

Universidad
Nacional de
General Sarmiento-
Proyecto
Profesional 1

TP Inicial: Segunda Entrega



Docentes: Juan Carlos Monteros y Evelyn
Aragon

Alumnos: Gabriel Althaparro, Federico
Gonzalez, Esteban Ezequiel Lucero

Índice

1. Descripción General:	2
2. Funcionalidad	2
3. Arquitectura:	3
4. Tecnologías Utilizadas:	3
5. ¿Qué es NLTK, Tokenizar y Lematizar?	4
6. Entrenamiento del Modelo:	4
7. Implementación:	4
8. Bibliografía:	5

Documentación Técnica y Funcional del Chatbot

1. Descripción General:

En esta entrega presentaremos el ChatBot en una aplicación para reservas de cancha que utilizará inteligencia artificial (IA) para interactuar con usuarios mediante un chat de texto que resolverá las dudas del mismo. Esta IA utiliza una red neuronal artificial para comprender y responder a las consultas de los usuarios de manera automatizada y con un entrenamiento supervisado para su mejor funcionamiento.

2. Funcionalidad

En este apartado definiremos las principales funcionalidades que debe presentar la aplicación siendo la misma la primera entrega y dando mayor profundidad a lo que debe realizar el ChatBot lo cual es:

- Como usuario debemos poder preguntar la hora de apertura y cierre de las canchas de diferentes maneras al ChatBot para que nos dé la respuesta adecuada.
- Como usuario debemos poder preguntar el precio de la cancha de diferentes maneras al ChatBot para que nos dé la respuesta adecuada.
- Como usuario debemos poder preguntar los horarios de las canchas disponibles de diferentes maneras al ChatBot para que nos dé la respuesta adecuada.
- Como usuario debemos poder preguntar la ubicación del centro deportivo de diferentes maneras al ChatBot para que nos dé la respuesta adecuada.
- Como usuario debemos poder tener una charla fluida con el ChatBot.
- Como usuario debemos poder preguntar otras formas de comunicación de diferentes maneras al ChatBot para que les dé el número de atención al cliente.
- Como usuario debemos poder reservar o cancelar alguna cancha de diferentes maneras al ChatBot para que les dé el número de atención al cliente.

Las Funcionalidades básicas de este ChatBot cubre las necesidades del usuario para que pueda tener información básica del centro de canchas de manera rápida y sencilla sin tener que esperar a la respuesta del personal.

3. Arquitectura:

El ChatBot se basa en una arquitectura en donde el cliente a través de la interfaz de usuario enviará mensajes y el ChatBot responderá a partir de su entrenamiento, con un banco de datos propio, cuál es la mejor respuesta. La arquitectura incluye los siguientes componentes principales:

- Interfaz de Usuario: Interfaz a través de la cual los usuarios envían mensajes al chatbot.
- Procesamiento de Texto: Preprocesamiento de texto para convertir los mensajes de los usuarios en datos que la red neuronal pueda procesar.
- Red Neuronal: Modelo de red neuronal artificial que clasifica los mensajes de los usuarios en categorías y genera respuestas adecuadas.
- Integración con Datos: Integración con archivos de datos que contienen patrones de mensajes y respuestas asociadas para entrenar el modelo de red neuronal.

4. Tecnologías Utilizadas:

El ChatBot está desarrollado utilizando las siguientes tecnologías principales:

- Python: Lenguaje de programación utilizado para escribir el código del ChatBot.
- Keras: Biblioteca de aprendizaje profundo utilizada para construir y entrenar el modelo de red neuronal.
- NLTK (Natural Language Toolkit): Biblioteca de procesamiento de lenguaje natural utilizada para tokenizar y lematizar palabras en los mensajes de los usuarios.
- JSON: Formato utilizado para almacenar los datos de entrenamiento del chatbot.
- Terminal: en esta entrega esta va a ser nuestra interfaz de usuario por el cual se va a poder interactuar con el ChatBot.

5. ¿Qué es NLTK, Tokenizar y Lematizar?

NLTK es una librería de Python que se utiliza para el procesamiento de lenguaje natural y que nos brinda herramientas para tokenizar y lematizar palabras que nos sirven como datos de entrenamiento para la red neuronal que va a utilizar el ChatBot.

Ya visto lo que es NLTK nos consta de ver el significado de tokenizar que refiere a la acción de separar una frase en unidades más pequeñas tales pueden ser palabras, sub palabras o caracteres para poder hacer un correcto análisis léxico.

Por último, después de tokenizar tendremos que lematizar que refiere al proceso lingüístico y de procesamiento de lenguaje natural que consiste en reducir las palabras a su forma base o raíz como puede ser el infinitivo de un verbo, el singular de un sustantivo y el singular masculino para los adjetivos.

6. Entrenamiento del Modelo:

El modelo de red neuronal del ChatBot se entrena utilizando un conjunto de datos que contiene patrones de mensajes de usuarios y las respuestas asociadas. El proceso de entrenamiento implica los siguientes pasos:

- Tokenización y lematización de palabras en los mensajes de los usuarios.
- Generación de una representación de bolsa de palabras para cada mensaje.
- Entrenamiento de la red neuronal utilizando un modelo secuencial y un optimizador descenso de gradiente estocástico (SGD).
- Guardado del modelo entrenado para su posterior uso en la generación de respuestas.

7. Implementación:

El ChatBot se implementa como una aplicación de consola en Python. La implementación incluye los siguientes archivos principales:

- train_chatbot.py: Script utilizado para entrenar el modelo de red neuronal utilizando los datos de entrenamiento. En donde las primeras líneas de código se utiliza la librería NLTK para poder procesar los datos y así poder generar un modelo de IA que esté bien entrenada
- chatbot.py: Script principal que contiene la lógica del ChatBot, incluyendo la clasificación de mensajes de usuarios y la generación de respuestas.
- intents.json: Archivo JSON que contiene los patrones de mensajes y respuestas asociadas utilizadas para entrenar el modelo.

8. Bibliografía:

<https://www.adictosaltrabajo.com/2023/07/27/nltk-python/>
<https://es.wikipedia.org/wiki/Lematizaci%C3%B3n>