todas las palabras que conforman la oración y segundo debe saber que el perro es un animal, conocimiento que debió ser memorizada anteriormente.

De igual forma, las RNAs tiene que pasar por un proceso de aprendizaje para aprender las palabras y las oraciones. En general las personas aprenden a edades tempranas las palabras y conocimiento básico sobre los objetos que nos rodean por indicaciones ya sea de los padres o profesores en el colegio, proceso conocido como aprendizaje

supervisado ya que conocemos la respuesta que debemos dar ante una oración planteada. Utilizando este concepto de aprendizaje supervisado se desarrolla el algoritmo para el aprendizaje de las sub RNAs y se muestra en la figura 12.

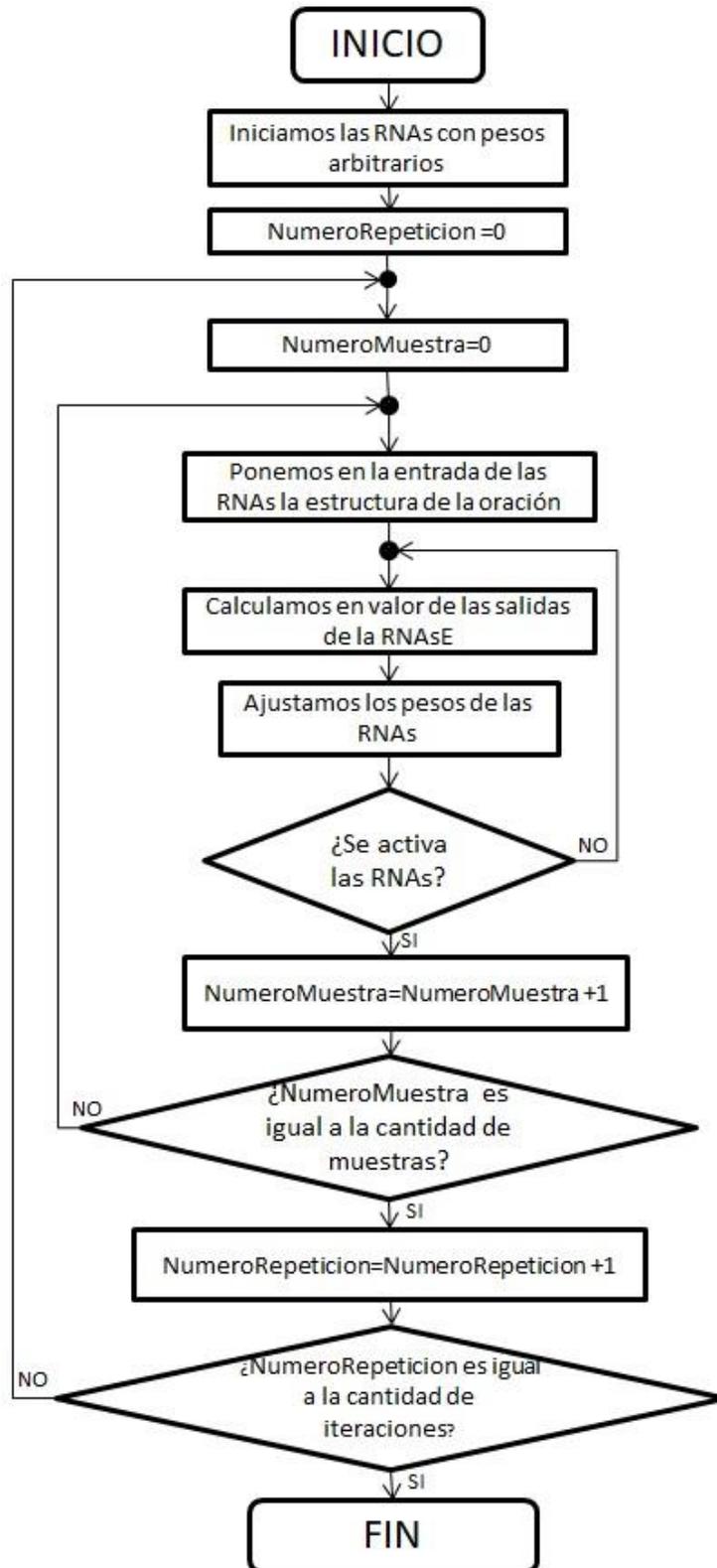


Figura N 12. Algoritmo de aprendizaje de la RNAs.

Para realizar el proceso de aprendizaje se utiliza una cantidad de muestras predefinida y para indicar que muestra se está procesando utilizamos la variable NumeroMuestra, y se repite el aprendizaje con todo el conjunto de muestras por una cantidad de iteraciones hasta que la red neuronal resuelva correctamente todas las muestras, utilizando la variable NumeroRepeticion para indicar que repetición se realiza.

El algoritmo de aprendizaje inicia con NumeroRepeticion=0 y toma el NumeroMuestra=0, se pone la muestra en la entrada de la RNA, se calcula los valores de las salidas, se ajusta los pesos y se verifica si no se activa al RNA, entonces se repite el cálculo de las salidas y se actualiza los pesos hasta que se active. Cuando se activa con la muestra en proceso, se toma la siguiente muestra incrementando NumeroMuestra en 1 y se repite el proceso con todas las muestras. Después de procesar todas las muestras se repite el proceso con todas las muestras la cantidad de repeticiones programadas.

Para la actualización de los pesos se usa el algoritmo de Backpropagation modificado, en el cual en lugar de usar el error medio cuadrático se usa el error instantáneo de la RNA (Mario G. Borja-Borja, 2009). En base a esto, para actualizar los pesos se usa la fórmula:

$$peso_{i,j} = peso_{j,i-1} + error * k_a * entrada_j$$

Donde:

$peso_{j,i}$ – Valor del peso j actual.

$peso_{j,i-1}$ – Valor del peso j anterior.

error = Valor deseado de salida de la RNA - salida real de la RNA

k_a – Velocidad de aprendizaje.

$Entrada_j$ – Entrada de la cual se actualiza el peso.

5.1.3. Operación de la red neuronal artificial

En la figura 13 se presenta el algoritmo de funcionamiento, en el cual primero se carga los pesos calculados durante el proceso de aprendizaje, se pone en la entrada la oración que ingresa al sistema, se calcula las salidas y si se activa la RNA entonces significa que el sistema tiene en la memoria esa consulta y es positiva la respuesta, en caso contrario el sistema desconoce el contenido de la oración.

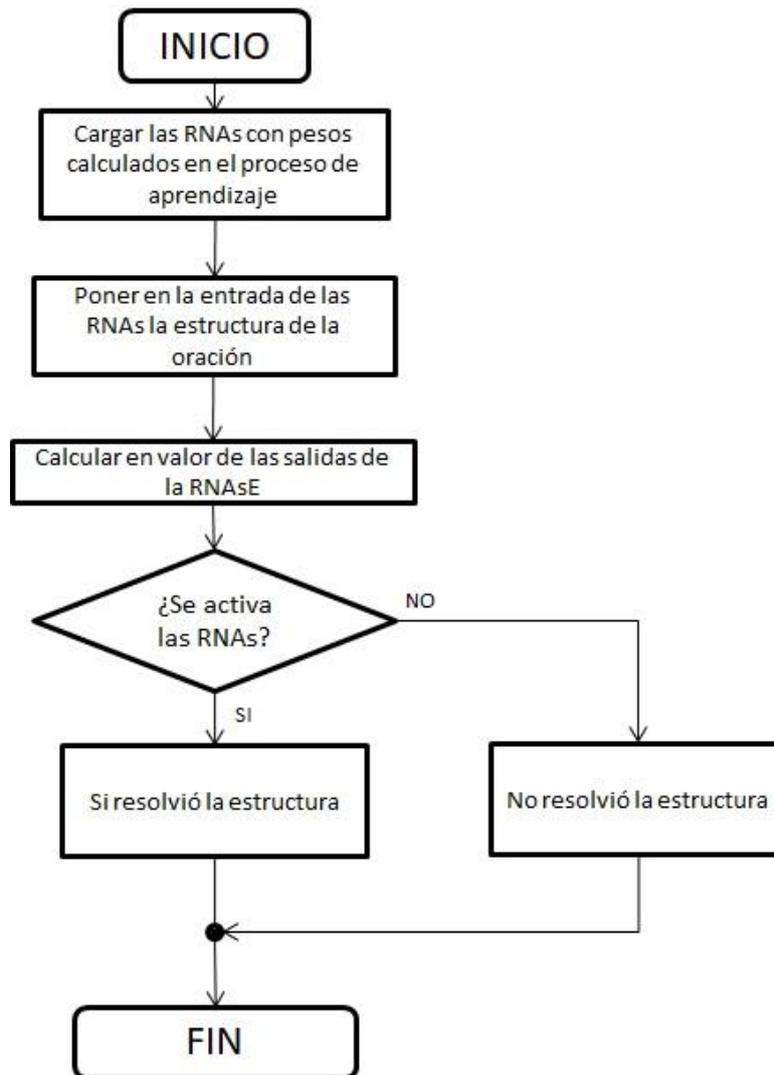


Figura N 13. Algoritmo de operación de la RNA.

5.2. Implementación.

5.2.1. Implementación de la red neuronal artificial.

La implementación de la RNA se realizó utilizando el gestor de bases de datos MariaDB y lenguaje de programación php que corre sobre plataforma Linux.

De acuerdo al diseño de la RNA del punto 5.1, se requiere crear RNAs con la cantidad de entradas, capas y salidas configurables de acuerdo a las necesidades para realizar las pruebas.

Para almacenar las RNAs se utiliza la base de datos que se muestra en la figura 14, en la tabla rna se almacenan las RNAs necesarias para el sistema, a cada RNA le corresponde

una estructura en la tabla estructura_rna en donde se especifica la cantidad de capas en el primer registro y en los siguientes registros la cantidad de neuronas por capa.

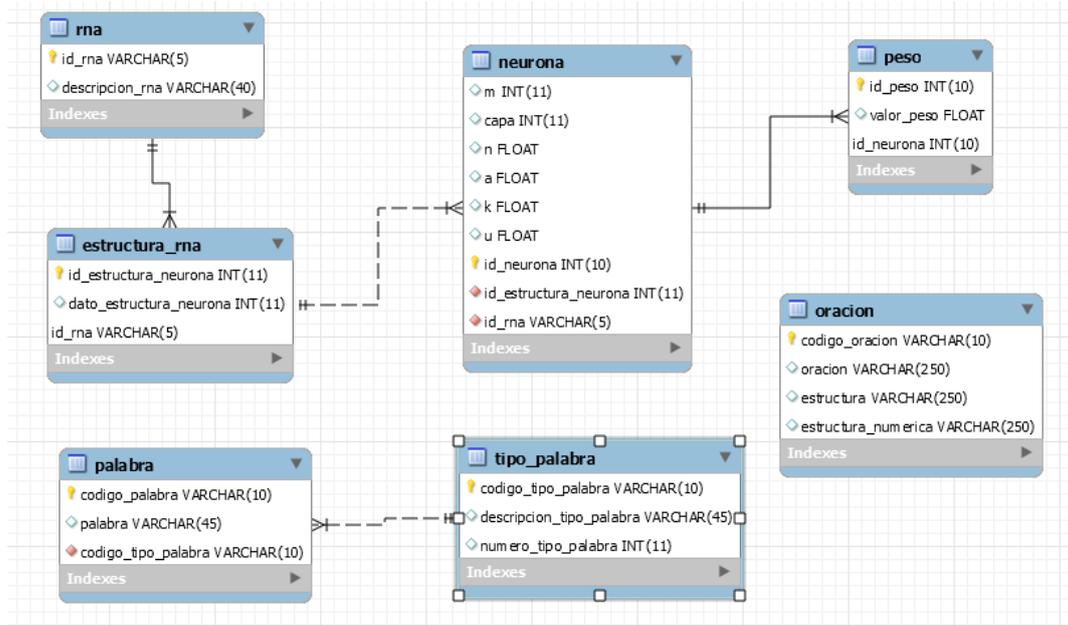


Figura 14. Base de datos.

Los datos de las estructuras de la RNAs que se almacenan en la tabla estructura_rna, que son generados con el programa del Anexo A, como se muestran en la figura 15, para la RNA con id_rna 00000, en el primero indica la cantidad de capas de la RNA y en los siguientes registros indica la cantidad de neuronas en cada capa.

	id_estructura_neurona	dato_estructura_neurona	id_rna
<input type="checkbox"/>	1	5	00000
<input type="checkbox"/>	2	5	00000
<input type="checkbox"/>	3	10	00000
<input type="checkbox"/>	4	8	00000
<input type="checkbox"/>	5	4	00000
<input type="checkbox"/>	6	1	00000
<input type="checkbox"/>	1	5	00001
<input type="checkbox"/>	2	5	00001
<input type="checkbox"/>	3	10	00001
<input type="checkbox"/>	4	8	00001
<input type="checkbox"/>	5	4	00001
<input type="checkbox"/>	6	1	00001

Figura 15. Estructura de la RNA

Utilizando la información de la estructura de la RNA que se muestra en la figura 15, se genera las neuronas en la tabla neurona y los pesos de cada neurona en la tabla pesos con el programa del Anexo B, como se presenta en la figura 16 los datos de las neuronas

y en la figura 17 los datos de los pesos de las entradas de cada neurona generados con números aleatorios.

cada neurona contiene la cantidad de entradas m , la capa a la cual pertenece $capa$, el valor de la sumatoria n , el estado de la salida de la neurona a , el coeficiente de activación lineal k , el umbral de activación con valor aleatorio u , el identificador de la neurona $id_neurona$ y el identificador de RNA id_rna , presentado en la figura 16.

	m	capa	n	a	k	u	id_neurona	id_rna
<input type="checkbox"/>	20	1	0	0	1	1721	1	00000
<input type="checkbox"/>	20	1	0	0	1	1183	2	00000
<input type="checkbox"/>	20	1	0	0	1	807	3	00000
<input type="checkbox"/>	20	1	0	0	1	1144	4	00000
<input type="checkbox"/>	20	1	0	0	1	1324	5	00000
<input type="checkbox"/>	5	2	0	0	1	1778	6	00000
<input type="checkbox"/>	5	2	0	0	1	1074	7	00000
<input type="checkbox"/>	5	2	0	0	1	1123	8	00000
<input type="checkbox"/>	5	2	0	0	1	1020	9	00000
<input type="checkbox"/>	5	2	0	0	1	1522	10	00000
<input type="checkbox"/>	5	2	0	0	1	812	11	00000
<input type="checkbox"/>	5	2	0	0	1	1056	12	00000
<input type="checkbox"/>	5	2	0	0	1	925	13	00000

Figura 16. Datos de la RNA.

Cada registro de la tabla pesos tiene el identificador de la RNA ide_rna , el identificador de neurona $id_neurona$, el identificador de peso id_peso y el valor del peso inicializado con un valor aleatorio $valor_peso$.

	id_peso	id_neurona	valor_peso	id_rna
<input type="checkbox"/>	1	1	10.2005	00000
<input type="checkbox"/>	1	1	8.96121	00001
<input type="checkbox"/>	1	1	7.65459	00002
<input type="checkbox"/>	1	1	10.6587	00003
<input type="checkbox"/>	1	1	7.9887	00004
<input type="checkbox"/>	1	1	8.55325	00005
<input type="checkbox"/>	1	1	9.26419	00006
<input type="checkbox"/>	1	1	7.57467	00007
<input type="checkbox"/>	1	1	9.15361	00008
<input type="checkbox"/>	1	1	9.98684	00009
<input type="checkbox"/>	1	1	9.17523	00010

Figura 17. Pesos de las entradas de las neuronas.

La cantidad exacta de neuronas en las capas, la cantidad de capas y cantidad de entradas para cada problema en específicamente se determina experimentalmente durante el

proceso de aprendizaje. Primeramente, se implementa las RNAs con una estructura inicial que no muy grande para reducir el tiempo de aprendizaje y si no converge el aprendizaje se aumenta el tamaño de la RNA, el tamaño mínimo recomendable es una RNA de tres capas con 5, 10, 1 neuronas correspondientemente en cada capa.

La cantidad de entradas depende de problema, en este caso específico de la cantidad de letras de la oración, pero el vector de entrada debe ser fijo para la RNA por lo que se debe asignar un vector de entrada con un valor que pueda albergar a la oración con mayor longitud.

5.2.2. Implementación del algoritmo de aprendizaje

La implementación del algoritmo de aprendizaje del Anexo C de la RNA se desarrolló en php, utilizando el diseño del punto 5.1.2, este programa lee los datos de la RNAs desde la base de datos, realiza el proceso de aprendizaje y actualiza los pesos obtenidos en la base de datos.

En las figuras 18, 19, 20 y 21 se presenta la evolución de aprendizaje de las RNAs para las palabras ANIMAL, EL, CELULAR y COSA, en donde podemos ver que converge hacia el umbral de cada RNA asignada a la palabra. En el caso de las palabras se usa una red para cada palabra del vocabulario.

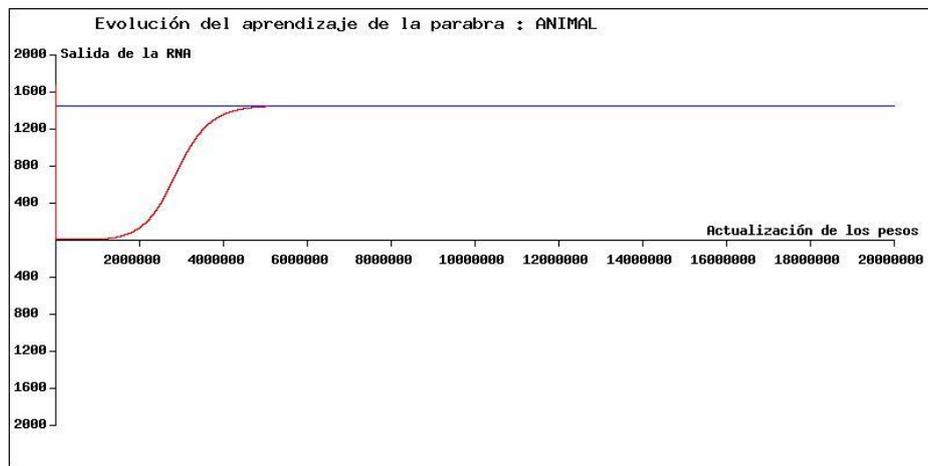


Figura 18. Evolución de aprendizaje de la palabra ANIMAL.

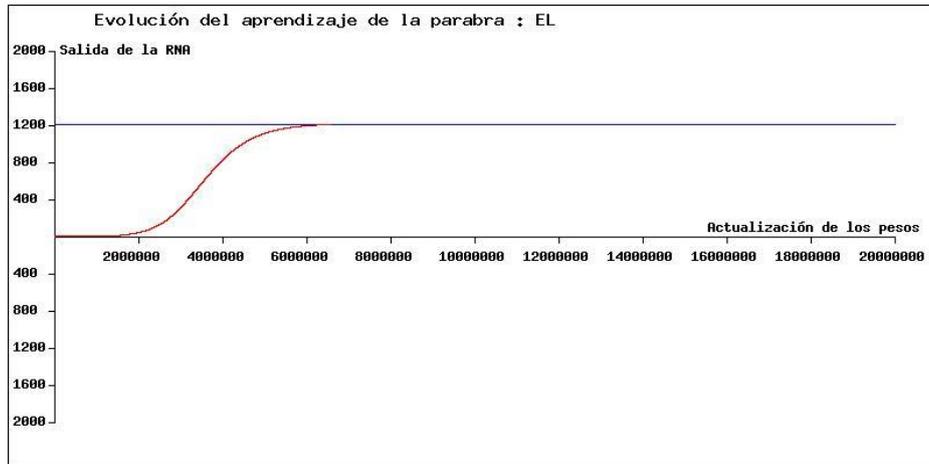


Figura 19. Evolución de aprendizaje de la palabra EL.

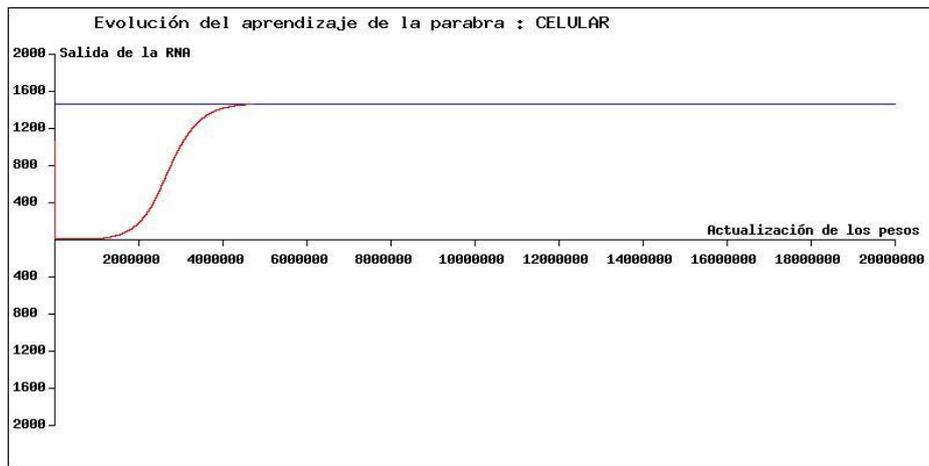


Figura 20. Evolución de aprendizaje de la palabra CELULAR.

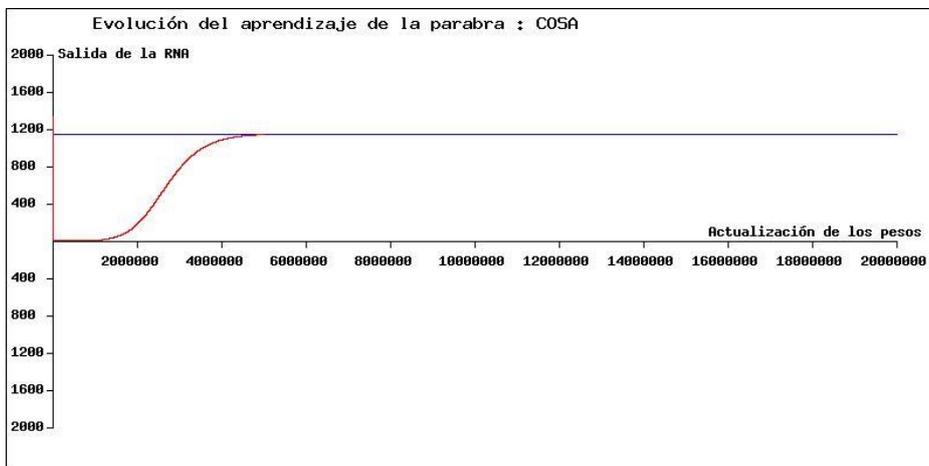


Figura 21. Evolución de aprendizaje de la palabra COSA.

El programa de aprendizaje se corre en el entorno de Linux como se muestra en la figura 22, en donde se muestra la RNA asignada a la palabra, el umbral de la RNA, la palabra,

el vector de las letras de la palabra en código ASCII, el valor inicial de la salida de la RNA, el proceso de aprendizaje en donde se muestra el número de repeticiones de actualización de todos los pesos de la RNA, cantidad de actualizaciones de pesos y el valor de la salida y el valor final de la salida de la RNA que muestra que alcanza el umbral de la RNA.

```

RNA : 00034
Umbral : 1002
Palabra : BONITO
Vector : 66 79 78 73 84 79 65 65 82 0 0 0 0 0 0 0 0 0 0 0 0

Valor inicial de la salida de la RNA : 2.6166271109058E-11

      No Rep.      Can. Actual.      Salida RNA
      3760 :      1000000 :      6.370655
      7519 :      2000000 :      175.358675
     11279 :      3000000 :      664.754231
     15038 :      4000000 :      938.274528
     18797 :      5000000 :      999.418504
Valor final de la salida de la RNA : 1002.0008240498

RNA : 00035
Umbral : 1535
Palabra : PEZ
Vector : 80 69 90 73 84 79 65 65 82 0 0 0 0 0 0 0 0 0 0 0 0

Valor inicial de la salida de la RNA : 4.5321178797179E-11

      No Rep.      Can. Actual.      Salida RNA
      3760 :      1000000 :      6.754047
      7519 :      2000000 :      209.991688
     11279 :      3000000 :      1151.137440
     15038 :      4000000 :      1507.955823
Valor final de la salida de la RNA : 1535.0006707656

```

Figura 22. Ejecución del programa de aprendizaje.

5.2.3. Interface de usuario

Se desarrolló la interface web del usuario para ejecutar diversas fases de inicialización del sistema que se muestra en la figura 23. En donde se puede generar estructuras de RNAs como se muestra en la figura 24 la generación de diez estructuras de RNAs adicionales de cinco capas, con 20, 30, 20, 10 y 1 neuronas correspondientemente. En la figura 25 se presenta la interface de registro de oraciones que serán procesadas posteriormente.

PROCESAMIENTO DE LENGUAJE NATURAL (PLN)	
Genera Estructuras RNAs Tipo palabra Oración Genera palabras Asigna tipos de palabras Genera estructuras de Oraciones Genera RNA Aprendizaje Reconocimiento Semántica SemánticaM Semántica 2.0 palabras Semántica 2.0 oraciones	Contenido Lista contenido
Procesamiento de lenguaje natural	

Figura 23. Interface del usuario.

Estructuras de RNAs		
No	Código	Estructura (Numero de capas Numero de neuronas en la capa 1 Numero de neuronas en la capa N)
1	00060	5 20 30 20 10 1
2	00061	5 20 30 20 10 1
3	00062	5 20 30 20 10 1
4	00063	5 20 30 20 10 1
5	00064	5 20 30 20 10 1
6	00065	5 20 30 20 10 1
7	00066	5 20 30 20 10 1
8	00067	5 20 30 20 10 1
9	00068	5 20 30 20 10 1
10	00069	5 20 30 20 10 1

Figura 24. Generación de diez neuronas adicionales.

Cada oración registrada se separa en palabras como se muestra en la figura 26, que se usan para realizar el proceso de aprendizaje.

Registro de Oraciones		
Oración :	<input type="text" value="Carga el word"/>	
<input type="button" value="Grabar"/>		<input type="button" value="Borrar"/>
Oraciones		
	Código	Oración
	000000000	EL PERRO ES UN ANIMAL
	000000001	EL GATO ES UN ANIMAL
	000000002	EL LORO ES UN ANIMAL
	000000003	EL ELEFANTE ES UN ANIMAL
	000000004	LA VACA ES UN ANIMAL

Figura 25. Interface de registro de oraciones.

Lista de palabras		
Codigo	Palabra	Tipo palabra
000000000	EL	
000000001	PERRO	
000000002	ES	
000000003	UN	
000000004	ANIMAL	
000000005	GATO	
000000006	LORO	
000000007	ELEFANTE	

Figura 26. Generación de palabras.

5.3. Pruebas de operación

Las pruebas se realizaron para 32 oraciones que se muestran en la tabla 1, registrados en la base de datos con sus correspondientes códigos.

Tabla 1. Oraciones para las pruebas.

Oraciones		
	Código	Oración

	000000000	EL PERRO ES UN ANIMAL
	000000001	EL GATO ES UN ANIMAL
	000000002	EL LORO ES UN ANIMAL
	000000003	EL ELEFANTE ES UN ANIMAL
	000000004	LA VACA ES UN ANIMAL
	000000005	EL TORO ES UN ANIMAL
	000000006	EL CABALLO ES UN ANIMAL
	000000007	LA YEGUA ES UN ANIMAL
	000000008	EL RATON ES UN ANIMAL
	000000009	EL GALLO ES UN ANIMAL
	000000010	LA MESA ES UNA COSA
	000000011	LA SILLA ES UNA COSA
	000000012	EL CELULAR ES UNA COSA
	000000013	EL CARRO ES UNA COSA
	000000014	EL LAPIZ ES UNA COSA
	000000015	EL CUADERNO ES UNA COSA
	000000016	EL TELEVISOR ES UNA COSA
	000000017	EL BUS ES UNA COSA
	000000018	EL TREN ES UNA COSA
	000000019	EL AUTO ES UNA COSA

	0000000020	LA VENTANA ES UNA COSA
	0000000021	EL PARLANTE ES UNA COSA
	0000000022	EL TIGRE ES UN ANIMAL
	0000000023	EL EDIFICIO ES UNA COSA
	0000000024	LA ALMOHADA ES UNA COSA
	0000000025	EL PERRO ES UNA COSA
	0000000026	LA BALLENA ES UN ANIMAL
	0000000027	EL FOCO ES UNA COSA
	0000000028	EL BONITO ES UN PEZ
	0000000029	EL PERRO ES UNA COSA
	0000000030	CUALES SON LOS ANIMALES
	0000000031	EL PERRO ES ROJO

Procesamiento de lenguaje natural

Las oraciones se separan en palabras y se crea la base de datos de palabras que se muestran en la tabla 2, y hacen un total de 41 palabras.

Tabla 2. Lista de palabras

Lista de palabras	
Codigo	Palabra
0000000000	EL
0000000001	PERRO
0000000002	ES
0000000003	UN

0000000004	ANIMAL
0000000005	GATO
0000000006	LORO
0000000007	ELEFANTE
0000000008	LA
0000000009	VACA
0000000010	TORO
0000000011	CABALLO
0000000012	YEGUA
0000000013	RATON
0000000014	GALLO
0000000015	MESA
0000000016	UNA
0000000017	COSA
0000000018	SILLA
0000000019	CELULAR
0000000020	CARRO
0000000021	LAPIZ
0000000022	CUADERNO
0000000023	TELEVISOR
0000000024	BUS
0000000025	TREN
0000000026	AUTO
0000000027	VENTANA
0000000028	PARLANTE
0000000029	TIGRE
0000000030	EDIFICIO
0000000031	ALMOHADA

000000032	BALLENA
000000033	FOCO
000000034	BONITO
000000035	PEZ
000000036	CUALES
000000037	SON
000000038	LOS
000000039	ANIMALES
000000040	ROJO

Para memorizar las palabras se asignan 41 RNAs de 5 capas con 5, 10, 8, 4, 1 neuronas en cada capa correspondientemente, haciendo un total de 28 neuronas cada RNA y en las 41 RNAs son 1148 neuronas con un total de 10906 pesos, el vector de entradas tiene hasta 20 elementos para almacenar las palabras.

Se realizó el proceso de aprendizaje de las 41 palabras utilizando las redes asignadas y se muestra en las figuras 27, 28 y 29 ejemplos de convergencia del aprendizaje para las palabras CUADERNO, GATO y BUS, en donde se puede ver que la curva de aprendizaje converge al valor del umbral.

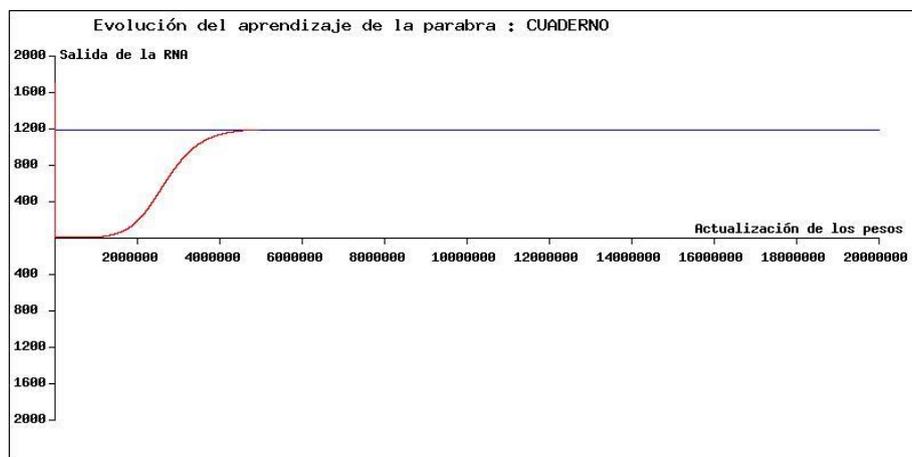


Figura 27. Evolución del aprendizaje de la palabra CUADERNO.

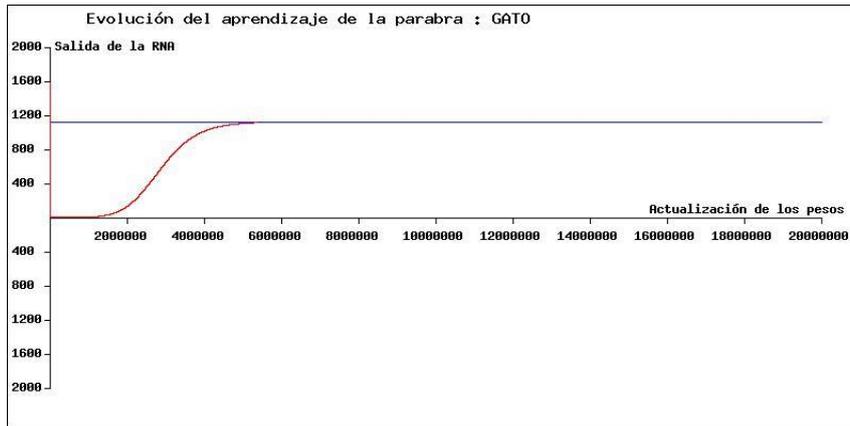


Figura 28. Evolución del aprendizaje de la palabra GATO.

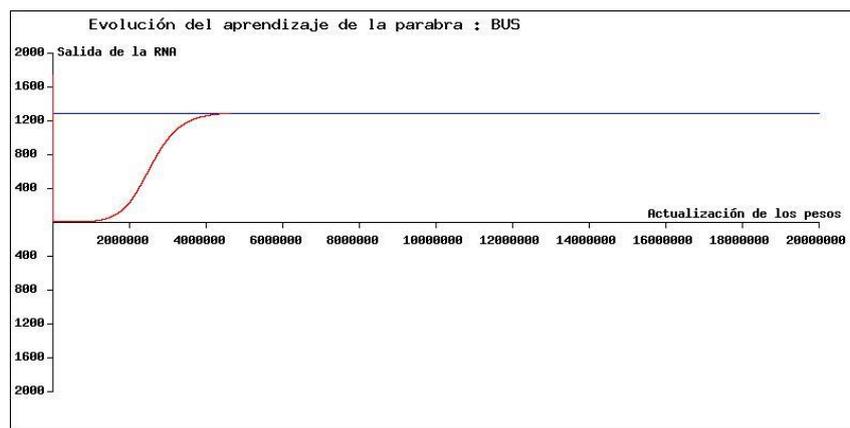


Figura 29. Evolución del aprendizaje de la palabra BUS.

Se realizaron las pruebas de comprobación si las redes neuronales que pasaron el proceso de aprendizaje reconocen las palabras que aprendieron y se muestra los resultados en la tabla 3.

Tabla 3. Resultado de reconocimiento de todas las palabras de prueba.

Reconocimiento de palabras con la Red Neuronal Artificial		
1	¿Conoce la palabra EI? Valor RNA : 1499.003680074	Si
2	¿Conoce la palabra Perro? Valor RNA : 1206.0015351753	Si
3	¿Conoce la palabra Es? Valor RNA : 1700.0014755025	Si

4	¿Conoce la palabra Un? Valor RNA : 1479.9963927218	Si
5	¿Conoce la palabra Animal? Valor RNA : 1585.9973154283	Si
6	¿Conoce la palabra Gato? Valor RNA : 1114.0032303329	Si
7	¿Conoce la palabra Loro? Valor RNA : 1293.9984810584	Si
8	¿Conoce la palabra Elefante? Valor RNA : 1162.9972248629	Si
9	¿Conoce la palabra La? Valor RNA : 1132.00129935	Si
10	¿Conoce la palabra Vaca? Valor RNA : 1594.0031443358	Si
11	¿Conoce la palabra Toro? Valor RNA : 1664.0008533281	Si
12	¿Conoce la palabra Caballo? Valor RNA : 1352.9997416767	Si
13	¿Conoce la palabra Yegua? Valor RNA : 1308.0014628631	Si
14	¿Conoce la palabra Raton? Valor RNA : 1302.0047683758	Si
15	¿Conoce la palabra Gallo? Valor RNA : 1301.000426196	Si
16	¿Conoce la palabra Mesa? Valor RNA : 1267.9992065731	Si
17	¿Conoce la palabra Una? Valor RNA : 1661.0047774979	Si
18	¿Conoce la palabra Cosa? Valor RNA : 953.00148755809	Si
19	¿Conoce la palabra Silla? Valor RNA : 1299.999669314	Si
20	¿Conoce la palabra Celular? Valor RNA : 915.00003449625	Si
21	¿Conoce la palabra Carro? Valor RNA : 1138.0005870213	Si

22	¿Conoce la palabra Lapiz? Valor RNA : 1687.9990558824	Si
23	¿Conoce la palabra Cuaderno? Valor RNA : 1187.0004844192	Si
24	¿Conoce la palabra Televisor? Valor RNA : 1726.0017222143	Si
25	¿Conoce la palabra Bus? Valor RNA : 1284.002472859	Si
26	¿Conoce la palabra Tren? Valor RNA : 1075.0002555822	Si
27	¿Conoce la palabra Auto? Valor RNA : 1263.9971611294	Si
28	¿Conoce la palabra Ventana? Valor RNA : 1293.9989233243	Si
29	¿Conoce la palabra Parlante? Valor RNA : 1351.0017828776	Si
30	¿Conoce la palabra Tigre? Valor RNA : 971.00185064534	Si
31	¿Conoce la palabra Edificio? Valor RNA : 1507.9974450149	Si
32	¿Conoce la palabra Almohada? Valor RNA : 1114.0036866805	Si
33	¿Conoce la palabra Ballena? Valor RNA : 1489.0028826833	Si
34	¿Conoce la palabra Foco? Valor RNA : 1235.0003046704	Si
35	¿Conoce la palabra Bonito? Valor RNA : 1214.0014821633	Si
36	¿Conoce la palabra Pez? Valor RNA : 1488.0009442356	Si
37	¿Conoce la palabra Cuales? Valor RNA : 1488.9974367873	Si
38	¿Conoce la palabra Son? Valor RNA : 1469.0011335783	Si
39	¿Conoce la palabra Los? Valor RNA : 1241.0028195735	Si

40	¿Conoce la palabra Animales? Valor RNA : 1622.0012355267	Si
41	¿Conoce la palabra Rojo? Valor RNA : 1023.9988462876	Si

Se realizan pruebas para verificar si las RNAs solo se activan con la palabra a la que fueron asignadas como se muestra en la tabla 4, la prueba para la RNA 7 asignada a la palabra Loro, y este sí reconoce la palabra Loro, pero el resto de palabras no las reconoce.

Tabla 4. Prueba de la RNA 7, asignada a la palabra Loro.

Reconocimiento de palabras con la Red Neuronal Artificial		
1	¿Conoce la palabra El? Valor RNA : 416.05899080668	No
2	¿Conoce la palabra Perro? Valor RNA : 1104.7950165964	No
3	¿Conoce la palabra Es? Valor RNA : 1114.7816178002	No
4	¿Conoce la palabra Un? Valor RNA : 1145.1505883702	No
5	¿Conoce la palabra Animal? Valor RNA : 1227.010000643	No
6	¿Conoce la palabra Gato? Valor RNA : 1245.0984809665	No
7	¿Conoce la palabra Loro? Valor RNA : 1293.9984810584	Si
8	¿Conoce la palabra Elefante? Valor RNA : 1205.4671431661	No
9	¿Conoce la palabra La? Valor RNA : 1193.0020716108	No
10	¿Conoce la palabra Vaca? Valor RNA : 1200.4376291859	No

11	¿Conoce la palabra Toro? Valor RNA : 1322.1321810722	No
12	¿Conoce la palabra Caballo? Valor RNA : 1164.8146331191	No
13	¿Conoce la palabra Yegua? Valor RNA : 1285.5112334453	No
14	¿Conoce la palabra Raton? Valor RNA : 1307.2980217859	No
15	¿Conoce la palabra Gallo? Valor RNA : 1245.7470385696	No
16	¿Conoce la palabra Mesa? Valor RNA : 1263.3923321102	No
17	¿Conoce la palabra Una? Valor RNA : 1257.5844598305	No
18	¿Conoce la palabra Cosa? Valor RNA : 1264.5015452187	No
19	¿Conoce la palabra Silla? Valor RNA : 1269.1995401651	No
20	¿Conoce la palabra Celular? Valor RNA : 1234.353769224	No
21	¿Conoce la palabra Carro? Valor RNA : 1239.3106961417	No
22	¿Conoce la palabra Lapiz? Valor RNA : 1258.6934638234	No
23	¿Conoce la palabra Cuaderno? Valor RNA : 1229.0295610015	No
24	¿Conoce la palabra Televisor? Valor RNA : 1281.9433517375	No
25	¿Conoce la palabra Bus? Valor RNA : 1299.3278479395	No
26	¿Conoce la palabra Tren? Valor RNA : 1324.99793477	No
27	¿Conoce la palabra Auto? Valor RNA : 1328.7706597014	No
28	¿Conoce la palabra Ventana? Valor RNA : 1301.0276337619	No

29	¿Conoce la palabra Parlante? Valor RNA : 1261.2290239972	No
30	¿Conoce la palabra Tigre? Valor RNA : 1289.6358211057	No
31	¿Conoce la palabra Edificio? Valor RNA : 1182.5511799433	No
32	¿Conoce la palabra Almohada? Valor RNA : 1225.0856836811	No
33	¿Conoce la palabra Ballena? Valor RNA : 1213.2864239882	No
34	¿Conoce la palabra Foco? Valor RNA : 1246.9092273648	No
35	¿Conoce la palabra Bonito? Valor RNA : 1290.3717183081	No
36	¿Conoce la palabra Pez? Valor RNA : 1336.9055402728	No
37	¿Conoce la palabra Cuales? Valor RNA : 1255.2319432174	No
38	¿Conoce la palabra Son? Valor RNA : 1322.1100309577	No
39	¿Conoce la palabra Los? Valor RNA : 1317.59726511	No
40	¿Conoce la palabra Animales? Valor RNA : 1227.04930558	No
41	¿Conoce la palabra Rojo? Valor RNA : 1286.6397930748	No

Se realizó el aprendizaje de las 10 primeras oraciones en las redes neuronales del segundo nivel destinadas a interpretar las oraciones, las RNAs destinadas a esta función son desde la 00050 hacia arriba. En la figura 30 se muestra la ejecución del proceso de aprendizaje para las oraciones "El raton es un animal" y "", en donde se puede ver que el valor inicial de la RNA es $2.646118477768E-9$ y $1.8620317107718E-9$ correspondientemente y finaliza el aprendizaje con los valores de 1693.0000411402 y 1054.0000352503 que significa que alcanzo el valor del umbral de cada una de la RNAs asignadas.

```

Oracion : El raton es un animal
RNA : 00058
Valor inicial RNA : 2.646118477768E-9
4630 : 1000000 : 1.197606
9260 : 2000000 : 30.515563
13889 : 3000000 : 214.117827
18519 : 4000000 : 853.566229
23149 : 5000000 : 1490.980772
27778 : 6000000 : 1665.301121
Actualizando la Neuron : 15 Interacion : 0 Valor Red : 1693.0000411402 Valor suma : 0
Oracion : El gallo es un animal
RNA : 00059
Valor inicial RNA : 1.8620317107718E-9
4630 : 1000000 : 1.119379
9260 : 2000000 : 28.803972
13889 : 3000000 : 191.385524
18519 : 4000000 : 541.112397
23149 : 5000000 : 849.376361
27778 : 6000000 : 993.321138
32408 : 7000000 : 1042.618673
Actualizando la Neuron : 24 Interacion : 0 Valor Red : 1054.0000352503 Valor suma : 0
Cantidad de iteraciones : 1

```

Figura 30. Ejecución del aprendizaje de las oraciones “El raton es un animal” y “El gallo es un animal”.

En las figuras 31 y 32 se presenta la evolución del aprendizaje de las oraciones en donde podemos observar que en los dos casos converge hacia el valor del umbral de cada RNA asignada a la oración.

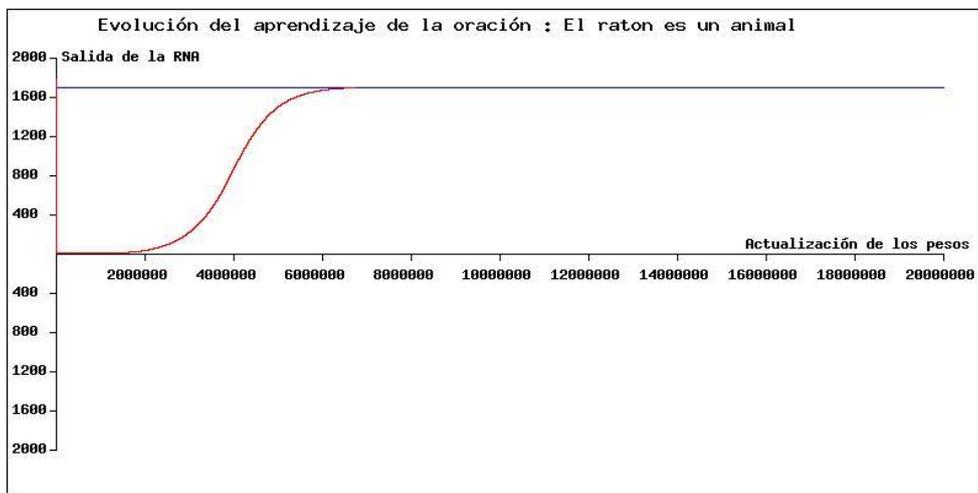


Figura 31. Evolución del aprendizaje de la oración “El raton es un animal”.

En la tabla 5 se muestra la prueba en donde se muestra que interpreta las primeras diez oraciones para los cuales fue realizado un proceso de aprendizaje.

En la tabla 6 se presenta la prueba de interpretación de la sexta oración, esta prueba demuestra que la RNA asignada solo resuelve la oración correspondiente y el resto de oraciones no se activan ya que fueron entrenadas para otras oraciones.

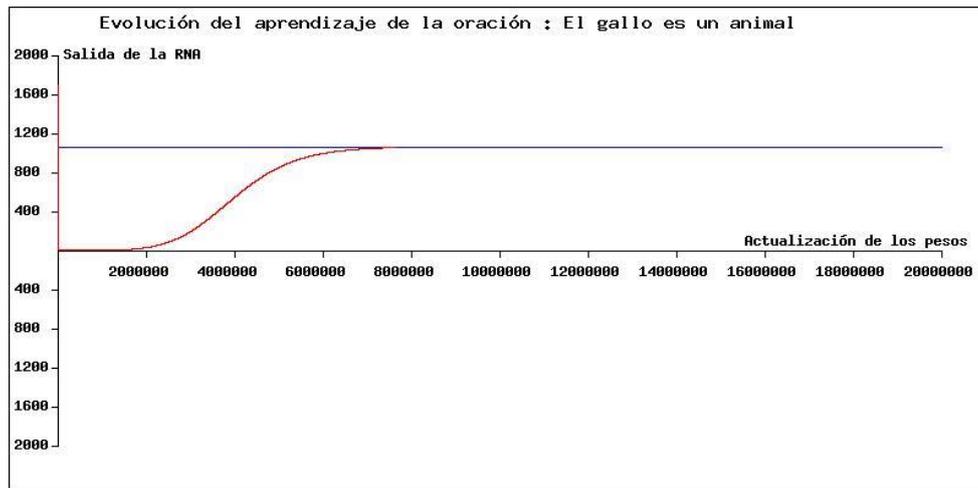


Figura 32. Evolución del aprendizaje de la oración “El gallo es un animal”.

Tabla 5. Prueba de interpretación de las 10 primeras oraciones.

Reconocimiento de oraciones con la Red Neuronal Artificial	
Consulta : ¿EL PERRO ES UN ANIMAL? Valor RNA : 1256.0014010369	Si
Consulta : ¿EL GATO ES UN ANIMAL? Valor RNA : 1648.002086431	Si
Consulta : ¿EL LORO ES UN ANIMAL? Valor RNA : 1295.9980625639	Si
Consulta : ¿EL ELEFANTE ES UN ANIMAL? Valor RNA : 802.00203753756	Si
Consulta : ¿LA VACA ES UN ANIMAL? Valor RNA : 1180.9975299889	Si
Consulta : ¿EL TORO ES UN ANIMAL? Valor RNA : 1389.0030588433	Si
Consulta : ¿EL CABALLO ES UN ANIMAL? Valor RNA : 1518.0018510538	Si
Consulta : ¿LA YEGUA ES UN ANIMAL? Valor RNA : 1100.9997524152	Si
Consulta : ¿EL RATON ES UN ANIMAL? Valor RNA : 1674.9977634977	Si
Consulta : ¿EL GALLO ES UN ANIMAL? Valor RNA : 1404.9986908843	Si

Consulta : ¿LA MESA ES UNA COSA? Valor RNA : 5.7790166487734E-10	No
Consulta : ¿LA SILLA ES UNA COSA? Valor RNA : 1.3035995444403E-10	No
Consulta : ¿EL CELULAR ES UNA COSA? Valor RNA : 1.1243323608904E-9	No
Consulta : ¿EL CARRO ES UNA COSA? Valor RNA : 7.3887656759499E-10	No
Consulta : ¿EL LAPIZ ES UNA COSA? Valor RNA : 7.4446591281938E-10	No
Consulta : ¿EL CUADERNO ES UNA COSA? Valor RNA : 2.9368423976519E-10	No
Consulta : ¿EL TELEVISOR ES UNA COSA? Valor RNA : 1.0669878507488E-9	No
Consulta : ¿EL BUS ES UNA COSA? Valor RNA : 9.1170479127435E-10	No
Consulta : ¿EL TREN ES UNA COSA? Valor RNA : 6.3403997379129E-10	No
Consulta : ¿EL AUTO ES UNA COSA? Valor RNA : 5.2179204331535E-10	No
Consulta : ¿LA VENTANA ES UNA COSA? Valor RNA : 5.7891432828645E-10	No
Consulta : ¿EL PARLANTE ES UNA COSA? Valor RNA : 1.163047320135E-9	No
Consulta : ¿EL TIGRE ES UN ANIMAL? Valor RNA : 9.1122490855274E-10	No
Consulta : ¿EL EDIFICIO ES UNA COSA? Valor RNA : 6.8080509992394E-10	No
Consulta : ¿LA ALMOHADA ES UNA COSA? Valor RNA : 4.3009101744877E-10	No
Consulta : ¿EL PERRO ES UNA COSA? Valor RNA : 5.3003582801896E-10	No
Consulta : ¿LA BALLENA ES UN ANIMAL? Valor RNA : 9.7634618924057E-10	No
Consulta : ¿EL FOCO ES UNA COSA? Valor RNA : 4.6858071702456E-10	No

Consulta : ¿EL BONITO ES UN PEZ? Valor RNA : 1.4937284855234E-9	No
--	----

Tabla 6. Prueba de interpretación de la sexta oración.

Reconocimiento de oraciones con la Red Neuronal Artificial	
Consulta : ¿EL PERRO ES UN ANIMAL? Valor RNA : 1581.4088407873	No
Consulta : ¿EL GATO ES UN ANIMAL? Valor RNA : 1428.6747424907	No
Consulta : ¿EL LORO ES UN ANIMAL? Valor RNA : 1561.1397374613	No
Consulta : ¿EL ELEFANTE ES UN ANIMAL? Valor RNA : 1480.779147509	No
Consulta : ¿LA VACA ES UN ANIMAL? Valor RNA : 1404.7609223848	No
Consulta : ¿EL TORO ES UN ANIMAL? Valor RNA : 1389.0030588433	Si
Consulta : ¿EL CABALLO ES UN ANIMAL? Valor RNA : 1445.0256548392	No
Consulta : ¿LA YEGUA ES UN ANIMAL? Valor RNA : 1561.4132116589	No
Consulta : ¿EL RATON ES UN ANIMAL? Valor RNA : 1608.3840786709	No
Consulta : ¿EL GALLO ES UN ANIMAL? Valor RNA : 1588.2280000847	No
Consulta : ¿LA MESA ES UNA COSA? Valor RNA : 915.64735483271	No
Consulta : ¿LA SILLA ES UNA COSA? Valor RNA : 663.71520996114	No
Consulta : ¿EL CELULAR ES UNA COSA? Valor RNA : 927.3545030471	No

Consulta : ¿EL CARRO ES UNA COSA? Valor RNA : 618.49556426182	No
Consulta : ¿EL LAPIZ ES UNA COSA? Valor RNA : 767.91427019283	No
Consulta : ¿EL CUADERNO ES UNA COSA? Valor RNA : 586.62258765627	No
Consulta : ¿EL TELEVISOR ES UNA COSA? Valor RNA : 871.85935583672	No
Consulta : ¿EL BUS ES UNA COSA? Valor RNA : 653.94765762569	No
Consulta : ¿EL TREN ES UNA COSA? Valor RNA : 789.50199548241	No
Consulta : ¿EL AUTO ES UNA COSA? Valor RNA : 568.72700386473	No
Consulta : ¿LA VENTANA ES UNA COSA? Valor RNA : 694.98538795477	No
Consulta : ¿EL PARLANTE ES UNA COSA? Valor RNA : 923.81306120204	No
Consulta : ¿EL TIGRE ES UN ANIMAL? Valor RNA : 1256.6510886125	No
Consulta : ¿EL EDIFICIO ES UNA COSA? Valor RNA : 704.62041594066	No
Consulta : ¿LA ALMOHADA ES UNA COSA? Valor RNA : 844.85619284472	No
Consulta : ¿EL PERRO ES UNA COSA? Valor RNA : 904.07140666148	No
Consulta : ¿LA BALLENA ES UN ANIMAL? Valor RNA : 1337.0591032829	No
Consulta : ¿EL FOCO ES UNA COSA? Valor RNA : 665.96595495095	No
Consulta : ¿EL BONITO ES UN PEZ? Valor RNA : 1057.3673837688	No
Consulta : ¿CUALES SON LOS ANIMALES? Valor RNA : 894.99381516009	No

Las pruebas realizadas para 30 oraciones en idioma español se necesitan una RNA organizado en dos niveles, en la primera con 41 sub RNAs para el reconocimiento de las palabras y el segundo nivel con 30 sub RNAs para la interpretación de las oraciones, haciendo un total de 71 sub RNAs interconectadas de 5 capas con 5, 10, 8, 4, 1 neuronas en cada capa correspondientemente, haciendo un total de 28 neuronas cada sub RNA y en las 71 sub RNAs son 1988 neuronas con un total de 1886 pesos, el vector de entradas tiene hasta 20 elementos para almacenar las palabras.

Considerando la estructura de la RNA descrita para implementar el análisis semántico el indicador Complejidad del analizador semántico tendría un valor de 100%.

Indicadores de la hipótesis general:

Indicador : Complejidad del analizador semántico.

Índice : Número de neuronas en la red (0% - 100%).

Según los resultados el sistema proporcionó el 100% de respuestas y la comprensión de las oraciones representa el 100% para las oraciones utilizadas en el proceso de aprendizaje.

Indicador de la hipótesis secundaria:

Indicador 1 : Cantidad de respuestas del sistema para hacer la comprensión.

Índice 1 : Porcentaje de respuestas correctas (0 – 100%).

Indicador 2 : Precisión en la comprensión de oraciones en idioma castellano.

Índice 2 : Porcentaje promedio de precisión (0 – 100%).

CAPITULO VI: DISCUSIÓN DE RESULTADOS

6.1. Contrastación y demostración de la hipótesis con los resultados

Los resultados demuestran que el conocimiento del grupo de oraciones en lenguaje natural en idioma español utilizadas, fueron posibles representar con RNAs en la computadora después de un proceso de aprendizaje, lo que demuestra la segunda hipótesis específica que dice “Las redes neuronales artificiales permitirán representar el conocimiento transmitido a través de oraciones dadas en lenguaje natural”.

Por otro lado, se prueba que comprende las oraciones haciendo la consulta con la oración al sistema y responde positivamente si está en el grupo utilizado para el proceso de aprendizaje y negativamente a cualquier otra pregunta, lo que demuestra la segunda hipótesis específica que dice “Las redes neuronales artificiales permitirán desarrollar un software que comprenda las oraciones del lenguaje natural”

La comprensión de las consultas sobre el conocimiento aprendido por la RNAs utilizadas, al responder adecuadamente, prueba que el sistema implementado con RNAs realiza el análisis semántico, demostrando la hipótesis general “Las redes neuronales artificiales permitirán desarrollar un analizador semántico de lenguaje natural para el idioma castellano”.

6.2. Contrastación de los resultados con otros estudios similares

En general los investigadores en procesamiento de lenguaje natural no han trabajado directamente en el análisis semántico y menos aún con RNAs, como se puede ver en el *A State of Art for Semantic Analysis of Natural Language Processing* de (Dastan Hussen Maulud, 2021). Los trabajos previos más bien se orientan a desarrollar soluciones específicas de chatbots, clasificación de textos, etc, sin centrarse en el proceso natural de aprendizaje y comunicación, por lo que no existen estudios similares para contrastar los resultados obtenidos.

6.3. Responsabilidad ética

En el desarrollo del presente proyecto se ha cumplido con todos los principios éticos de investigación y uso de recursos, reconociendo la autoría de fuentes bibliográficas referenciando debidamente, reportando el resultado obtenido, desarrollando

herramientas propias para las pruebas y sin usar resultados de otros autores para reportarlo con resultados de este proyecto.

CONCLUSIONES

Las RNAs multicapa de perceptrones permite representar el conocimiento contenido en las oraciones, si el proceso de aprendizaje es similar al proceso natural de aprendizaje de las personas en donde primero aprenden las palabras y luego aprenden a formular y comprender oraciones.

El proceso de aprendizaje de la RNAs similar al natural garantiza la secuencia de análisis innato de realizar primero el análisis léxico, luego el análisis semántico, siguiendo con el análisis semántico y finalmente la construcción de la respuesta.

RECOMENDACIONES

La estructura de la RNAs y el proceso de aprendizaje diseñados en esta investigación dieron buenos resultados en la implementación de un grupo de oraciones, es recomendable hacer pruebas para otros tipos de oraciones.

Con los resultados obtenidos de una solución que permite realizar el análisis semántico de forma innata es importante continuar con la investigación en la siguiente fase de construcción de la respuesta utilizando RNAs, ya que eso permitiría implementar interfaces hombre maquina en lenguaje natural en idioma español.

REFERENCIAS BIBLIOGRÁFICAS

- Baliri, S. (2014). Teoria de Lenguajes Formales, Una introduccion para Linguistica. *Barcelona: Universidad Autinoma de Barcelona.*
- Chen, X. L. (2020). English text quality analysis based on recurrent neural network and semantic segmentation. *Future Generation Computer Systems*, 507-511.
- Dastan Hussen Maulud, S. R. (2021). A State of Art for Semantic Analysis of Natural Language Processing. *Qubahan Academic Journal*, 1-8.
- Jingyun Xu, Y. C.-f. (2019). Incorporating Context-Relevant Concepts into Convolutional Neural Networks for Short Text Classification. *Neurocomputing*.
- Li, W. (2020). Analysis of Semantic Comprehension Algorithms of Natural Language Based on Robot's Questions and Answers. *IEEE Xplore*, 1021-1024.
- Mario G. Borja-Borja, S. T. (2009). Mario G. Borja-Borja, S. K.-A. *ECI ISSN Nº 1813-9401*, 77 - 81.
- Musil, T. (2020). Semantic Holism and Word Representations in Artificial Neural Networks.
- Tapsai, H. U. (2020). Thai Natural Language Processing. *Studies in Computational Intelligence*.
- Xie, Z. C. (2020). SemBRS: A Semantic Analysis based Book Retrieval Approach. *IEEE Xplore*, 101-104.

ANEXOS

Anexo A: Programa de generación de estructuras de las RNAs.

```
function generaRNAs($db){
    $basic = new datosbasicos();
    $html= "<table border=0 align=center>
        <tr>
            <td colspan=4 class=titulo>Estructuras de RNAs</td>
        </tr>";
    $html=$html."<br><br>
        <tr>
            <td class=cabCentro>No</td>
            <td class=cabIzquierda>C&ocute;digo</td>
            <td class=cabCentro>Estructura <br><span>(Numero de capas
                | Numero de neuronas en la capa 1 |
                ....|Numero de neuronas en la
                capa N)</span>
            </td>
        </tr>";
    for ($i=1; $i<=10; $i++){
        $cod_rna = $basic->generaCodigoN("estructura_neurona","id_rna",5,$db);
        $html=$html."<tr>
            <td class=cabCentro>$i</td>
            <td class=cabIzquierda>$cod_rna</td>
            <td class=cabCentro>5 | 20 | 30 | 20 | 10 | 1</td>
        </tr>";
    }
    try {
        $st1 = $db->prepare("insert into estructura_neurona id_rna,
            dato_estructura_neurona,
            id_estructura_neurona) value('$cod_rna', '5',
            '1')");
        $st2 = $db->prepare("insert into estructura_neurona id_rna,
            dato_estructura_neurona,
            id_estructura_neurona) value('$cod_rna', '5',
            '2')");
        $st3 = $db->prepare("insert into estructura_neurona
            (id_rna,dato_estructura_neurona,
            id_estructura_neurona) value('$cod_rna', '10',
            '3')");
        $st4 = $db->prepare("insert into estructura_neurona
            (id_rna,dato_estructura_neurona,
            id_estructura_neurona) value('$cod_rna', '8',
            '4')");
        $st5 = $db->prepare("insert into estructura_neurona
            (id_rna,dato_estructura_neurona,
            id_estructura_neurona) value('$cod_rna', '4',
            '5')");
        $st6 = $db->prepare("insert into estructura_neurona
            (id_rna,dato_estructura_neurona,
```

```

        id_estructura_neurona) value('$cod_rna', '1',
        '6')");
    $db->beginTransaction();
    $st1->execute();
    $st2->execute();
    $st3->execute();
    $st4->execute();
    $st5->execute();
    $st6->execute();
    $db->commit();
} catch (Exception $e) {
    $db->rollBack();
    echo "Failed: " . $e->getMessage();
}
}
$html=$html."</table>";
echo $html;
}

```

Anexo B: Programa de generación de las RNAs.

```
function generaRed($db){
    $basic = new datosbasicos();
    $capa=0;
    $numNeuro=0;
    $numEntra=0;
    try {
        $bpes = $db->prepare("delete from peso");
        $bneu = $db->prepare("delete from neurona");
        $db->beginTransaction();
        $bpes->execute();
        $bneu->execute();
        $db->commit();
    } catch (Exception $e) {
        $db->rollBack();
        echo "Failed: " . $e->getMessage();
    }
    $html= "<table border=0 align=center>
        <tr>
            <td colspan=4 class=titulo>Estructura de Red Neuronal
            Artificial</td>
        </tr>";
    try {
        $st = $db->prepare("select id_estructura_neurona,
            dato_estructura_neurona, id_rna from
            estructura_neurona");
        $st->execute();
        $html=$html."<br><br>
            <tr>
                <td class=cabCentro>Estructura</td>
                <td class=cabIzquierda>Dato</td>
            </tr>";
        while ($result = $st->fetch(PDO::FETCH_ASSOC)) {
            $idEstruc = $result['id_estructura_neurona'];
            $datoEstruct = $result['dato_estructura_neurona'];
            $idrna = $result['id_rna'];
            if ($idEstruc==1){
                $numero_capas = $datoEstruct;
                $capa=0;
                $numNeuro=0;
            }
            if ($idEstruc==2){
                print("***** RNA : ".$idrna." Capa :
                    ".$capa." *****\n");
                $capa++;
                $numEntra=20;
                for ($i=1; $i<=$datoEstruct; $i++){
```

```

$numNeuro++;
print("Neurona : ".$numNeuro."\n");
try {
    $m[$numNeuro] = $numEntra;
    $caparna[$numNeuro] = $capa;
    $n[$numNeuro] = 0;
    $a[$numNeuro] = 0;
    $k[$numNeuro] = 1;
    $u[$numNeuro] = rand(800, 1800);
    $stmt = $db->prepare("insert into neurona (id_rna,
id_neurona, m, capa, n, a, k, u) values ('$idrna',
'$numNeuro', '$numEntra', '$capa', '0', '0', '1',
'$u[$numNeuro]')");
    $db->beginTransaction();
    $stmt->execute();
    $db->commit();
    for ($j=1; $j<=$numEntra; $j++){
        try {
            $peso = rand(0, 10000) / 10000000;
            $stpeso = $db->prepare("insert into peso
(id_rna, id_peso, id_neurona, valor_peso)
values ('$idrna', '$j', '$numNeuro', '$peso')");
            $db->beginTransaction();
            $stpeso->execute();
            $db->commit();
            $pesorna[$numNeuro][$j] = $peso;
            print("        Peso : ".$j." Valor :
                ".$peso."\n");
        } catch (Exception $e) {
            $db->rollBack();
            echo "Failed: " . $e->getMessage();
        }
    }
} catch (Exception $e) {
    $db->rollBack();
    echo "Failed: " . $e->getMessage();
}
}
$numEntra=$datoEstruct;
}
if ($idEstruc>2){
    print("***** RNA : ".$idrna."        Capa :
    ".$capa." *****\n");
    $capa++;
    for ($i=1; $i<=$datoEstruct; $i++){
        $numNeuro++;
        print("Neurona : ".$numNeuro."\n");
    }
}

```

```

try {
    $m[$numNeuro] = $numEntra;
    $caparna[$numNeuro] = $capa;
    $n[$numNeuro] = 0;
    $a[$numNeuro] = 0;
    $k[$numNeuro] = 1;
    $u[$numNeuro] = rand(800, 1800);
    $stmt = $db->prepare("insert into neurona
                        (id_rna, id_neurona, m, capa,
                         n, a, k, u) values ('$idrna',
                         '$numNeuro', '$numEntra',
                         '$capa', '0', '0', '1',
                         '$u[$numNeuro]')");
    $db->beginTransaction();
    $stmt->execute();
    $db->commit();
    for ($j=1; $j<=$numEntra; $j++){
        try {
            $peso=rand(0, 10000) / 10000000;
            $stpeso = $db->prepare("insert into peso
            (id_rna, id_peso, id_neurona, valor_peso)
            values ('$idrna', '$j', '$numNeuro', '$peso')");
            $db->beginTransaction();
            $stpeso->execute();
            $db->commit();
            $pesorna[$numNeuro][$j] = $peso;
            print("          Peso : ".$j." Valor : ".$peso."\n");
        } catch (Exception $e) {
            $db->rollBack();
            echo "Failed: " . $e->getMessage();
        }
    }
} catch (Exception $e) {
    $db->rollBack();
    echo "Failed: " . $e->getMessage();
}
}
$numEntra=$datoEstruct;
}
$html=$html."<tr>
                <td class=filaCentro>".$idEstruct."</td>
                <td class=filaIzquierda>".$datoEstruct."</td>
            </tr>";
}
$html=$html."<tr>
                <td colspan=4 class=titulo> Red Neuronal
                Artificial $numNeuro

```

```

        </td>
    </tr>";
    for ($ineuro=1; $ineuro <= $numNeuro; $ineuro++){
        $html=$html."<tr>
            <td class=filacentro>Neurona ".$ineuro."</td>
            <td class=filalquierda>
                Capa : ".$caparna[$ineuro]."<br>
                Entradas : ".$m[$ineuro]."<br>
                Sumatoria : ".$n[$ineuro]."<br>
                Resultado : ".$a[$ineuro]."<br>
                Umbral : ".$u[$ineuro]."<br>
                K adaline : ".$k[$ineuro]."<br>";
                $ne = $m[$ineuro];
                for ($npes=1; $npes <= $ne; $npes++){
                    $html=$html."peso[".$npes."]:
                    ".$pesorna[$ineuro][$npes]."<br>";
                }
                $html=$html."        </td>
            </tr>";
        }
        $html=$html."</table>";
    }
    catch (Exception $e) {
        Echo "Failed: " . $e->getMessage();
    }
}

```

Anexo C: Programa de aprendizaje de las RNAs.

```
function aprendeNominativoLetras($db){
    $basic = new datosbasicos();
    $cRNA = new libRNA();
    $capa=0;
    $numNeuro=0;
    $numEntra=0;
    print "***** Aprendizaje de la RNA *****\n\n"
    try {
        $st = $db->prepare("select id_rna, id_estructura_neurona,
        dato_estructura_neurona from estructura_neurona order by id_rna,
        id_estructura_neurona");
        $st->execute();
        while ($result = $st->fetch(PDO::FETCH_ASSOC)) {
            $idRNA = $result['id_rna'];
            $idEstruc = $result['id_estructura_neurona'];
            $datoEstruct = $result['dato_estructura_neurona'];
            $EstrucRNA[$idEstruc-1] = $datoEstruct;
            $EstrucRNAM[$idRNA][$idEstruc-1] = $datoEstruct;
        }
    }
    try {
        $stneuro = $db->prepare("select id_neurona, u, k, a, n, capa, m, id_rna from
        neurona");
        $stneuro->execute();
        $numNeuro=0;
        while ($rneuro = $stneuro->fetch(PDO::FETCH_ASSOC)) {
            $numNeuro++;
            $m[$rneuro['id_neurona']] = $rneuro['m'];
            $caparna[$rneuro['id_neurona']] = $rneuro['capa'];
            $n[$rneuro['id_neurona']] = $rneuro['n'];
            $a[$rneuro['id_neurona']] = $rneuro['a'];
            $k[$rneuro['id_neurona']] = $rneuro['k'];
            $u[$rneuro['id_neurona']] = $rneuro['u'];
            $mM[$rneuro['id_rna']][$rneuro['id_neurona']] = $rneuro['m'];
            $caparnaM[$rneuro['id_rna']][$rneuro['id_neurona']] = $rneuro['capa'];
            $nM[$rneuro['id_rna']][$rneuro['id_neurona']] = $rneuro['n'];
            $aM[$rneuro['id_rna']][$rneuro['id_neurona']] = $rneuro['a'];
            $kM[$rneuro['id_rna']][$rneuro['id_neurona']] = $rneuro['k'];
            $uM[$rneuro['id_rna']][$rneuro['id_neurona']] = $rneuro['u'];
        }
    }
    catch (Exception $e) {
        Echo "Failed: " . $e->getMessage();
    }
    try {
        $stpeso = $db->prepare("select id_rna, id_neurona, id_peso, valor_peso from
        peso order by id_neurona, id_peso");
        $stpeso->execute();
        while ($rpeso = $stpeso->fetch(PDO::FETCH_ASSOC)) {
            $pesorna[$rpeso['id_neurona']][$rpeso['id_peso']] = $rpeso['valor_peso'];
            $pesornaM[$rpeso['id_rna']][$rpeso['id_neurona']][$rpeso['id_peso']] =
```

```

        $rpeso['valor_peso'];
    }
} catch (Exception $e) {
    Echo "Failed: " . $e->getMessage();
}
print "Numero de neuronas : ".$numNeuro."\n";
print "\nDatos de la matriz : ". "\n\n";
for ($indRNA=0; $indRNA<sizeof($pesornaM); $indRNA++){
    $id_RNA=str_pad($indRNA, 5, "0", STR_PAD_LEFT);
    print "\n*****RNA :
    ".$id_RNA."*****
    *****\n\n";
    for ($indneuro=1; $indneuro<=sizeof($pesornaM[$id_RNA]); $indneuro++){
        $id_neu=str_pad($indneuro, 1, "0", STR_PAD_LEFT);
        print "\nNeurona : ".$id_neu."-----\n\n";
        for ($npes=1; $npes<=sizeof($pesornaM[$id_RNA][$id_neu]); $npes++){
            $npe=str_pad($npes, 1, "0", STR_PAD_LEFT);
            $pesoVal=$pesornaM[$id_RNA][$id_neu][$npe];
            print sprintf("%15f",$pesoVal);
        }
        print "\n";
    }
    print "\n";
}
print "\n";
/***** APRENDIZAJE *****/
try {
    try {
        $sconfig = $db->prepare("select maximo_valor_variable_x,
        maximo_valor_variable_y, maximo_valor_x_pixel, maximo_valor_y_pixel,
        centro_coordenadas_x, centro_coordenadas_y, numero_iteraciones,
        velocidad_aprendizaje, coeficiente_x_sigma, coeficiente_y_sigma,
        umbral_salida from configuracion where codigo_configuracion='00001'");
        $sconfig->execute();
        if ($rconfig = $sconfig->fetch(PDO::FETCH_ASSOC)) {
            $xMax = $rconfig['maximo_valor_x_pixel'];
            $yMax = $rconfig['maximo_valor_y_pixel'];
            $xc1 = $rconfig['centro_coordenadas_x'];
            $yc1 = $rconfig['centro_coordenadas_y'];
            $xValMax = $rconfig['maximo_valor_variable_x'];
            $yValMax = $rconfig['maximo_valor_variable_y'];
            $ci = $rconfig['numero_iteraciones'];
            $ka = $rconfig['velocidad_aprendizaje'];
            $kx = $rconfig['coeficiente_x_sigma'];
            $ky = $rconfig['coeficiente_y_sigma'];
            $sus = $rconfig['umbral_salida'];
        }
    } catch (Exception $e) {
        Echo "Failed: " . $e->getMessage();
    }
}
$yc1 = $yMax/2;
$xzg = $xMax - 80;

```

```

$yzg = $yMax/2 - 40;
$sex = $xzg/$xValMax;
$sey = $yzg/$yValMax;
$ixcoor = $xzg/10;
$ivxmax = $xValMax/10;
$iycoor = $yzg/5;
$ivymax = $yValMax/5;
$cActuaPesos = 0;
$cActuaPesosa = 0;
$ineurona = 1;
$SumRNA = 0;
$resRNA = 0;
$resRNAa = 0;
print "Numero de iteraciones : ".$ci."\n\n";
$numNeuro=0;
for ($i_e=1; $i_e<=sizeof($EstrucRNAM[str_pad(0, 5, "0")])-1; $i_e++){
    $numNeuro = $numNeuro + $EstrucRNAM[str_pad(0, 5, "0",
        STR_PAD_LEFT)][$i_e];
}
print "Numero de neuronas : ".$numNeuro."\n";
$id_rnaN = 0;
$ci=1;
for ($iit=0; $iit<$ci; $iit++){
    print "\n\n***** Iteracion : ".$iit."*****\n";
    $stEstruc = $db->prepare("select codigo_palabra, palabra from palabra");
    $stEstruc->execute();
    while ($rEstruc = $stEstruc->fetch(PDO::FETCH_ASSOC)) {
        $codPal = trim($rEstruc['codigo_palabra']);
        $palabO = trim($rEstruc['palabra']);
        $palabm = ucfirst(strtolower(trim($palabO)));
        $narc = strtolower(preg_replace('/\s+/', '', $palabO));
        $id_rna = str_pad($id_rnaN, 5, "0", STR_PAD_LEFT);
        $numNeuRNA = sizeof($pesornaM[$id_rna]);
        $sus = $uM[$id_rna][$numNeuRNA];
        $id_rnaN++;
        print "\n\n RNA : ".$id_rna."\n";
        print "Umbral : ".$sus."\n";
        print "Palabra : ".$palabO."\n";
        print "Vector : ";
        $palL = str_split($palabO);
        for ($i=0; $i<sizeof($palL); $i++){
            $entrada[$i]=ord($palL[$i]);
        }
        $max = sizeof($entrada);
        for ($i=$max; $i<20; $i++){
            $entrada[$i]=0;
        }
        for ($i=0; $i<20; $i++){
            print $entrada[$i]." ";
        }
        print "\n\n ";
        $cActuaPesos = 0;
    }
}

```

```

$cActuaPesosa = 0;
$im = imagecreatetruecolor($xMax, $yMax);
$text_color = imagecolorallocate($im, 233, 14, 91);
$line_color = imagecolorallocate($im, 0, 0, 0);
$resul_color = imagecolorallocate($im, 255, 0, 0);
$umbral_color = imagecolorallocate($im, 0, 0, 255);
$itera_color = imagecolorallocate($im, 0, 255, 0);
$bg_Color = imagecolorallocate($im, 255, 255, 255);
imagefilledrectangle($im,0,0,$xMax, $yMax,$line_color);
imagefilledrectangle($im,1,1,$xMax-2, $yMax-2,$bg_Color);
    imagestring($im, 5,$xc1+35, 5, 'Evolución del aprendizaje de la
        palabra : '.$palabO, $line_color);
imageline ($im, $xc1, $yc1, $xc1+$xzg, $yc1,$line_color);
imagestring($im, 3, $xc1+$xzg-160, $yc1-15, "Actualización de los
        pesos", $line_color);
imageline ($im, $xc1, $yc1-$yzg, $xc1, $yc1+$yzg,$line_color);
imagestring($im, 3, $xc1+5, $yc1-$yzg-8, "Salida de la RNA",
        $line_color);

for ($ix=1; $ix<=10; $ix++){
    imageline ($im, $xc1+$ix*$ixcoor, $yc1+5, $xc1+$ix*$ixcoor,
        $yc1,$line_color);
    imagestring($im, 3, $xc1+$ix*$ixcoor-30, $yc1+10,$ix*$ixvmax,
        $line_color);
}

for ($iy=1; $iy<=5; $iy++){
    imageline ($im, $xc1, $yc1-$iy*$iycoor, $xc1-5, $yc1-
        $iy*$iycoor,$line_color);
    imageline ($im, $xc1, $yc1+$iy*$iycoor, $xc1-5,
        $yc1+$iy*$iycoor,$line_color);
    imagestring($im, 3, $xc1-35, $yc1+$iy*$iycoor-8,$iy*$iyvmax,
        $line_color);
    imagestring($im, 3, $xc1-35, $yc1-$iy*$iycoor-8,$iy*$iyvmax,
        $line_color);
}
imageline ($im,$xc1+$ex*$cActuaPesosa,$yc1-$ey*$us, $xc1+$xzg,
        $yc1-$ey*$us, $umbral_color);
$resRNA = $cRNA->calculoRNAM($entrada, $nM, $mM, $caparnaM ,
        $aM, $kM, $uM, $pesornaM,
        $numNeuro, $EstrucRNAM , $id_rna);
print "Valor inicial de la salida de la RNA : ".$resRNA."\n\n";
print sprintf("%15s %15s %20s", "No Rep.", "Can. Actual.", "Salida
        RNA")."\n";

$ineurona=1;
$irep=1;
do{
    $resRNA = $cRNA->calculoRNAM($entrada, $nM, $mM,
        $caparnaM, $aM, $kM, $uM, $pesornaM, $numNeuro,
        $EstrucRNAM,$id_rna);
    $AP = $us+10-$resRNA;
    if ($caparnaM[$id_rna][$ineurona]==1){
        if ($AP>0){
            if ($AP>1000){

```

```

    $AP=1000;
}
for ($ip=1; $ip<=$mM[$id_rna][$ineurona]; $ip++){
    $cActuaPesos++;
    $pesornaM[$id_rna][$ineurona][$ip] =
    $pesornaM[$id_rna][$ineurona][$ip] +
    $ka*$AP*$entrada[$ip-1];
    if (($cActuaPesos%1000000)==0) {
        print sprintf("%15d : %15d : %20f", $irep, $cActuaPesos,
        $resRNA)."\n
    }
}
}
else{
    if ($AP<-1000){
        $AP=-1000;
    }
    for ($ip=1; $ip<=$mM[$id_rna][$ineurona]; $ip++){
        $cActuaPesos++;
        $pesornaM[$id_rna][$ineurona][$ip] =
        $pesornaM[$id_rna][$ineurona][$ip] +
        $ka*$AP*$entrada[$ip-1];
        if (($cActuaPesos%1000000)==0) {
            print sprintf("%15d : %15d :
            %20f", $irep, $cActuaPesos, $resRNA)."\n";
        }
    }
}
}
else{
    if ($AP>0){
        if ($AP>1000){
            $AP=1000;
        }
        for ($ip=1; $ip<=$mM[$id_rna][$ineurona]; $ip++){
            $cActuaPesos++;
            $pesornaM[$id_rna][$ineurona][$ip] =
            $pesornaM[$id_rna][$ineurona][$ip] + $ka*$AP;
            if (($cActuaPesos%1000000)==0) {
                print sprintf("%15d : %15d :
                %20f", $irep, $cActuaPesos, $resRNA)."\n";
            }
        }
    }
}
else{
    if ($AP<-1000){
        $AP=-1000;
    }
}
for ($ip=1; $ip<=$mM[$id_rna][$ineurona]; $ip++){
    $cActuaPesos++;
    $pesornaM[$id_rna][$ineurona][$ip] =
    $pesornaM[$id_rna][$ineurona][$ip] + $ka*$AP;
}
}

```

```

        if (($cActuaPesos%1000000)==0)
            print sprintf("%15d : %15d :
                %20f", $irep, $cActuaPesos, $resRNA)."\n";
        }
    }
}
if ($ineurona<$numNeuro){
    $ineurona++;
}
else{
    $ineurona=1;
    $irep++;
}
imageline ($im,$xc1+$ex*$cActuaPesosa,$yc1-$ey*$resRNAa,
    $xc1+$ex*$cActuaPesos, $yc1-$ey*$resRNA, $resul_color);
$cActuaPesosa=$cActuaPesos;
$resRNAa = $resRNA;
}while(abs($us+10-$resRNA)>10);
print "Valor final de la salida de la RNA : ".$resRNA."\n";
imagejpeg($im, 'images/'. $narc.'.jpg');
imagedestroy($im);
}
}
print "Cantidad de iteraciones : ".$ci."\n\n";
print "Cantidad de actualizaciones : ".$cActuaPesos."\n\n";
}catch (Exception $e)
    Echo "Failed: " . $e->getMessage();
}
print "***** Grabando los pesos de las neuronas *****\n\n";
for ($indrRNA=0; $indrRNA<sizeof($pesornaM); $indrRNA++){
    $id_RNA=str_pad($indrRNA, 5, "0", STR_PAD_LEFT);
    for ($indneuro=1; $indneuro<=sizeof($pesornaM[$id_RNA]); $indneuro++){
        $id_neu=str_pad($indneuro, 1, "0", STR_PAD_LEFT);
        for ($npes=1; $npes<=sizeof($pesornaM[$id_RNA][$id_neu]); $npes++){
            $npe=str_pad($npes, 1, "0", STR_PAD_LEFT);
            $pesoVal=$pesornaM[$id_RNA][$id_neu][$npe];
            try {
                $pesoVal=$pesornaM[$id_RNA][$indneuro][$npes];
                $udpeso = $db->prepare("update peso set valor_peso='$pesoVal'
                    where id_peso='$npe' and id_neurona='$id_neu' and
                    id_rna='$id_RNA'");
                $db->beginTransaction();
                $udpeso->execute();
                $db->commit();
            } catch (Exception $e) {
                $db->rollBack();
                echo "Failed: " . $e->getMessage();
            }
        }
    }
}
}
}

```

```
} catch (Exception $e) {  
    Echo "Failed: " . $e->getMessage();  
}  
}
```

Anexo D: Matriz de consistencia.

PROBLEMA PRINCIPAL	OBJETIVO GENERAL	HIPOTESIS GENERAL	VARIABLES	INDICADORES	TIPO DE INVESTIGACION
<p>¿Se puede desarrollar un analizador semántico para el procesamiento del lenguaje natural para el idioma castellano?</p> <p>Problemas Específicos: ¿Se puede desarrollar un software que comprenda las oraciones del lenguaje natural utilizando redes neuronales artificiales? ¿Se puede representar el conocimiento transmitido a través de oraciones del lenguaje natural por medio de redes neuronales artificiales?</p>	<p>Desarrollar un analizador semántico de lenguaje natural para el idioma castellano</p> <p>Objetivos específicos: Desarrollar un software que comprenda las oraciones del lenguaje natural utilizando redes neuronales artificiales. Representar el conocimiento transmitido a través de oraciones dadas en lenguaje natural por medio de redes neuronales artificiales.</p>	<p>Las redes neuronales artificiales permitirán desarrollar un analizador semántico de lenguaje natural para el idioma castellano.</p> <p>Hipótesis específicas: Las redes neuronales artificiales permitirán desarrollar un software que comprenda las oraciones del lenguaje natural. Las redes neuronales artificiales permitirán representar el conocimiento transmitido a través de oraciones dadas en lenguaje natural</p>	<p>Variable Independiente: Métodos y Técnicas de aplicación de las redes neuronales artificiales para realizar un analizador semántico de lenguaje natural.</p> <p>Variable dependiente: Eficacia en el Comprensión de oraciones en Idioma castellano.</p>	<p>Indicador Complejidad del analizador semántico Índice Nro de neuronas en la red. (0% - 100%).</p> <p>Indicador 1 Cantidad de respuestas del sistema para hacer la comprensión Índice 1 Porcentaje de respuestas correctas Índice (0 – 100%) Indicador 2 Precisión en la comprensión de oraciones en idioma castellano. Índice 2 Porcentaje promedio de precisión. Índice (0 – 100%)</p>	<ul style="list-style-type: none"> • BASICA • Científica • Experimental • Aplicada