

Asistente Virtual Mediante Un Chatbot Basado En Aprendizaje Automático

Ing Jose Gabriel Silva Paredes

JUNIOR Carrera de Ingeniería de Sistemas, Escuela Militar de Ingeniería

La Paz, Bolivia

jsilvap@est.emi.edu.bo



Virtual Assistant Through A Chatbot Based On Machine Learning

Resumen – La inteligencia artificial, es la ciencia que busca crear maquinas que presenten las mismas capacidades que las del ser humano, uno de los campos que estudia es el Aprendizaje automático que se basa en el aprendizaje de las maquinas, de tal manera que se vuelvan autónomas. Esta tecnología puede ser aplicada a distintas áreas, una de ellas es la comunicación entre la maquina con las personas, una herramienta muy utilizada son los bots conversacionales o mejor conocidos como chatbots, estos tienen la función de interactuar con los clientes que requieran algún tipo de servicio o información, esto sin importar la hora o la disponibilidad física. Implementar inteligencia artificial a estos bots, mejoran notablemente la conversación con los clientes, es por esto que en el presente artículo se detalla la investigación del desarrollo de chatbots con inteligencia artificial, en el campo de el aprendizaje automático.

Palabras Claves— Asistente virtual, aprendizaje automático, chatbot

Abstract - Artificial intelligence is the science that seeks to create machines that have the same capabilities as those of the human being, one of the fields it studies is Machine Learning that is based on the learning of machines, in such a way that they become autonomous. This technology can be applied to different areas, one of them is communication between the machine and people, a widely used tool is conversational bots or better known as chatbots, these have the function of interacting with customers who require some type of service or information, this regardless of time or physical availability. Implementing artificial intelligence to these bots notably improves the conversation with customers, which is why this article details the research on the development of chatbots with artificial

intelligence, in the field of machine learning.

Keywords— Virtual assistant, machine learning, chatbot

I. INTRODUCCION

En la actualidad la Inteligencia artificial ya está muy presente en nuestra vida cotidiana, la comunicación de las personas con las organizaciones, instituciones y empresas ha evolucionado en la última década, ya que anteriormente las llamadas telefónicas y las reuniones personales eran los medios principales para comunicarse.

Se desarrollaron aplicaciones basadas en la web, que han mejorado bastante la comunicación, como las redes sociales o correo electrónico, aplicaciones web y móviles. Entonces la inteligencia artificial ha llegado a resolver o automatizar varios problemas para la comunicación entre personas e instituciones privadas o públicas pudiendo ejercer tareas que solo podrían realizar anteriormente los funcionarios de una institución.

En esta década se ha incursionado mucho lo que se refiere a la asistencia virtual, robots, Chatbots. Asimismo, un Chatbot es un software con inteligencia artificial, que está diseñado para permitir la interacción con usuarios para satisfacer alguna petición, demanda de información o un servicio sin importar la hora o la disponibilidad física de los funcionarios de las instituciones.

Uno de los problemas que hoy en día es la mayoría de las consultas que se realicen a los Chatbots permanecen sin respuesta debido a una comprensión deficiente.

Es por eso que existen posibilidades de optimización que nos ofrece el aprendizaje

automático, lo cual nos representan una gran ventaja. Nos permite que el Chatbot descifre la intención oculta detrás de la pregunta para proporcionar al usuario una respuesta pertinente y personalizada.

Dentro del Aprendizaje automático, existen varios algoritmos, cada uno son de estos son destacables según el problema que se requiera solucionar. Entre los más utilizados están:

- Regresión lineal.
- Árboles de decisión.
- Redes neuronales.
- Redes bayesianas.
- Inducción de reglas.
- Regresión logística.

Para el presente artículo nos enfocaremos en la utilización de las redes neuronales para mejorar el entendimiento del chatbot a la hora de establecer conversaciones.

II. APRENDIZAJE AUTOMÁTICO

Aprendizaje automático es una rama de la inteligencia artificial que se dedica al estudio de los programas que aprenden a realizar una tarea en base a la experiencia[1]

Objetivo: Utilizar la evidencia conocida para crear un modelo que sirva para dar respuestas a situaciones desconocidas.

El aprendizaje automático se divide en 2 tipos, supervisado y no supervisado.

A. Aprendizaje No Supervisado

Se trabaja con datos sin etiqueta o con datos de estructura desconocida. Al utilizar este tipo de técnicas no supervisadas, se pueden explorar estructuras de los datos para extraer información importante sin la guía de una variable o resultado conocido o una función objetivo [2]

B. Aprendizaje Supervisado

Tiene como objetivo principal aprender y generar un modelo, a partir de datos de entrenamiento etiquetados que permitan realizar predicciones sobre datos invisibles o futuros. [2]

El término supervisado, se refiere a un conjunto de muestras donde ya se conocen las señales de salida deseadas.

En el presente artículo nos enfocaremos en el Aprendizaje Supervisado, utilizando redes neuronales artificiales, para el desarrollo del chatbot. En puntos posteriores se profundizará en

los conceptos fundamentales de las redes neuronales.

III. REDES NEURONALES

A. Definición.

Existen numerosas formas de definir lo que son las redes neuronales, sin embargo, se puede decir que son modelos matemáticos conformados por un gran número de elementos de procesamiento que están interconectados masivamente y que tienen una organización jerárquica, los cuales intentan interactuar con los objetos del mundo real al igual que lo hace el cerebro humano para conseguir resolver problemas relacionados con el reconocimiento de patrones, predicción, codificación, control, optimización entre otros. [3]

En las RNA, los elementos de procesamiento corresponden a las neuronas biológicas, las interconexiones se realizan por medio de las ramas de salida (axones) que producen un número variable de conexiones (sinapsis) con otras neuronas. [3]

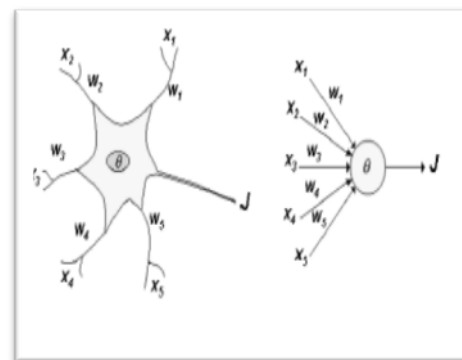


Fig. 1. Similitudes entre una neurona biológica y una artificial. Donde las entradas X_i representan todas las entradas que provienen de las demás neuronas y que son capturadas por la dendrita. Y las W_i son los pesos sinápticos, representan todas las conexiones entre todas las neuronas en la red. El θ es el umbral que debe sobrepasar toda neurona para poder activarse y generar una respuesta. Y J es la salida que ofrece la neurona a través del axón.

Las redes neuronales tienen diferentes estructuras o topologías, ya sea por el número de capas o el sentido de su conexión.

Ahora veremos la clasificación de las redes neuronales según el tipo de conexiones:

- Perceptrón simple
- Perceptrón multicapa
- Redes neuronales recurrentes (RNN)

Estas arquitecturas tienen un eficiente desempeño, según el tipo de datos con los que se trabaje. Para el desarrollo del chatbot, necesitamos tener múltiples entradas y múltiples salidas, entonces la arquitectura escogida es el Perceptron multicapa.

Se eligió trabajar con un perceptrón multicapa ya que es óptimo en problemas de clasificación de patrones, son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos, de abstraer características esenciales a partir de las entradas.

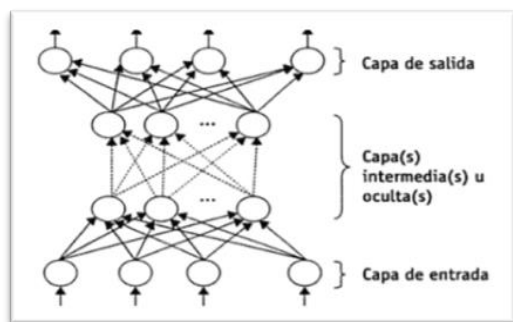


Fig. 2. Arquitectura de una red neuronal perceptrón multicapa. Donde la topología contiene una capa de entrada, 2 o mas capas ocultas y una capa de salida.

Estas arquitecturas de red neuronal se han usado ampliamente en problemas de clasificación, es decir, en problemas en los que se trata de determinar el tipo de categoría a la que pertenece un elemento desconocido.

Esto lo hace idóneo para poder trabajar en el desarrollo del chatbot basado en aprendizaje automático.

IV. CHATBOT

Un asistente virtual que ayuda a usuarios de sistemas computacionales, automatizando y realizando tareas con la mínima interacción hombre-maquina.

A veces el termino Chatbot o bot conversacional se utiliza para referirse a los asistentes virtuales en general o específicamente a aquellos a los que se accede por chat. [4]

A. Definición

Los Chatbots son programas, software, (entiéndase un conjunto de algoritmos), que utilizan procesamiento de lenguaje natural (NLP) en un sistema de preguntas y respuestas. Estos sistemas han sido definidos también como

sistemas expertos que usan razonamiento basado en casos. [5]

Los Chatbot deben tener una personalidad y responder emocionalmente, para esto se debe ingresar una gran cantidad de comandos para que aprenda el Chatbot. Existen dos tipos de Chatbot o agentes conversacionales. [6]

- Los sistemas basados en reglas.
- Basados en Inteligencia Artificial (Autoaprendizaje).

La diferencia se da al momento de responder a una pregunta, y por lo general para solventar servicios públicos se utilizan los sistemas basados en reglas y recuperación por su comunicación mecánica.

B. Chatbots basados en reglas

La creación de estos Chatbots es sencilla utilizando un enfoque basado en reglas, pero el Chatbot no es eficiente para responder preguntas, cuyo patrón no coincide con las reglas sobre las que se entrena el bot. Sin embargo, estos sistemas no pueden responder a patrones de entrada o palabras clave que no coinciden con las reglas existentes.

Entonces estos tipos de chatbots son simples, por que no contemplan inteligencia artificial para poder mantener una conversación.

Ahora a continuación veremos el tipo de chatbot del cual tratara el presente artículo.

C. Chatbots basados en Inteligencia Artificial

Los Chatbots basados en inteligencia artificial son los que utilizan algunos enfoques basados en el aprendizaje automático y son definitivamente más eficientes que los robots basados en reglas. [7]

Estos son más complejos que los Chatbots basados en reglas y tienden a ser más conversacionales, utilizar más datos y ser predictivos.

Estos Chatbots pueden ser de otros dos tipos:

- Basados en recuperación
- Generativos

Ahora para comprender mejor los enfoques de los chatbots con Inteligencia artificial, es necesario comprender sus definiciones.

1) Chatbots Generativos

Los Chatbots generativos pueden generar las respuestas y no siempre responden con una de las respuestas de un conjunto de respuestas. Esto los hace más inteligentes a medida que toman palabra por palabra de la consulta y generan las respuestas. Es un modelo bastante avanzado, se usa muy poco porque requiere algoritmos complejos. Son difíciles de construir y desarrollar. El entrenamiento de estos Chatbots requiere de mucho tiempo y esfuerzos, creando millones de ejemplos. [7]

2) Chatbots basados en Recuperación

En este modelo, un Chatbot utiliza cierta heurística para seleccionar una respuesta de una biblioteca de respuestas predefinidas. Entonces el chatbot utiliza el mensaje y el contexto de conversación para seleccionar la mejor respuesta de una lista predefinida de mensajes a dar. [7]

Este modelo de chatbots, rara vez comete errores, ya que sus respuestas ya que para responder ya tienen un repositorio donde obtener las respuestas adecuadas acordes a las consultas.

Entonces la manera de seleccionar una respuesta, se puede realizar de diferentes formas, tanto como la lógica condicional o clasificadores de aprendizaje automático.

En el presente artículo se hace el estudio de utilizar un Perceptrón multicapa como clasificador de aprendizaje automático.

D. Arquitectura de una chatbot

Para poder desarrollar un chatbot, con aprendizaje automático, basado en el enfoque de recuperación visto en el anterior punto, es necesario entender la arquitectura de los chatbots de este tipo.

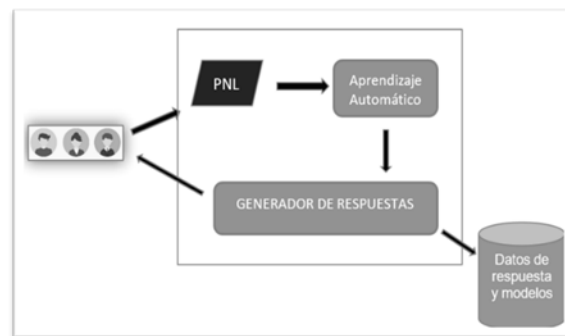


Fig. 3. Arquitectura de un chatbot basado en aprendizaje automático.

Ahora veremos las definiciones y la función que desempeñan los elementos dentro del chatbot.

1) Procesamiento de Lenguaje natural

El procesamiento de lenguaje natural (PNL) es el componente que ayuda a los chatbots a comprender el lenguaje humano. Entonces el chatbot depende de la utilización de las técnicas de PLN.

A diferencia de un motor de búsqueda, PNL se enfoca en más de una palabra clave específica, sino que utiliza la estructura de las oraciones, los patrones y el contexto para identificar la intención.

2) Aprendizaje automático

El aprendizaje automático es la rama de la inteligencia artificial basada en la idea de que los sistemas pueden aprender de los datos, identificar patrones y tomar decisiones con la mínima intervención humana, (la mayoría de estos modelos se basan en estadística). El Aprendizaje Profundo es una familia particular dentro del Aprendizaje Automático que hace uso de Redes Neuronales. [8]

3) Generador de Respuesta

Los generadores de respuestas definen lo que queremos que diga el Chatbot cuando identifica qué es lo que ha querido decir el usuario (es decir, el intent). De este modo, y de forma visual, se va creando una conversación en modo árbol, donde iremos indicando qué deberá contestar el Chatbot cuando reconoce la intención que ha querido decir el usuario.

Ahora comprendidos los conceptos, ahora es necesario utilizar las definiciones para el

desarrollo de el asistente virtual mediante un chatbot basado en aprendizaje automatico.

V. DESARROLLO

Para poder desplegar Chatbots inteligentes necesitamos que el Chatbot entienda las preguntas de los usuarios y no solo palabras predeterminadas, a esto se le llama Procesamiento de lenguaje natural (PLN).

Utilizar el aprendizaje automático para implementar el Procesamiento del Lenguaje Natural es un área en continuo desarrollo y mejora. [9]

Muchas de las tecnologías de Chatbots que se basan en la inteligencia artificial se apoyan, utilizando redes neuronales para lograr tareas que se esperaría que los Chatbots hagan mejor la cual es la de adivinar la intención o poder entender mejor las declaraciones de las palabras.

Las declaraciones se tratan de diversas y múltiples maneras de preguntar algo con la misma intención.

Ejemplo: “cuál es el costo de la tarjeta de crédito” y “cuáles son las tarifas de la tarjeta”, son esencialmente 2 declaraciones distintas con la misma intención
Tratar de entender la intención de la pregunta, de esta manera generar las respuestas desde el Chatbot.

A. Variables

Las variables necesarias para poder desarrollar el chatbot mencionado, son las siguientes.

- Intenciones
- Patrones
- Respuestas

Buenos días	Saludar
Buenas tardes	
Buenas noches	
Hola	
Me gustaría viajar	Viajar
Quiero ir de viaje	
Me quiero ir de vacaciones	

Fig. 4. Ejemplos de intenciones y patrones.

1) Intenciones

Es lo que el sistema de Inteligencia Artificial, trata de identificar cuando

alguien le escribe es la “intención” de lo que el usuario le ha querido decir.

Según el ejemplo visto anteriormente se puede observar que un dialogo puede tener diferentes declaraciones, pero si tener la misma intención.

2) Patrones

Son aquellos “inputs” del usuario (palabras, categorías, frases) que determinarán la respuesta del chatbot, puesto que son necesarias para poder llevar a cabo la acción.

3) Respuestas

Es la propia estructura de la conversación, en los diálogos, se definen lo que queremos que diga el chatbot cuando identifica qué es lo que ha querido decir el usuario (es decir, el intent). De este modo, y de forma visual, se va creando una conversación en modo árbol, donde iremos indicando qué deberá contestar el Chatbot cuando reconoce la intención que ha querido decir el usuario.

Todas estas variables son almacenadas en el repositorio de respuestas.

El repositorio de respuestas contendrá todas las intenciones con sus patrones correspondientes, y las respuestas acordes a la intención.

B.Red Neuronal y el chatbot

Un chatbot basado en recuperación, utilizando algoritmos de aprendizaje supervisado, estos trabajan con datos etiquetados, en nuestro caso salidas que llegaría a ser intenciones, los algoritmos intentan encontrar una función que, dadas las variables de entradas, se les asigne la salida adecuada.

El algoritmo de aprendizaje supervisado, tal como se mencionó en anteriores puntos es una red neuronal perceptrón multicapa. Entonces la arquitectura de un chatbot con redes neuronales para poder identificar la intención de una consulta, según las palabras ingresadas es la siguiente.

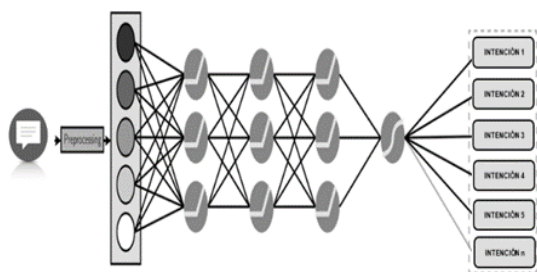


Fig. 4. Red neuronal perceptrón multicapa para la clasificación de intenciones.

1) Definición de la Red Neuronal

Un perceptrón multicapa tiene una o mas capas ocultas. Entonces para la implementación de una red neuronal para la clasificación de intenciones debe ser acorde a el repositorio de respuestas del chatbot.

Capa de Entrada: El numero de neuronas en la capa de entrada será igual a la cantidad de palabras – patrones dentro del repositorio de respuestas.

Capa de Salida: La cantidad de neuronas en la capa de salida será igual a el numero de intenciones del repositorio de intenciones, antes mencionado.

Capas ocultas: No existe una fórmula que permita seleccionar el numero óptimo de capas ocultas, así como la cantidad de nodos en cada una de ellas, se recomienda que se deba comenzar con una o dos capas ocultas, pero esto será respondido una vez realizado el entrenamiento de la red neuronal.

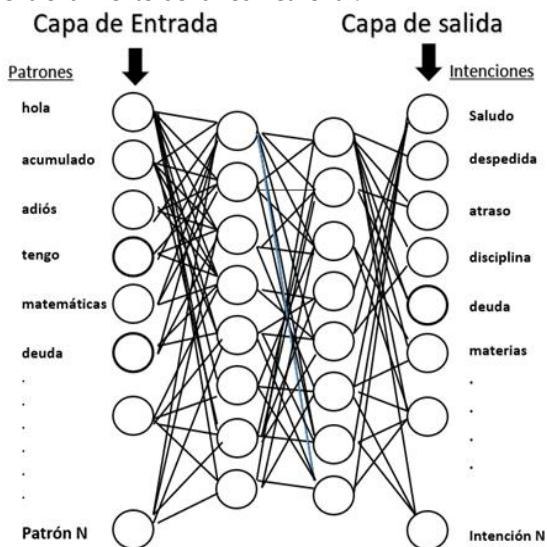


Fig. 5. Arquitectura de la Red Neuronal, donde en la capa de entrada van los patrones y en la de salida las intenciones.

Para poder realizar el entrenamiento de la red

neuronal es necesario convertir las palabras del repositorio de respuestas, en un formato numérico vectorial para que la red neuronal pueda realizar el entrenamiento con estos datos.

Este proceso se lo conoce como Pre-Procesamiento que veremos a continuación.

2) Pre-procesamiento

Para poder entrenar la red neuronal es necesario poder convertir el texto a formato numérico vectorial. Este proceso se realiza gracias a las técnicas de Procesamiento de Lenguaje Natural (PNL).

Las técnicas necesarias son las siguientes:

- Tokenización
- Normalización
- Bolsa de Palabras

Tokenización: Esta técnica se encarga de dividir las oraciones en palabras y separar las palabras de los signos de interrogación, puntos, comas, etc. [10]

etc. [16]				
	Hola	Adiós	tengo	deudas
Hola	=	[1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...,0]		
Adiós	=	[0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...,0]		
tengo deudas	=	[0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, ...,0]		
control de atrasos	=	[0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, ...,0]		
registro de faltas	=	[0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, ...,0]		
	Control	de	atrasos	

Normalización: Pone todas las palabras en igualdad, reduce todas las palabras tokenizadas de mayúsculas a minúsculas, para que no exista confusiones a la hora de emplear el entrenamiento. [10]

Bolsa de Palabras: Una de las técnicas más simples de codificación vectorial es bolsa de palabras (BOW) o Codificación One-hot donde el tamaño de los vectores es igual al número total de palabras únicas, y donde todos los valores del vector valen 0 excepto la posición que corresponde al índice de cada palabra. [10]

Fig. 6. Ejemplo de la bolsa de palabras, donde cada palabra tiene una representación numérica vectorial.

El pre-procesamiento se realiza tanto a los patrones como a las intenciones, entonces el resultado esperado es el siguiente.

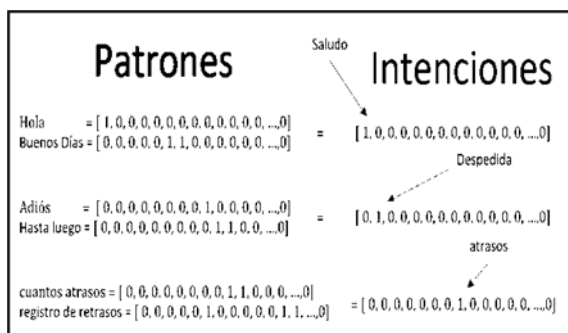


Fig. 7. Pre-procesamiento a los patrones e intenciones del repositorio de respuestas.

Ahora que se ha convertido las palabras del repositorio de respuestas, comienza el entrenamiento de la red neuronal.

3) Algoritmo de Aprendizaje

Como el algoritmo es de aprendizaje supervisado, el más recomendable es el Backpropagation, el cual es el método de entrenamiento más utilizado en redes con conexión hacia delante. Es un método de aprendizaje supervisado de gradiente descendente, en el que se distinguen claramente dos fases: primero se aplica un patrón de entrada, el cual se propaga por las distintas capas que componen la red hasta producir la salida de la misma. Esta salida se compara con la salida deseada y se calcula el error cometido por cada neurona de salida. Estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de las capas intermedias. [11]

Cada neurona recibe un error que es proporcional a su contribución sobre el error total de la red. Basándose en el error recibido, se ajustan los errores de los pesos sinápticos de cada neurona.

4) Utilización de la Red Neuronal

Una vez realizado el entrenamiento, ahora se debe de verificar el funcionamiento de la red neuronal, con el ingreso de consultas. Para determinar si las respuestas tienen relación a las preguntas ingresadas por el usuario.

En la siguiente Figura observaremos el esquema del funcionamiento de la red neuronal, donde se especifica el funcionamiento desde el ingreso de la consulta hasta la identificación de la Intención, para posteriormente responder al usuario con una de las respuestas del repositorio de respuestas.

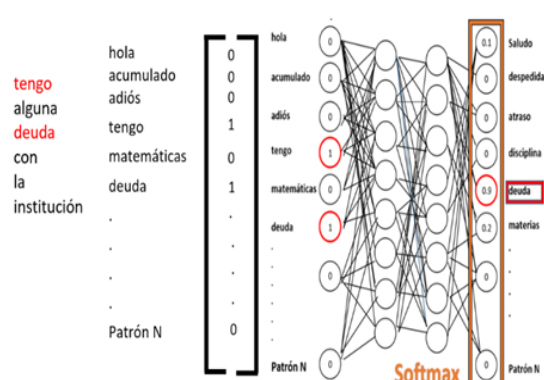


Fig. 8. Funcionamiento de la Red Neuronal, al ingreso de las consultas del usuario.

Como se observa en la imagen las consultas ingresadas en formato texto para posteriormente ser Pre-procesadas por técnicas de procesamiento de lenguaje natural (PLN), convirtiéndolas a un valor numérico vectorial.

En la técnica de bolsa de palabras, se crea un vector lleno de ceros, pero con los índices de las palabras-patrones del repositorio de respuestas, esto sirve para el caso en el que ingresen algunas palabras que se encuentren dentro del mencionado repositorio, pasen a tener un valor de 1 dentro del vector.

Entonces estos pesos, ingresan a la red neuronal entrenada, esta realizará el proceso de clasificación o identificación de intenciones según el entrenamiento que tuvo, con el histórico de datos del repositorio de respuestas, tal y como se observa en la Fig. 5. La red neuronal identificará la intención de la oración ingresada.

Fig. 9. Conversación con el chatbot, donde se evidencia que las respuestas van acorde a las preguntas.

VI. CONCLUSIÓN

1. El Desarrollo de un chatbot con aprendizaje automático, implementando una red neuronal, específicamente un perceptrón multicapa, es un avance en el estudio de poder elaborar un asistente que comprenda las intenciones de las oraciones y pueda responder coherentemente.
2. La red neuronal fue entrenada bajo un enfoque de aprendizaje supervisado, lo que beneficia a el chatbot para que no cometa errores a la hora de responder, así los usuarios recibirán respuestas acordes.
3. Un límite con el que nos encontramos es que para que el chatbot pueda aprender nuevas intenciones, es necesario

adicionarlas a el repositorio de respuestas manualmente.

RECONOCIMIENTO

El presente trabajo se desarrolla en el marco del proyecto de investigación, elaborado en la Escuela Militar de Ingeniería – Carrera de Ingeniería de sistemas en la gestión 2020.

Referencias

- [1] Raschka, S. (2015). Python Machine Learning Unlock deeper insights into machine learning with this vital guide to cutting-edge predictive analytics. (R. Banerjee & S. Copestake, Eds.). BIRMINGHAM - MUMBAI: Packt Publishing Ltd.
- [2] Mario Rodríguez Cantelar (2018). Estudio y análisis de las arquitecturas de redes neuronales aplicadas a agentes conversacionales. Lugar: “Universidad Politécnica de Madrid”
- [3] Br. Argenis A. Moreno V. (2005). Proyecto de grado “Redes neuronales artificiales para resolver problemas de programación lineal”
- [4] Luis Enrique Cubero Caba, 2015 “Asistente virtual para la web de la facultad informática” Proyecto de sistemas informáticos. Universidad Complutense de Madrid.
- [5] Wallace, R. (2003). The Elements of AIML Style. ALICE A.I. Foundation.
- [6] Ezquerro, A. N. (2018). Implementing ChatBots using Neural Machine Translation techniques, 2017– 2018.
- [7] Parul Pandey. (2018). Construyendo un Chatbot simple desde cero en Python (usando NLTK) - Recuperado de: <https://planetachatbot.com/construyendo-un-chatbot-simple-desde-cero-en-python-usando-nltk-31b9ae4f71db>
- [8] Ramiro Savoie - Chatbots con inteligencia artificial: Un mejor servicio de atención – 29 de agosto del 2018-Recuperado de: <https://blog.cliengo.com/chatbots-inteligencia-artificial-servicio-de-atencion-al-cliente/>
- [9] Carlos F. González (2019)- Cómo utilizar el Aprendizaje de Máquina para implementar Chatbot - Recuperado de: <https://www.inbenta.com/es/blog/como-utilizar-el-aprendizaje-de-maquina-para-implementar-chatbots>
- [10] Hunter Heidenreich. Introduction to Word Embeddings. (2018) Recuperado: <http://hunterheidenreich.com/blog/intro-to-word-embeddings/>
- [11] Fritsch, J. (1996). Modular Neural Networks for Speech Recognition. Master’s Thesis. Carnegie Mellon University.

Fecha de Envío del Artículo: 14/05/2020

Fecha de Aceptación de artículo: 28/05/2020