

Diseño, desarrollo y entrenamiento de un Bot para una plataforma transmedia del páramo de  
Santurbán

Juan Andres López Cubides

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Ingeniería Eléctrica, Electrónica y de Telecomunicaciones

Bucaramanga

2022

Diseño, desarrollo y entrenamiento de un Bot para una plataforma transmedia del páramo de  
Santurbán

Juan Andres López Cubides

Trabajo de Grado para Optar al Título de Ingeniero Electrónico

Director

Homero Ortega Boada

Doctor en ciencias de la ingeniería, radio comunicaciones

Codirector (Externo)

Nelson Enrique Trillos León

Ingeniero Electrónico

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Ingeniería Eléctrica, Electrónica y de Telecomunicaciones

Bucaramanga

2022

## **Dedicatoria**

A mi madre infinitas gracias por su amor, sacrificio y apoyo incondicional, quien nunca dejó de creer en mis capacidades durante todos estos años, quien estuvo para mí en todas las etapas más difíciles de mi vida, quien me dio voz de aliento cuando más la necesitaba, quien nunca dejó de ser un ejemplo a seguir, quien siempre fue una de mis más grandes motivaciones para darlo todo en mi formación académica.

A mi tío Iván Cuervo gracias, porque fue una de las personas que me brindó su apoyo incondicional, quien creyó en mí y me dio el empujón que necesitaba para poder iniciar mi proceso académico en una de las mejores universidades del país.

A mis tíos Angela Cuervo y Orlando Vargas, quienes me han apoyado incondicionalmente, quienes siempre estuvieron pendientes de mí todos estos años, quienes sembraron en mí el sueño de querer ser un profesional e ir más allá en el futuro, a ellos gracias por siempre estar ahí.

A mis abuelos, eternamente agradecido por el amor que me han brindado, por el apoyo incondicional, por su respaldo siempre, por el tiempo brindado, quienes han estado presentes en todo este proceso académico, infinitas gracias, porque sin ellos, haber llegado hasta aquí no hubiera sido posible.

A mis hermanos les agradezco por siempre creer en mí, por siempre estar para mí, por ser junto a mi madre, una las razones por las que cada día de mi formación, lo di todo.

Finalmente, pero no menos importante, quiero agradecerles a los demás integrantes de mi familia, quienes siempre también estuvieron ahí para apoyarme incondicionalmente, quienes pusieron su granito de arena en este largo proceso académico. Hoy puedo llegar a decir que estoy orgulloso de la espectacular familia con la que fui bendecido.

## **Agradecimientos**

Quiero agradecer a todos los docentes y tutores que me acompañaron en este proceso de formación en la Universidad Industrial de Santander, por haberme brindado sus conocimientos, experiencias y tiempo, que aportaron un grano de arena para lograr formarme como Ingeniero Electrónico.

Quiero agradecer al ingeniero Homero Ortega, director del presente proyecto de grado y del grupo de investigación Radiogis, y al ingeniero Nelson Trillos, Codirector externo, por haber compartido sus experiencias, conocimientos, asesorías y tiempo para poder dar desarrollo a cada etapa que conllevó al éxito de este proyecto de grado.

Un agradecimiento especial al profesor Humberto Rodríguez, Ingeniero de telecomunicaciones y docente en tecnología e informática en el colegio Lizcano Flores, quién me abrió el espacio para poder trabajar con los 11 niños de la institución y así poder realizar las diferentes actividades propuestas y finalmente también quiero agradecerles a estos 11 niños, quienes participaron de forma activa en las actividades programadas para dar culminación al proyecto.

## Tabla de Contenido

|  | <b>Pág.</b> |
|--|-------------|
| Introducción .....                                   | 17          |
| 1. Objetivos.....                                    | 21          |
| 1.1. Objetivo General.....                           | 21          |
| 1.2. Objetivos Específicos.....                      | 21          |
| 2. Marco teórico .....                               | 22          |
| 2.1. Machine Learning .....                          | 22          |
| 2.1.1. Aprendizaje Supervisado .....                 | 23          |
| 2.1.1.1. Regresión. ....                             | 24          |
| 2.1.1.2. Clasificación .....                         | 24          |
| 2.1.2. Aprendizaje No supervisado .....              | 24          |
| 2.1.2.1. Clustering.....                             | 24          |
| 2.1.2.2. Detección anómala.....                      | 24          |
| 2.1.2.3. Reducción de dimensionalidad .....          | 24          |
| 2.1.2.4. Asociación de reglas .....                  | 25          |
| 2.2. Aprendizaje por refuerzo .....                  | 25          |
| 2.3. Deep Learning.....                              | 25          |
| 2.3.1. Redes neuronales convolucionales .....        | 26          |
| 2.4. Procesamiento del lenguaje natural (NLP) .....  | 27          |
| 2.4.1. Entendimiento del lenguaje natural (NLU)..... | 28          |
| 2.5. Chatbot.....                                    | 28          |
| 2.5.1. Basados en caja de texto (chatterboxes) ..... | 28          |

|           |  |    |
|-----------|--|----|
| 2.5.2.    | Asistentes virtuales personificados (embodied conversational agents) ..... | 28 |
| 2.5.3.    | Físicos .....  | 29 |
| 2.5.4.    | Terminología general en Chatbots .....                                     | 29 |
| 2.5.4.1.  | Agentes .....  | 29 |
| 2.5.4.2.  | Flujos.....  | 29 |
| 2.5.4.3.  | Paginas .....  | 30 |
| 2.5.4.4.  | Tipos de entidad.....  | 30 |
| 2.5.4.5.  | Parámetros.....  | 30 |
| 2.5.4.6.  | Intentos.....  | 31 |
| 2.5.4.7.  | Webhook.....   | 31 |
| 2.5.4.8.  | Entrega .....  | 31 |
| 2.5.4.9.  | Integraciones .....  | 32 |
| 2.5.4.10. | Interacción.....   | 32 |
| 2.6.      | Métricas de rendimiento .....  | 33 |
| 2.6.1.    | Tasa de aciertos.....  | 33 |
| 2.6.2.    | Matriz binaria.....  | 34 |
| 2.6.2.1.  | Exactitud .....  | 34 |
| 2.6.2.2.  | Precisión.....   | 34 |
| 2.6.2.3.  | Recall .....   | 34 |
| 2.6.2.4.  | F1-score.....  | 35 |
| 2.6.3.    | Matriz multiclase .....  | 35 |
| 2.6.3.1.  | Exactitud .....  | 35 |
| 2.6.3.2.  | Precisión.....   | 35 |

|          |   |    |
|----------|---|----|
| 2.6.3.3. | Recall .....  | 36 |
| 2.6.3.4. | F1-score.....   | 36 |
| 3.       | Marco de referencia .....   | 36 |
| 3.1.     | PATH .....  | 37 |
| 3.2.     | CheMBL .....  | 37 |
| 3.3.     | Chatbot como herramienta para el aprendizaje de un nuevo idioma ..... | 37 |
| 3.4.     | Chatbot utilizado en el campo de la educación.....                    | 38 |
| 3.5.     | Florence.....   | 38 |
| 3.6.     | Trabajo realizado en la Universidad Industrial de Santander .....     | 39 |
| 4.       | Requerimientos y restricciones de diseño para el Chatbot .....        | 40 |
| 5.       | Selección de plataforma.....  | 40 |
| 5.1.     | Dialogflow .....  | 41 |
| 5.1.1.   | Dialogflow ES.....  | 41 |
| 5.1.1.1. | Características .....   | 41 |
| 5.1.1.2. | Costos Dialogflow ES.....   | 42 |
| 5.1.2.   | Dialogflow CX.....  | 42 |
| 5.1.2.1. | Características .....   | 42 |
| 5.1.1.2. | Costos.....   | 43 |
| 5.2.     | Microsoft – Azure Bot Service .....                                   | 44 |
| 5.2.1.   | Características .....   | 44 |
| 5.2.2.   | Costos.....   | 44 |
| 5.3.     | IBM - Watson .....  | 45 |
| 5.3.1.   | Características .....   | 46 |

|        |   |    |
|--------|---|----|
| 5.3.2. | Costos.....   | 46 |
| 5.4.   | Amazon Lex.....   | 46 |
| 5.4.1. | Características.....  | 46 |
| 5.4.2. | Costos.....   | 47 |
| 5.5.   | Criterio de selección de plataforma.....                          | 48 |
| 6.     | Recopilación de preguntas para la nueva versión del Chatbot.....  | 49 |
| 7.     | Casos de uso.....   | 50 |
| 7.1.   | Rol de desconocido.....   | 50 |
| 7.2.   | Rol de estudiantes del semillero Radiogis.....                    | 50 |
| 7.3.   | Rol de profesores.....  | 51 |
| 8.     | Diseño de diagramas conversacionales.....                         | 53 |
| 8.1.   | Diagrama de la arquitectura del Chatbot.....                      | 53 |
| 8.2.   | Diagrama conversacional 1.....                                    | 54 |
| 8.3.   | Diagrama conversacional 2.....                                    | 55 |
| 8.4.   | Diagrama conversacional 3.....                                    | 56 |
| 8.5.   | Diagrama conversacional 4.....                                    | 57 |
| 8.6.   | Diagrama conversacional 5.....                                    | 58 |
| 9.     | Implementación del Chatbot en Dialogflow CX.....                  | 59 |
| 9.1.   | Creación de las entidades.....                                    | 59 |
| 9.1.1. | Creación y configuración de los intentos.....                     | 60 |
| 10.    | Análisis de resultados.....                                       | 66 |
| 10.1.  | Sondeo previo.....  | 66 |
| 10.2.  | Actividad para la identificación de necesidades de los niños..... | 67 |

|         |                                     |    |
|---------|-------------------------------------|----|
| 10.3.   | Test de rendimiento 1.....          | 69 |
| 10.3.1. | Matriz de confusión binaria .....   | 69 |
| 10.3.2. | Matriz de confusión multiclase..... | 71 |
| 10.4.   | Test de rendimiento 2.....          | 72 |
| 10.4.1. | Matriz de confusión binaria .....   | 72 |
| 10.4.2. | Matriz de confusión multiclase..... | 74 |
| 11.     | Trabajos futuros .....              | 75 |
| 12.     | Conclusiones .....                  | 75 |
| 13.     | Recomendaciones .....               | 77 |
|         | Referencias Bibliográficas .....    | 78 |
|         | Apéndices.....                      | 82 |

## Lista de Tablas

|  | <b>Pág.</b> |
|--|-------------|
| Tabla 1. Precios para la plataforma Dialogflow CX.....                                     | 43          |
| Tabla 2 Precios de Azure Bot Service .....   | 45          |
| Tabla 3. Precio de Amazon lex - Precios de interacción de respuesta y solicitud.....       | 47          |
| Tabla 4. Precio de Amazon Lex - Precios de la conversación en Streaming.....               | 48          |
| Tabla 5. Tabla de la métrica de rendimiento de la tasa de aciertos .....                   | 66          |
| Tabla 6. Matriz de confusión binaria en el Test de rendimiento 1 .....                     | 70          |
| Tabla 7. Resultados de la matriz de confusión multiclase en el Test de rendimiento 1 ..... | 71          |
| Tabla 8. Matriz de confusión binaria en el Test de rendimiento 2 .....                     | 73          |
| Tabla 9. Resultados de la matriz de confusión multiclase en el Test de rendimiento 2 ..... | 74          |

## Lista de Figuras

|   | <b>Pág.</b> |
|---|-------------|
| Figura 1. Ejemplo de clasificación de mensajes Spam .....   | 23          |
| Figura 2. Ilustración de un modelo de Deep Learning.....  | 26          |
| Figura 3. Diagrama sobre la distribución general del procesamiento natural del lenguaje como parte de la inteligencia artificial..... | 27          |
| Figura 4. Diagrama de flujos en Dialogflow CX.....  | 30          |
| Figura 5. Grafica de interacciones con API.....   | 33          |
| Figura 6. Grafica de los casos de uso identificados .....   | 51          |
| Figura 7. Caso de uso - Chat conversacional.....  | 52          |
| Figura 8. Arquitectura del Chatbot basado en diagramas UML .....  | 53          |
| Figura 9. Diagrama conversacional 1 .....   | 54          |
| Figura 10. Diagrama conversacional 2 .....  | 55          |
| Figura 11. Diagrama conversacional 3 .....  | 56          |
| Figura 12. Diagrama conversacional 4 .....  | 57          |
| Figura 13. Diagrama conversacional 5 .....  | 58          |
| Figura 14. Grafica de identidades integradas en Dialogflow CX.....  | 59          |
| Figura 15. Grafica de los intentos integrados en Dialogflow CX.....   | 60          |
| Figura 16. Grafica de la configuración de un intento en Dialogflow Cx.....  | 61          |
| Figura 17. Grafica de la página Start en la plataforma de Dialogflow CX.....  | 62          |
| Figura 18. Grafica de la configuración de un intento en Dialogflow Cx.....  | 63          |
| Figura 19. Grafica del interfaz de una página en Dialogflow CX .....  | 64          |

Figura 20. Grafica del interfaz grafico del Chatbot diseñado e implementado en Dialogflow Cx.

..... 65

## Lista de Apéndices

|   | <b>pág.</b> |
|---|-------------|
| Apéndice A. Recopilación de frases de entrenamiento de la versión preexistente..... | 82          |
| Apéndice B. Sondeo previo .....   | 82          |
| Apéndice C. Test de rendimiento 1.....  | 82          |
| Apéndice D. Test de rendimiento 2 .....   | 82          |

Los apéndices están adjuntos y puede visualizarlos en la base de datos de la biblioteca UIS

## Glosario

**Métricas de rendimiento:** Son estrategias que se emplean para determinar la eficacia del aprendizaje de un modelo que incorpore inteligencia artificial. (Ricardo Borja-Robalino et al., 2020)

**Páramo:** Los páramos son un tipo de ecosistema montañoso que se encuentra los bosques altoandinos, los cuales juegan un papel muy importante en la producción de agua y en el cual viven diversas plantas, animales y microorganismos. (Alexander von Humboldt, 2011)

**UML:** Es un estándar que se emplea para representar la estructura y comportamiento mediante modelos de un sistema, el cual permite describir como interactúan todos los elementos de este.

*(Precise Semantics of UML State Machines*

*(PSSM)2019)*

**Transmedia:** Es una narrativa que se desarrolla en los diferentes medios y plataformas digitales, los cuales intentan contar una historia que puede partir de un cómic, un dibujo animado entre otros. (Alejandro Caloguera, 2012)

## Resumen

**Título:** Diseño, desarrollo y entrenamiento de un Bot para una plataforma transmedia del páramo de Santurbán \*

**Autor:** Juan Andres López Cubides \*\*

**Palabras Clave:** Inteligencia Artificial, Chatbot

**Descripción:** Con el auge tecnológico que actualmente hay en torno a las IA (Inteligencia Artificial), ha sido posible realizar tareas que hace algunos años era impensable, como apoyar, optimizar y agilizar procesos a nivel industrial y social.

En el presente documento, se describe el proceso de desarrollo detallado que se llevó a cabo en el diseño, implementación y entrenamiento de un Chatbot para una plataforma transmedia que apoya la formación de niños y jóvenes habitantes de los alrededores del páramo de Santurbán en Colombia, que hace parte de la iniciativa investigadores por naturaleza: Expedición Santurbán, el cual es una iniciativa por parte de la vicerrectoría de investigación y extensión de la Universidad Industrial de Santander. Para este desarrollo se utilizó la plataforma de Dialogflow CX que incorpora modelos de inteligencia artificial tales como procesamiento y entendimiento del lenguaje natural (NLP/NLU). Para el diseño e implementación, se identificaron los posibles casos de uso, los cuales en su mayoría son del tipo conversacional, el cual sirvió como punto de partida para el diseño de los diagramas conversacionales basados en máquinas de estado utilizando la norma UML. El diseño del Chatbot incluye pruebas de exactitud para lo cual se utilizaron 3 tipos de métricas de rendimiento tales como la tasa de acierto, matriz de confusión binaria y multiclase.

---

\* Trabajo de grado de investigación

\*\* Escuela de ingenierías eléctrica, electrónica y de telecomunicaciones. Ingeniería electrónica. Director: Homero Ortega Boada. Título académico completo de mayor rango. Codirector externo: Nelson Enrique Trillos León. Ingeniero electrónico.

### Abstract

**Title:** Design, Development, and training of a Bot for a transmedia platform of the moor of Santurbán\*

**Author:** Juan Andres López Cubides\*\*

**Key words:** Artificial Intelligence, Chatbot, transmedia

**Description:** With the technological boom that currently exists around AI (Artificial Intelligence), it has been possible to perform tasks that were unthinkable a few years ago, such as supporting, optimizing, and streamlining processes at an industrial and social level.

This document describes the detailed development process for the design, implementation, and training of a Chatbot for a transmedia platform that supports the education of children and young people living in the surroundings of the Santurbán moor in Colombia that belong to the initiative of the program: researchers by nature: expedition Santurbán, who is an initiative for the part vice rector in research and extension at the Universidad Industrial de Santander. For this development, it used the Dialogflow CX platform that it has artificial intelligence models for processing and understanding natural language (NLP/NLU). For the design and implementation, it identified the possible cases use, which in most are conversation type, which it was utilized as a start point for the design of diagram conversational based in state machine using the rule UML. The design of Chatbot includes test accuracy for which it used 3 types of performance metrics as hit rate, matrix confusion binary, and multiclass.

---

\* Research degree work

\*\*School of electrical, electronic and telecommunications engineering. Electronic Engineering. Director: Homero Ortega Boada. Academic title: Philosophy Doctor (PhD) of Engineering Science. Codirector extern: Nelson Enrique Trillos León. Engineering Electronic

## Introducción

El páramo de Santurbán hace parte de los 37 páramos que tiene Colombia, abasteciendo de agua a aproximadamente 2,3 millones de personas y es el hábitat de aproximadamente 3.200 especies de plantas y de algunas especies de animales que están en peligro de extinción. Al mismo tiempo los niños y jóvenes con edades entre los ocho (8) y quince (15) años, habitantes del páramo y sus alrededores, pasan grandes dificultades para acceder al conocimiento debido a que es un territorio rural y tiene limitaciones en cuanto la conectividad a internet, algo que se ha evidenciado principalmente durante la pandemia por COVID-19.

En el marco del proyecto transmedia - investigadores por naturaleza: Expedición Santurbán, que hace parte de una iniciativa de la vicerrectoría de Investigación y Extensión de la Universidad Industrial de Santander, (Comunicaciones, 2022) , que ha venido invirtiendo esfuerzos en involucrar a los niños y jóvenes del páramo Santurbán, en actividades propias de algunos grupos de investigación, de los cuales es destacable el aporte por parte de Radiogis, perteneciente a la escuela de ingeniería eléctrica, electrónica y telecomunicaciones (E3T) con el desarrollo de Kits y de cartillas para encaminar a los niños hacia el internet de las cosas.

Otro de los recientes aportes realizados por el grupo de investigación Radiogis, es el desarrollo de una primera versión de un Chatbot para una plataforma transmedia, donde este aborda la narrativa de ser el páramo de Santurbán, con la disposición de responder información relacionada con el páramo e información de interés.

Los Chatbots son un tipo de asistentes virtuales especializados sobre un área en particular, siendo actualmente una de las tecnologías emergentes de mayor dinamismo, los cuales se han desarrollado en diferentes campos del conocimiento debido a que incorporan inteligencia artificial, con la facilidad de añadirse en páginas web y redes sociales, lo que facilita su acceso a estos. Por

lo cual se cree que hacia el futuro serán los actores esenciales para el aprovechamiento de la inteligencia artificial en la prestación de servicios para las personas, (Kasthuri & Balaji, 2021).

En Colombia ya se ha venido trazando un camino hacia dicho futuro, como lo es el caso del proyecto “La primera humana virtual colombiana”, (Vanguardia, 2021), que promueve el ingeniero electrónico Nelson Trillos, codirector del presente proyecto.

Cuando se diseñó la primera versión del Chatbot, se contaban con pocos avances en este tipo de tecnologías emergentes y poca experiencia en la E3T, por tal razón se proyectó como una forma de entretenimiento para los niños, sin embargo, hoy se tiene otra visión y es que este Chatbot debe ser implementado de manera que pueda ir creciendo como lo hace un niño, aprendiendo cada vez más de las experiencias que se van suscitando y que ofrezca, para así mejorar cada vez más el apoyo, en las interacciones que tienen los niños con dicha plataforma.

Debido a las limitaciones que tiene la plataforma en que se desarrolló la versión preexistente del Chatbot, respecto a los pocos flujos conversacionales que restringe una conversación con mayor complejidad, el presente proyecto de investigación desarrolla un Chatbot para la plataforma transmedia investigadores por naturaleza: expedición Santurbán, que hereda la narrativa de dicha versión y que incorpora flujos conversacionales de mayor complejidad y así mejorar la interacción de los niños con la plataforma transmedia.

Para el desarrollo del Chatbot se realiza un estudio de 5 plataformas, de las cuales se destaca Dialogflow CX por incorporar modelos de inteligencia artificial como NLU/NLP, complejos flujos conversacionales, herramientas de pre-testeo y la posibilidad de incorporar otros servicios de Google Cloud con facilidad.

Para el proceso de construcción de Chatbot, se realizó la identificación de los posibles casos de uso, los cuales en su mayoría son del tipo conversacional, en el cual, utilizando las frases

recolectadas de la versión preexistente, se diseñan los diagramas conversacionales basados en máquinas de estado con la normal UML.

Estos diagramas conversacionales son de alta importancia, ya que cuentan con la arquitectura conversacional necesaria para implementar el Chatbot en la plataforma de desarrollo de Dialogflow CX, que incorpora un entrenamiento previo, sin embargo, con el fin de tener una retroalimentación propiamente por los usuarios antes de pasar a trabajar con los niños, se lleva a cabo un sondeo previo con 10 personas de la comunidad Universitaria, en el cual se determina una exactitud del 96.87%, utilizando una métrica de la tasa de aciertos y recogiendo a su vez, comentarios respecto a las experiencias que tuvieron al interactuar con el Chatbot, que posteriormente fueron tomadas en cuenta y corregidas para las posteriores actividades.

Con el propósito de identificar si el Chatbot estaba cumpliendo las necesidades que habían en los niños y jóvenes respecto a la interacciones proporcionadas respecto al páramo de Santurbán e información de interés, o si por el contrario debían agregarse nuevas interacciones al entrenamiento, se realiza un encuentro de manera remota con 11 niños pertenecientes al Colegio Lizcano Flores, en donde se abarca una presentación introductoria referente a esta tecnología, finalizando con una actividad exploratoria en la que se recoge información para un nuevo entrenamiento.

Al realizar las respectivas integraciones en el entrenamiento del Chatbot, respecto a la actividad anteriormente mencionada, se llevan a cabo dos actividades para determinar el rendimiento del Chatbot, para las cuales se utilizan matrices de confusión binarias y multiclase, en las que para el primer Test de rendimiento se obtuvo una exactitud del 85.45% y en el segundo una exactitud del 97.66%.

Finalmente se logra diseñar un Chatbot con una eficiencia mayor al 80%, el cual satisface las necesidades desde el programa investigadores por naturaleza: Expedición Santurbán, el cual logra aportar a la apropiación social de la ciencia y del conocimiento de una forma más amigable para los niños, respecto a uno de los ecosistemas más importantes como lo es el páramo de Santurbán y también se logra sembrar el precedente de un desarrollo documentado con sus respectivas métricas, del primer Chatbot con narrativa sobre un páramo, lo que implica un gran aporte para la E3T ya que se involucra en el desarrollo de soluciones de nuevas tecnologías emergentes y avanzadas.

## **1. Objetivos**

### **1.1. Objetivo General**

Desarrollar un Bot para apoyar la “Plataforma Transmedia Investigadores por Naturaleza” facilitando la interacción de niños y jóvenes del Páramo de Santurbán.

### **1.2. Objetivos Específicos**

Seleccionar la plataforma en que se desarrollará el Bot en función de las necesidades de la plataforma transmedia investigadores por naturaleza.

Diseñar el flujo conversacional del Bot, basado en diálogos de entrenamiento preexistentes de una versión previa.

Implementar el Bot en la plataforma seleccionada.

Entrenar al Bot orientado a atender nuevas necesidades de los niños de colegios del páramo de Santurbán.

Obtener un mínimo del 80% de exactitud con las interacciones establecidas.

## 2. Marco teórico

En este capítulo se busca abarcar los fundamentos teóricos que están detrás del desarrollo, con el fin de preparar al lector y que entienda con una mayor facilidad, el diseño del Chatbot a proponer en los próximos capítulos y así poder tener una visión más clara de los modelos que se manejan internamente.

Gracias a los múltiples avances tecnológicos en el campo de las ciencias computacionales, se han ido implementando diversos métodos cada más complejos con el fin de poder crear algoritmos inteligentes. Permitiendo en diferentes campos, optimizar procesos en el que antes se requería la presencia de ciertos profesionales. Una de las tecnologías que ha implementado las diferentes ramas de la inteligencia artificial como el Machine Learning y el Deep Learning, son los Chatbot que utilizan la tecnología del procesamiento del lenguaje natural (NLP) que incorpora a su vez entendimiento del lenguaje natural (NLU), siendo entonces los Chatbots una de las herramientas tecnológicas que más aprovechan esta nueva tendencia informática actualmente. (Kasthuri & Balaji, February 2021).

### 2.1. Machine Learning

El Machine Learning o Aprendizaje Automático es un campo de estudio que se deriva de la inteligencia artificial, el cual tiene la capacidad de procesar diferentes datos y mediante la experiencia o aprendizaje previo, le permite a este realizar predicciones de forma discreta o continua. (Géron, 2019). Una forma de verlo a nivel de ingeniería es: “Se dice que un programa de computadora aprende de la experiencia E con respecto a alguna tarea T y alguna medida de desempeño P, si su desempeño en T, medido por P, mejora con la experiencia E.”. (Mitchell, 1997, pag.1-18).

Aunque existen múltiples tipos de Machine Learning, se hablará de grosso modo sobre los más comunes.

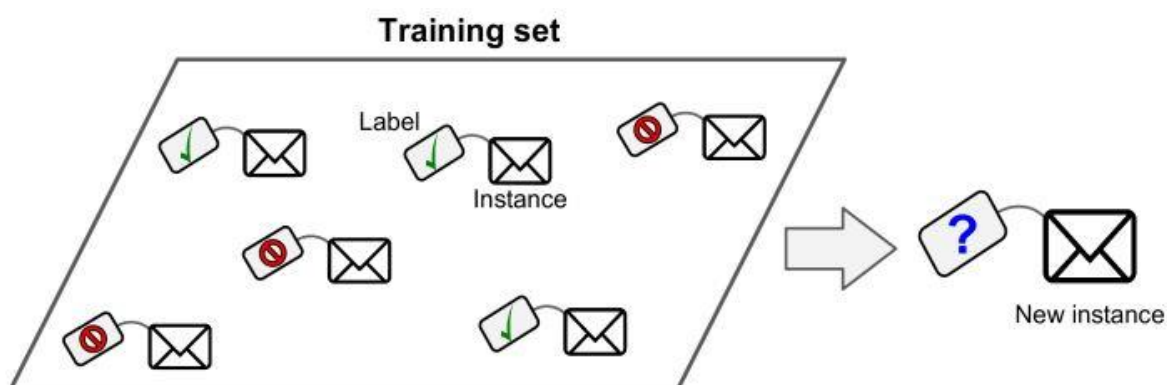
### 2.1.1. Aprendizaje Supervisado

Los algoritmos de Machine Learning supervisados son aquellos en donde se necesita de la intervención humana, para que las etiquetas puedan ser asignadas a los elementos que se utilizarán para el entrenamiento del algoritmo, los cuales incluirían la solución deseada. (Géron, 2019).

Un ejemplo en el que se puede entender el concepto de aprendizaje supervisado es cuando se quiere crear un algoritmo que clasifique si un mensaje es Spam o no, para lo cual una persona de manera previa, entrega una serie de datos de entrenamiento, en donde cada dato viene con su respectiva etiqueta, que permitirá al final con base a la experiencia adquirida, determinar si un mensaje es Spam. (Géron, 2019). Tal como se presenta en la figura 1.

#### Figura 1.

*Ejemplo de clasificación de mensajes Spam*



Nota. El grafico representa un ejemplo de aprendizaje supervisado, en donde se tiene un set de entrenamiento para la clasificación de mensajes spam, el que incluye sus respectivos label con las soluciones deseadas. (Géron, 2019)

**2.1.1.1. Regresión.** Los algoritmos de regresión son un tipo de Machine Learning supervisados, en donde su etiqueta o label es continua, es decir que las predicciones a realizar dan como resultado un valor numérico. (Géron, 2019).

**2.1.1.2. Clasificación.** Los algoritmos de clasificación son un tipo de Machine Learning no supervisado, en donde su etiqueta o label es discreta, es decir que la clasificación realizada, la hace dentro de las diferentes clases ya establecidas. (Géron, 2019).

### **2.1.2. Aprendizaje No supervisado**

Los algoritmos de Machine Learning no supervisados, son aquellos en donde a diferencia de los supervisados, no requieren de la intervención humana, es decir que los datos de entrenamiento no van etiquetados y el sistema trata de aprender sin ayuda. (Géron, 2019).

**2.1.2.1. Clustering.** En el clustering, el algoritmo por sí mismo intenta detectar grupos que compartan características o patrones en común y así poder realizar las diferentes agrupaciones. (Géron, 2019).

**2.1.2.2. Detección anómala.** En la detección anómala, se le enseña al algoritmo en el entrenamiento, instancias que son de situaciones normales y cuando sucede algo extraño dentro de esas instancias, este las marcará como un suceso anormal. (Géron, 2019). Algo común de ver en los bancos, ellos caracterizan cada cliente de acuerdo con lo que normalmente compran con sus tarjetas de crédito y cuando se hace una compra fuera del perfil ya caracterizado, el sistema lo indicara como movimiento extraño y en muchos casos bloqueara la tarjeta de crédito del cliente.

**2.1.2.3. Reducción de dimensionalidad.** Muchas veces en el conjunto de datos adquiridos, puede haber dos o más categorías que se relacionen entre sí y de estas obtener una sola categoría, lo cual se conoce como reducción de dimensionalidad. Lo cual suele ser una buena práctica para aplicar antes de ser enviado a la automatización de datos, ya que esto permitirá que

el algoritmo se ejecute mucho más rápido, optimizando el hardware utilizado. (Géron, 2019). Un ejemplo puede ser el kilometraje de un automóvil respecto a la antigüedad de este, estos pueden reducirse al desgaste de un automóvil.

**2.1.2.4. Asociación de reglas.** En los algoritmos por asociación de regla, el objetivo es tomar una gran cantidad de datos para luego detectar patrones o comportamientos en común y poder así extraer de ellos un tipo de comportamiento. (Géron, 2019). Un claro ejemplo puede ser cuando se le hace seguimiento a una persona en las compras que realiza en un supermercado y con base a eso se establece un comportamiento de compras y así el algoritmo le haga sugerencias al cliente, algo que se aplica constantemente en el mercadeo.

## **2.2. Aprendizaje por refuerzo**

Los algoritmos que utilizan aprendizaje por refuerzo son capaces de observar su entorno y con base a este ejecutar acciones, también se conoce como un entrenamiento por recompensa y castigo. (Géron, 2019). Es decir que en este tipo de entrenamiento el sistema aprende de sus propios errores.

## **2.3. Deep Learning**

El Deep Learning es un tipo de aprendizaje profundo, el cual a diferencia del Machine Learning, en este no se tienen un bloque de extracción de características directamente, ni tampoco requiere aplicar un pre – procesamiento de los datos como si se requiere en Machine Learning, ya que internamente en Deep Learning, los datos pasan por múltiples etapas que se denominan redes neuronales. Las cuales permiten construir modelos complejos a base de modelos más simples, el cual entre más neuronas tiene, se conoce como red neuronal profunda. (Yoshua Bengio, Ian Goodfello, Aaron Courville, 2016).

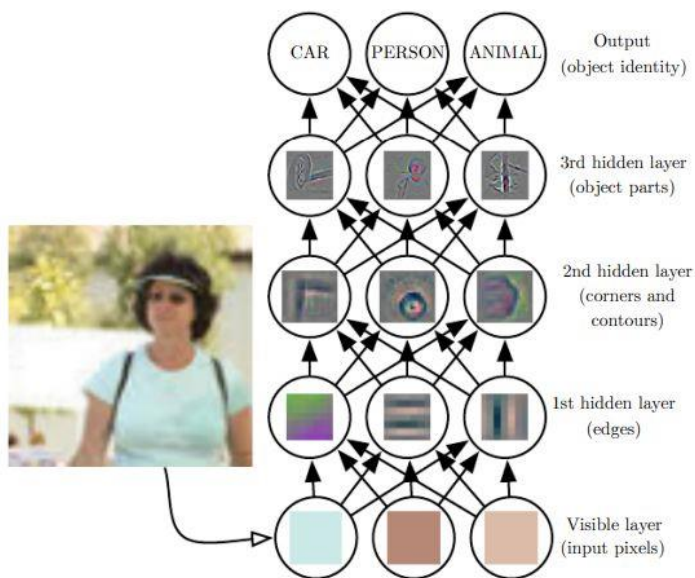
### 2.3.1. Redes neuronales convolucionales

Los conceptos de redes neuronales convolucionales surgieron del estudio de la corteza visual del cerebro, en donde algunas neuronas de nivel superior se basan en las neuronas de nivel inferior, algo que se aplica en las redes neuronales computacionales, que entre más neuronas contiene esta red, da como resultado redes neuronales profundas.(Géron, 2019).

Un claro ejemplo de esto se puede observar en la figura 2, en donde la imagen pasa por diferentes capas que contiene de 3 neuronas de manera respectiva y en cada una de ellas van identificando diferentes parámetros de esta, que al final da como resultado una clasificación que puede corresponder a carros, personas o animales.

**Figura 2.**

*Ilustración de un modelo de Deep Learning*



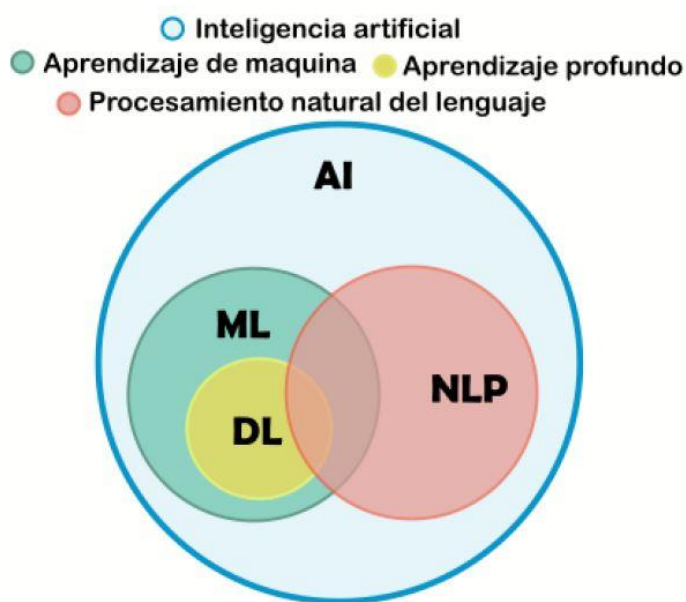
Nota: En el grafico se representa un modelo de Deep Learning que incorpora redes neuronales, el cual en las capas más cercanas a la capa principal extrae características más sencillas y las lejanas extrae las más complejas. (Yoshua Bengio, Ian Goodfello, Aaron Courville, 2016)

## 2.4. Procesamiento del lenguaje natural (NLP)

El procesamiento del lenguaje natural es una rama de la inteligencia artificial, que abarca tanto al Machine Learning como al Deep Learning como se presenta en la figura 3, las cuales aún no han llegado a un punto de automatización total y capacidad de pensamiento humano, por tal razón el NLP es utilizada con el fin de que las maquinas comprendan mucho mejor a los usuarios. (Muhammad et al., 2020). Este tipo de procesamiento tiene como objetivo transformar el texto en una estructura estandarizada, para que el NLU pueda extraer el sentido literal de lo dicho por el usuario. (*NLP vs NLU: What's The Difference?* ).

### Figura 3.

*Diagrama sobre la distribución general del procesamiento natural del lenguaje como parte de la inteligencia artificial.*



Nota: En el grafico se representan los niveles profundos del NLP, donde sus modelos estan basados de Machine Learning y Deep Learning. (Zuleika et al., 2022)

### **2.4.1. Entendimiento del lenguaje natural (NLU)**

El entendimiento del lenguaje natural es un derivado del procesamiento del lenguaje natural, el cual tiene como función identificar el contexto e intención de lo descrito por un usuario, el cual ya fue procesado por el NLP, lo que permite captar la idea principal de las preguntas realizadas de diferente forma. (*NLP vs NLU: What's The Difference?* ).

De acuerdo con lo mencionado respecto a NLU y NLP, es importante aclarar, que ambos son necesarios al momento de implementar Chatbots, ya que estos se complementan para mejorar la comunicación entre usuario - maquina.

## **2.5. Chatbot**

Los chatbots son considerados agentes conversacionales, pensados para interactuar con las personas, ya que incorporan NLP y NLU en algunos casos, lo cual les permite reducir la barrera hombre – máquina considerablemente. (Muhammad et al., September 2020).

### **2.5.1. Basados en caja de texto (chatterboxes)**

Los chatbots basados en cajas de texto, son aquellos en donde la interacción con el usuario se da por medio de entradas y salidas de texto o de voz, incorporando procesamiento de lenguaje natural el cual acorta la brecha comunicativa en la interacción entre una persona y un Chatbot.(Garcia Brustenga et al., 2018).

### **2.5.2. Asistentes virtuales personificados (embodied conversational agents)**

Con el fin de poder reducir la sensación de que se está hablando con una máquina, se suele utilizar una interfaz que es representada con la figura de un cuerpo o de una cara, es decir como un tipo de avatar, utilizando a su vez múltiples herramientas audiovisuales y de multimedia para mejorar la experiencia con el usuario. (Garcia Brustenga et al., 2018).

### **2.5.3. Físicos**

Existe un tercer tipo de chatbots los cuales son representados con un robot físico, el cual incorporara múltiples periféricos para poder comunicarse con el usuario. (Garcia Brustenga et al., 2018).

### **2.5.4. Terminología general en Chatbots**

Para la implementación abordada en capítulos posteriores es de vital importancia conocer acerca de la terminología de implementación que importante conocer al momento de implementar Chatbots.

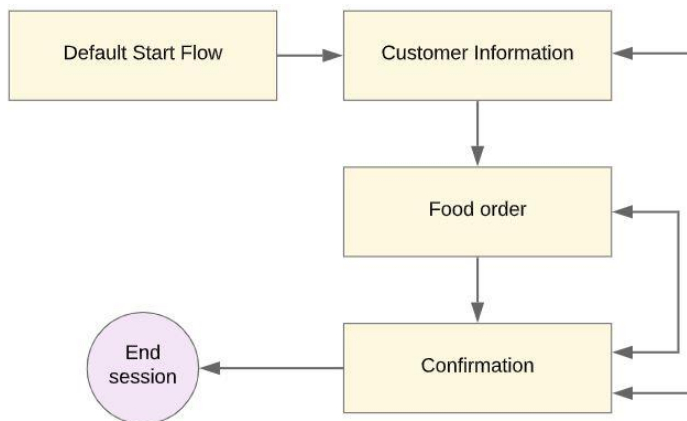
**2.5.4.1. Agentes.** Dialogflow Cx, permite la construcción de agentes que comúnmente podrían conocerse como un tipo de asistentes virtuales especializados, los cuales controlarían el proceso de conversación con los usuarios finales. Estos agentes incorporan en sus modelos el procesamiento del lenguaje natural, el cual permite detectar los diferentes matices del lenguaje humano. (Google Cloud, 2022).

**2.5.4.2. Flujos.** Muchas veces el poder entablar una conversación estructurada con una persona conlleva una serie de preguntas, sin olvidar las ya realizadas, lo mismo sucede en los agentes de Dialogflow CX que con el fin de poder entablar una conversación más compleja, se requiere de diferentes flujos conversacionales.

Estos flujos conversacionales son los que permitirán entablar conversaciones más elaboradas con el agente, un claro ejemplo es el caso de un repartidor de comidas rápidas, donde se realiza el pedido, se crea la orden, se confirma y esta confirmación llega al cliente y al negocio como se observa en la figura 4, que se logra apreciar que una solicitud implica una rama conversacional con varias preguntas que dependen siempre de la ya realizadas. (Google Cloud, 2022a)

**Figura 4.**

*Diagrama de flujos en Dialogflow CX*



Nota: Documentación Google Cloud, 2022,

<https://cloud.google.com/dialogflow/cx/docs/concept/flow?hl=es-419>

**2.5.4.3. Páginas.** Una forma de describir el funcionamiento de cómo interactúan los flujos conversacionales en una sesión, es con las máquinas de estado, en la cual existe un sistema con cierta entrada y salida, donde el cambio entre estados o más páginas estará a la espera de una respuesta por parte del usuario que permitirá internamente saltar entre flujos conversacionales. (Google Cloud, 2022)

**2.5.4.4. Tipos de entidad.** Tomando en cuenta que los modelos de procesamiento del lenguaje natural intentan extraer ciertas características de lo que se dice, Dialogflow CX incorpora lo que se conoce como entidades, las cuales extraen ciertos datos del usuario, un ejemplo de las que se trae por defecto en la plataforma, son las que identifican fechas, horas, colores, ubicaciones entre otras y a su vez permite crear entidades personalizadas. (Google Cloud, 2022)

**2.5.4.5. Parámetros.** De acuerdo con el enfoque del agente, habrá algunos que requieran tomar información puntual del usuario, para que el agente pueda procesarla ya sea para almacenar

o buscar en una base de datos, para almacenar de forma temporal la información y así hacer referencia a este valor durante la sesión o ejecutar sentencias condicionales, entre otras, para esto Dialogflow incluye una funcionalidad de parámetros permite capturar dichos valores. (Google Cloud, 2022)

**2.5.4.6. Intentos.** Los intentos son los que realizan la clasificación de la solicitud del usuario en cada turno de la conversación, en el caso de Dialogflow CX, permite reutilizar estos intentos cuantas veces sea necesario en diferentes flujos conversacionales, a diferencia de su hermano menor Dialogflow ES, donde cada intento está ligado a una respuesta ya que como se mencionó en capítulos anteriores, las ramas conversacionales son limitadas en dicha versión. (Google Cloud, 2022)

Estos intentos al momento de configurarse estarán ligados a las frases de entrenamiento, las cuales mediante los modelos de NLP que incorpora, al momento de recibir alguna entrada estará intentando identificar el intento basado a su entrenamiento, es de resaltar que no es necesario definir todos los ejemplos posibles en el entrenamiento, ya que Dialogflow expande dicha base con frases similares que incorpora internamente.

**2.5.4.7. Webhook.** Los Webhook son un servicio que se encuentra en Dialogflow CX, que permite realizar consultas a servidores asociados como por ejemplo una base de datos, en la cual se podría cargar información o consultar información de interés del usuario, como se mencionó anteriormente, esto dependería del enfoque del agente. (Google Cloud, 2022)

**2.5.4.8. Entrega.** El agente de Dialogflow al final tiene el objetivo de entregar una respuesta final al usuario por cada solicitud, en algunos casos estas respuestas pueden estar ligadas a bancos de preguntas con respuestas únicas y diagramas conversacionales, pero muchas veces pueden estar ligadas a consultas de bases de datos o servidores externos que lo hace por medio de

Webhook asociadas o incluso depender de algún parámetro que se capturará en la conversación. (Google Cloud, 2022)

**2.5.4.9. Integraciones.** Dialogflow CX, a diferencia de su hermano menor, cuenta con una lista limitada de integraciones que están asociadas a la ubicación de donde se va a desarrollar, esta integración permite que otras plataformas llamen a la API de Dialogflow, como por ejemplo WhatsApp, Messenger, Telegram, páginas web, llamadas con agentes externos entre otros. Para el presente desarrollo se utilizará especialmente la integración de páginas web.

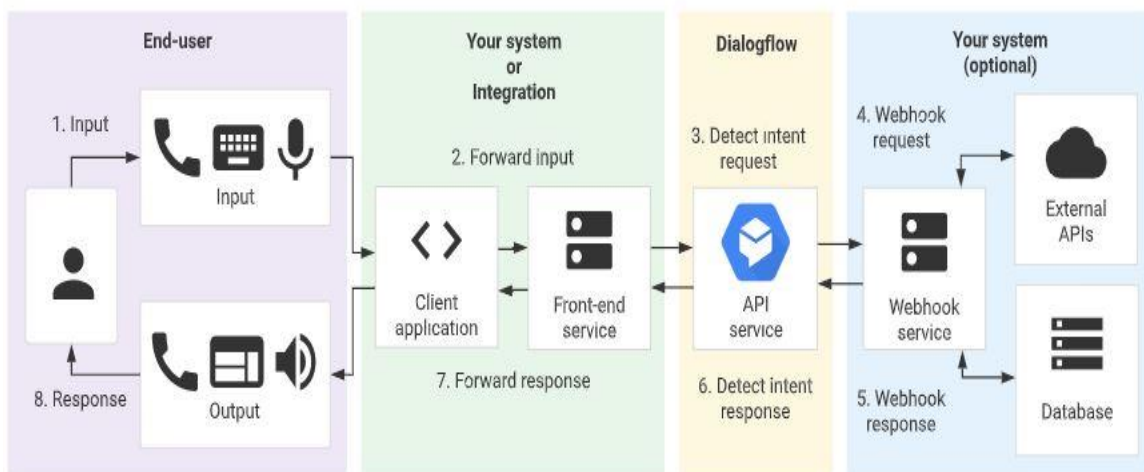
**2.5.4.10. Interacción.** En Dialogflow CX, cada que se realice una pregunta, se conoce como una interacción, donde en cada interacción tiene dos alternativas al momento de recibir la información por parte del usuario final, utilizar la API directamente de Dialogflow o utilizar integraciones.

Como se observaría en la figura 11. Cuando se utiliza la API, directamente se utiliza todo lo alojado en el agente, flujos conversacionales, parámetros y solicitudes a Webhook, pero no estaría alojado en ningún lugar, algo que ocasionalmente no ocurre, dicho proceso solo abarcaría la sección del End-user, Dialogflow y Webhook.

Cuando se utiliza una integración se puede abarcar un tipo de interfaz para el usuario, como es el ejemplo de las páginas web, la cual establece la conexión con el usuario, proporciona una interfaz y el directamente hace las solicitudes a la API de Dialogflow, como se observa en la figura 5, donde si incorpora todo el flujo que se sigue en una interacción. (Google Cloud, 2022)

**Figura 5.**

*Grafica de interacciones con API*



Nota: Documentación Dialogflow Cx, 2022,

<https://cloud.google.com/dialogflow/cx/docs/quick/api?hl=es-419>

## 2.6. Métricas de rendimiento

Las métricas de rendimiento son aquellas que son utilizadas para evaluar que tan eficiente es un modelo de inteligencia artificial, ya sea de Deep Learning, Machine Learning, NLU entre otros, que para su selección, dependerán de la aplicación de los modelos. (Ricardo Borja-Robalino et al., 2020).

### 2.6.1. Tasa de aciertos

La métrica de tasa de aciertos hace parte de las métricas rápidas y sencillas a utilizar al momento de querer calcular la exactitud de un modelo, la cual se calcula con el número de respuestas correctas de todos los usuarios, sobre el total de preguntas realizadas a todos. (Carlos F. Gonzales, 2020)

### 2.6.2. Matriz binaria

Para clasificar la información en esta matriz, el número de suposiciones verdaderas y falsas se resume con valores de conteo y se desglosa por cada clase especificada en la matriz, en esta matriz se puede determinar la forma en que el modelo de clasificación se confunde al momento de hacer la predicción y también permite determinar el error del modelo entrenado. (Das & Kumar, Dec 2018)

Para clasificar la información dentro de la matriz, es importante definir los verdaderos positivos (TP) que surgen cuando la observación es positiva, pero se clasifica como positiva, los falsos negativos (FN) que surgen cuando la observación es positiva, pero se clasifica como negativa, los negativos verdaderos (TN) que surgen cuando la observación es negativa, pero se clasifica como positiva y los falsos positivos (FP) que surgen cuando la observación es negativa pero se predice como positiva. (Das & Kumar, Dec 2018).

Para este tipo de métricas se puede calcular la exactitud, precisión, recall y F1-Score.

**2.6.2.1. Exactitud.** Esta métrica permite conocer el porcentaje de predicciones correctas. (Das & Kumar, Dec 2018)

$$Exactitud = \frac{TP + TN}{TP + TN + FP + FN}$$

**2.6.2.2. Precisión:** Permite determinar el porcentaje de predicción de la totalidad de los datos que fueron indicados como correctos. (Das & Kumar, Dec 2018)

$$Precisión = \frac{TP}{TP + FP}$$

**2.6.2.3. Recall.** Permite determinar el porcentaje de datos que realmente eran positivos. (Das & Kumar, Dec 2018)

$$Recall = \frac{TP}{TP + FN}$$

**2.6.2.4. F1-score.** Esta métrica permite conocer es un balance entre la precisión y el Recall en donde la métrica incrementa a medida que los falsos negativos y positivos disminuyen. (Das & Kumar, Dec 2018)

$$F1 - Score = \frac{TP}{TP + \frac{FP + FN}{2}}$$

### 2.6.3. Matriz multiclase

Las matrices de confusión multiclase, a diferencia de las binaria pueden llegar a clasificar un dato dentro de N clases, lo que implica tener una matriz de confusión de N x N, ya que se debe tener tantas etiquetas verdaderas como de predicción, normalmente este tipo de métricas las calculo por si solo el software debido a lo grande que pueden llegar a ser. (M. Heydarian et al., 2022).

**2.6.3.1. Exactitud.** Está métrica permite conocer el porcentaje de predicciones correctas, a diferencia de la matriz binaria, para calcular la exactitud del modelo se emplea la siguiente formula. (Arista-Jalife et al., 2017). M representa el número de clases.

$$Exactitud = \frac{\sum_{i=1}^M TP_i}{\text{número de muestras presentadas a clasificar}}$$

**2.6.3.2. Precisión:** Permite determinar el porcentaje de predicción de la totalidad de los datos que fueron indicados como correctos. Para calcular la exactitud del modelo se emplea la siguiente formula, la cual varía un poco respecto a la binaria. (Arista-Jalife et al., 2017). M representa el número de clases.

$$Precisión = \frac{\sum_{i=1}^M \frac{TP_i}{TP_i + FP_i}}{M}$$

**2.6.3.3. Recall.** Permite determinar el porcentaje de datos que realmente eran positivos. (Arista-Jalife et al., 2017). Para calcular la exactitud del modelo se emplea la siguiente formula, la cual varia un poco respecto a la binaria.  $M$  representa el número de clases.

$$Recall = \frac{\sum_{i=1}^M \frac{TP_i}{TP_i + NP_i}}{M}$$

**2.6.3.4. F1-score.** Esta métrica permite conocer es un balance entre la precisión y el Recall en donde la métrica incrementa a medida que los falsos negativos y positivos disminuyen. (Arista-Jalife et al., 2017). Para calcular la exactitud del modelo se emplea la siguiente formula, la cual varia un poco respecto a la binaria.  $M$  representa el número de clases.

$$F1 - score = \frac{\sum_{i=1}^M \frac{TP_i}{TP_i + \frac{FP_i + FN_i}{2}}}{M}$$

### 3. Marco de referencia

En este capítulo se busca presentar los diferentes desarrollos realizados recientemente utilizando Chatbots y así poder tener una idea hasta donde han llegado los recientes desarrollos en Chatbots y tener un punto de partida respecto a las plataformas que se suelen usar para sus implementaciones.

Desde que se empezó a emplear el procesamiento de lenguaje natural en los ya conocidos Chatbot, los cuales permiten un mayor acercamiento con el usuario, se han ido implementando en diferentes aplicaciones que tienden a ser repetitivas y en muchos casos, el personal no da abasto con todas las solicitudes por parte de los usuarios, existiendo la necesidad de implementar un asistente virtual especializado, es decir un Chatbot.

### **3.1. PATH**

PATH es un Chatbot inteligente el cual fue creado con fines turísticos, que tiene como función principal, guiar a los visitantes en la Universidad de Johannesburgo, indicando diferentes rutas para llegar a diferentes dependencias de la Universidad, resaltando un 90% de usabilidad y un 80% de utilidad para la comunidad, dicho Chatbot fue desarrollado en la plataforma Dialogflow ES que pertenece a Google Cloud. (Mabunda & Ade-Ibijola, November 2019).

### **3.2. CheMBL**

CheMBL es un Chatbot inteligente encargado de realizar una consulta detallada de bioactividades basada en ciertos compuestos (De usos medicinales, farmacología, entre otros) alojados en diferentes bases de datos químicas, optimizando las múltiples tareas en la práctica que antes eran realizadas por otras personas, dicho Chatbot fue desarrollado en la plataforma Dialogflow Es de Google Cloud, la cual permite incorporar NLP para así poder mejorar la interacción con el usuario y poder emplear una comunicación más informal. (Murali et al., December 2018).

### **3.3. Chatbot como herramienta para el aprendizaje de un nuevo idioma**

Actualmente el inglés es uno de los idiomas más hablados a nivel mundial, razón por la cual se hace importante aprenderlo, sin embargo, para muchas personas suele ser una tarea bastante difícil, pero debido al crecimiento tecnológico en los chatbots inteligentes, ha sido posible desarrollar un Chatbot conversacional que incorpora reconocimiento de voz y texto, desarrollado en Dialogflow ES de Google Cloud, lo que lo ha convertido en una herramienta de bastante utilidad para los estudiantes de inglés, siendo utilizada para practicar. Dicho agente conversacional cumplió con todas las expectativas con que se desarrolló, obteniendo así una tasa de precisión del 100%. (Muhammad et al., September 2020).

### **3.4. Chatbot utilizado en el campo de la educación**

Si bien es cierto que con el paso de los años se han ido desarrollando en gran medida la forma en cómo se transmite el conocimiento a los estudiantes, se ha observado que sigue siendo bastante limitado, debido a que los estudiantes solo tienen acceso a la enseñanza impartida por un docente por tiempo limitado cada día, por tal razón se desarrolló un Chatbot inteligente en la plataforma Discord, el cual le permite a cada estudiante, poder hablar con él las 24 horas del día y los 7 días de la semana, en el cual se espera que a futuro incorpore información detallada sobre los circuitos electrónicos y temáticas generales a las telecomunicaciones, sin embargo, durante el desarrollo se apreció que en esta plataforma aún está en crecimiento el procesamiento de lenguaje natural (NLP), comparado a otros ámbitos como lo es la medicina o los negocios que tienen mayor avance, llegando a la conclusión que a pesar de esta brecha que hay actualmente entre los Chatbots realizados en múltiples áreas respecto al campo de la educación, esta sigue prometiendo a futuro. (P et al., April 2021).

### **3.5. Florence**

Muchas veces debido a lo molesto que puede llegar a ser tener que esperar para poder acceder a una consulta médica, ya sea por la cantidad de gentes o el tiempo que hay que esperar, se desarrolló a Florence, un Chatbot inteligente con el fin de poder facilitar que las mismas personas puedan controlar su salud y que se acerque al servicio ofrecido de forma convencional, dicho Chatbot identificaría las enfermedades más probables con base a múltiples rutas conversacionales y poder indicar el mejor tratamiento recomendado, comportándose como un tipo de medico virtual que aunque no reemplazaría la medicina convencional si optimizaría el servicio en el diagnóstico de enfermedades comunes. (Gupta et al., March 2021).

### **3.6. Trabajo realizado en la Universidad Industrial de Santander**

En la E3T se implementó una plataforma transmedia del programa Investigadores por naturaleza: Expedición Santurbán, este programa ha sido impulsado por parte de la Vicerrectoría de investigación y extensión de la Universidad Industrial de Santander que fue diseñado con el fin de fomentar la vocación científica y la apropiación social del conocimiento en niños, niñas y jóvenes habitantes del páramo de Santurbán.

El proyecto transmedia, cuenta con la participación de los grupos de investigación de la Universidad Industrial de Santander tales como: CICAT, EDUMAT, CUYNACO, RADIOGIS y Grupo HALLEY, de los cuales el grupo de investigación Radiogis, tuvo la tarea de desarrollar un Chatbot en apoyo a dicha plataforma.

El Chatbot de la versión preexistente, fue desarrollado por el grupo de investigación Radiogis, con el fin de entretener a los niños que interactuaban con la plataforma transmedia, la cual seguía la narrativa de ser el páramo de Santurbán, que acompañaría a los niños, junto a los 5 personajes diseñados por el programa investigadores por naturaleza: Canis, Meyer, Chi Chuan Pan, Guardián del páramo y SENTIRE. Dicho Chatbot fue implementado en la plataforma de desarrollo de Dialogflow ES de Google Cloud y la recopilación de las frases de entrenamiento fueron recolectada por los demás grupos de investigación para su posterior entrenamiento.

De acuerdo con la recopilación anteriormente presentadas, de algunos de los avances que incorporan Chatbot incluso el desarrollado en la Universidad Industrial de Santander, se puede resaltar como ha impactado y beneficiado el uso de chatbots en los diferentes campos de aplicación y en donde el continuo crecimiento de la inteligencia artificial sigue abriendo paso a una mayor usabilidad y aprovechamiento de estos.

#### **4. Requerimientos y restricciones de diseño para el Chatbot**

El presente capítulo tiene como objetivo proporcionarle al lector los diferentes requerimientos y restricciones que se deben cumplir y tener en cuenta para el desarrollo de la nueva versión del Chatbot a desplegar en la plataforma Investigadores por naturaleza, el cual hereda las tomadas a consideración en la primera versión.

- Identificación de los casos de uso en función de las necesidades de los usuarios
- Mantener la narrativa en la cual el Chatbot es el ente representativo del páramo de Santurbán, tal como la versión previamente desarrollada.
- El caso de uso a desarrollar corresponde al conversacional
- Reutilizar la base de entrenamiento con la que se entrenó la versión previamente desarrollada.
- Identificar nuevas necesidades de los usuarios finales, que son los niños y jóvenes del páramo de Santurbán que se tendrán en cuenta para su posterior entrenamiento.
- El Chatbot desarrollado debe estar en la capacidad de poderse integrar fácilmente a la página Web ‘Investigadores por naturaleza’.
- El Chatbot estará en la capacidad de atender y responder a las solicitudes de los usuarios únicamente mediante entradas y salidas de texto de acuerdo con las interacciones establecidas.
- La exactitud del Chatbot será de mínimo el 80% respecto a las interacciones establecidas.

#### **5. Selección de plataforma**

En este capítulo se busca presentarle al lector, el proceso de selección de plataforma en la que se implementará el Chatbot, que como se pudo observar en los diferentes proyectos abordados en el capítulo del 3, se utilizaron diferentes plataformas de desarrollo, de las cuales resaltó

Dialogflow ES, por lo cual el primer paso que se abordará para el diseño del Chatbot propuesto, corresponde a la selección de plataforma, tomando a consideración el capítulo anterior, en donde se describen los requerimientos de éste.

## **5.1. Dialogflow**

Dialogflow es una plataforma de desarrollo de chatbots de Google, la cual incorpora procesamiento y entendimiento del lenguaje natural que permite diseñar diferentes interfaces de usuario y que dependiendo la versión a utilizar permite incorporar una mayor o menor opción de ramas conversacionales. Al ser una plataforma especializada en el desarrollo de chatbots ofrece la facilidad de su incorporación de forma rápida en dispositivos móviles, aplicaciones web entre otros.

Dialogflow ofrece 2 tipos de agentes virtuales, Dialogflow ES y Dialogflow CX que ofrecen la posibilidad de analizar la información proporcionada por el usuario por texto o audio y así mismo para poder responderle al usuario.

### **5.1.1. Dialogflow ES**

Dialogflow ES, es un agente virtual de la plataforma de Dialogflow de Google, el cual proporciona las funcionalidades estándar, pensado en utilizarse en donde se requieran agentes pequeños y simples. (Google Cloud, 2022b)

#### **5.1.1.1. Características**

- Entrada de texto y audio
- Respuesta de texto y audio
- Incorpora comunicación con múltiples API's (No incluye MySQL)
- Incorpora NLP
- Soporte para agentes multilinguaje

- Permite ejecutar el agente en el Backend de una página web con facilidad
- Permite incorporar código en múltiples lenguajes como Python, java, C++, Node.js, PHP entre otros.

**5.1.1.2. Costos Dialogflow ES.** Es una versión completamente gratuita que incorpora funcionalidades básicas respecto a su hermano mayor, Dialogflow CX el cual si incorpora tarifas por interacción.

### **5.1.2. Dialogflow CX**

Dialogflow CX es un agente virtual de la plataforma de Dialogflow de Google, el cual proporciona funcionalidades más avanzadas respecto a la versión ES, pensado para ser utilizado en donde se requieran agentes grandes o muy complejos. (Google Cloud, 2022)

#### **5.1.2.1. Características**

- Entrada de texto y audio
- Respuesta de texto y audio
- Incorpora nuevas comunicaciones con múltiples API's (Incluye MySQL)
- Incorpora ilimitados giros conversacionales, lo que permite implementar flujos conversacionales complejos.
- Soporte para agentes multilinguaje
- Incorpora NLU y NLP
- Permite embeber el agente para poder integrarlo a una página web con facilidad
- Incorpora una funcionalidad para realizar testeos de manera previa
- Permite incorporar código en múltiples lenguajes como Python, java, C++, Node.js, PHP entre otros.

**5.1.1.2. Costos.** Respecto a los precios que se pueden manejar en Dialogflow CX, dependerá respecto a las interacciones que se hagan y si estas serán netamente texto o incorporará las funcionalidades de audio, también dicho costo estaría asociado a si se incorpora el servicio de llamadas. Al acceder al servicio por primera vez la plataforma de Google regalará un acceso de 12 meses totalmente gratuito y luego de dicho tiempo se tendrían que asumir los costos presentados en la tabla 1.

**Tabla 1**

*Precios para la plataforma Dialogflow CX*

| Concepto  | Total [\$] (Pesos colombianos / interacción) |
|---|--|
| <p><b>Texto</b><br/>(incluidas todas las solicitudes DetectIntent, StreamingDetectIntent, FulfillIntent, AnalyzeContent y StreamingAnalyzeContent Texto (incluidas todas las solicitudes DetectIntent, StreamingDetectIntent, FulfillIntent, AnalyzeContent y StreamingAnalyzeContent que no contienen audio)).</p> | 26.5   |
| <p><b>Entrada y salida de audio</b><br/>(reconocimiento, transcripción, síntesis de voz, conversión de texto a voz y telefonía)</p>   | 228  |
| <p><b>Solicitudes durante el diseño</b><br/>Por ejemplo, llamadas para crear o actualizar un agente. También se incluyen las solicitudes de tiempo de ejecución de audio y texto enviadas por el simulador de la consola de Dialogflow.</p>   | 0  |
| <p><b>Otras solicitudes durante las sesiones</b><br/>Por ejemplo, para configurar u obtener entidades de la sesión.</p>   | 0  |

Nota: En la tabla se representan los precios recopilados para la plataforma Dialogflow CX de acuerdo con las interacciones del usuario. (Google Cloud, 2022)

## **5.2. Microsoft – Azure Bot Service**

Azure Bot Service es una herramienta brindada por los desarrolladores de Microsoft que proporcionan un entorno de desarrollo para el diseño y creación de inteligencia artificial conversacional aplicada a chatbots. (Microsoft, 2022)

Azure Bot Service incluye múltiples herramientas para el desarrollo de inteligencia artificial y procesamiento de lenguaje natural que facilitan la comunicación con el usuario. También integra una plataforma adicional que permite a los desarrolladores crear Bots sin necesidad de código. (Microsoft, 2022)

### **5.2.1. Características**

- Proporciona un lienzo para la creación visual y a su vez permite incorporar código abierto.
- Inteligencia artificial de primer nivel, la cual incorpora características de lenguaje natural, voz y visión de alta calidad.
- Datos seguros, alto control de datos privados.
- Integraciones rápidas con los servicios asociados a Microsoft y otros aliados.
- Enfocado a grandes corporaciones y organizaciones.

### **5.2.2. Costos**

La plataforma de Microsoft proporcionaría un acceso gratuito por 12 meses, en los cuales se tiene acceso gratuito a múltiples bases de datos de fácil integración, una vez finalizado el periodo de prueba, cada mes regalaría ciertas cantidades de espacio en la nube, los cuales, si se excede la capacidad, se tendría que acceder a un servicio de pago.

Una vez finalizado el periodo de prueba se podría acceder a múltiples servicios los cuales dependerían de la cantidad de mensajes que se manejarían por mes como se presenta en la tabla 2, adicionalmente habría otro tipo de servicios asociados que también tendrían costes adicionales cuando las interacciones se realizan por voz o chat, acceso al procesamiento del lenguaje natural el cual vararía si es de aplicación a gran escala, para cotizaciones específicas se debe comunicar con asesor especializado de Microsoft.

**Tabla 2**

*Precios de Azure Bot Service*

|                  | Gratis                | S11                         |
|------------------|-----------------------|-----------------------------|
| Canales Estándar | Mensajes ilimitados   | Mensajes ilimitados         |
| Canales Premium  | 10.000 mensajes / mes | \$1959 (pesos) por mensajes |

Nota: En la tabla se incorporan los precios a tener en cuenta en caso de querer acceder al servicio, sin embargo, de acuerdo al uso puede solicitar hablar con un especialista.  
<https://azure.microsoft.com/es-es/pricing/details/bot-services/>

### **5.3. IBM - Watson**

IBM – Watson es una plataforma en la cual se pueden desarrollar chatbots del tipo asistentes inteligentes, que basa sus modelos en el procesamiento de lenguaje natural que se utiliza para poder comprender preguntas, buscar o encontrar la mejor respuesta para las peticiones del usuario. Watson incorpora un interfaz gráfico con el fin de reducir la dificultad a la hora de la creación de sus Bots. (IBM, 2022)

### **5.3.1. Características**

- Sus modelos están basados en el procesamiento de lenguaje natural.
- Como entradas y salidas se puede utilizar texto o audio.
- Incorpora integraciones de servicios externos con aliados para mejorar la experiencia de usuario.
- Para el desarrollo de los Bots ofrece una interfaz gráfica intuitiva.
- Se puede integrar en dispositivos móviles con gran facilidad.

### **5.3.2. Costos**

Para acceder a la plataforma de desarrollo de IBM se cuenta con un acceso gratuito que incorpora diferentes soluciones para poder empezar a implementar el Bot, sin embargo, sigue siendo limitada por lo cual ofrecen una versión de pago mensual por \$140 (dólares) que incorpora todos los servicios e integraciones y también cuenta con una membresía de pago por uso la cual se cotiza directamente con los asesores especializados de IBM. (IBM, 2022)

## **5.4. Amazon Lex**

Amazon Lex es un servicio para el desarrollo e implementación de Chatbots que incorporan inteligencia artificial en sus modelos, aprovechando del procesamiento natural para mejorar la experiencia conversacional con sus usuarios. Amazon Lex destaca por sus aplicaciones en los servicios financieros, seguros, ventas, telecomunicaciones y viajes. (Amazon Web Service, 2022)

### **5.4.1. Características**

- Modelos basados en el procesamiento de lenguaje natural
- Integra Amazon Kendra que se utiliza para la búsqueda inteligente de gran precisión sobre algún documento y también Amazon Polly que utiliza la Conversión de texto a audio, permitiendo crear una mayor interacción con los usuarios.

- No tiene implementación directa en sitios Web.

(Amazon Web Service, 2022)

#### 5.4.2. Costos

Amazon Lex ofrece un periodo de prueba por un año y una vez termina dicho periodo de prueba, se pagaría un total de 33,5 USD por mes, el cual incluiría costos por solicitud de voz y texto como se muestra en la tabla 3 y en eventuales usos de conversaciones en Streaming un valor de 56,00 USD al mes como se muestra en la tabla 4.

**Tabla 3**

*Precio de Amazon lex - Precios de interacción de respuesta y solicitud*

| Solicitudes de entrada                   | Costos por unidad | Número de unidades | Totales   |
|--|-------------------|--------------------|-----------|
| 8000 intervalos de habla                 | 0,004 USD         | 8000 intervalos    | 32,00 USD |
| 2000 solicitudes de texto                | 0,00075 USD       | 2000 solicitudes   | 1,50 USD  |
| Cargos totales de Amazon Lex para el mes |                   |                    | 33,50 USD |

Nota: En la tabla se describen los precios relacionados a la interacción de las respuestas y las respectivas solicitudes. <https://aws.amazon.com/es/lex/pricing/>

**Tabla 4***Precio de Amazon Lex - Precios de la conversación en Streaming*

| Solicitudes de entrada                   | Costos por unidad | Número de unidades | Totales   |
|--|-------------------|--------------------|-----------|
| 8000 intervalos de habla                 | 0,0065 USD        | 8000 intervalos    | 52,00 USD |
| 2000 solicitudes de texto                | 0,0020 USD        | 2000 solicitudes   | 4,00 USD  |
| Cargos totales de Amazon Lex para el mes |                   |                    | 56,00 USD |

Nota: En caso de querer soporte especializado se puede contactar con un asesor.

<https://aws.amazon.com/es/lex/pricing/>

### **5.5. Criterio de selección de plataforma**

Como se pudo presentar anteriormente, existen múltiples plataformas especializadas para el desarrollo e implementación de chatbots, cada uno con sus beneficios y precios.

Asure Bot Service, aunque ofrece una gran cantidad de servicios para mejorar la experiencia, está pensado para ser utilizado a gran escala y a nivel empresarial, además se recomienda que sea utilizada en un ecosistema que se utilicen más programas de Microsoft, por lo cual se descartaría esta opción ya que el actual proyecto es de pequeña escala y con propósitos educativos.

IBM Watson ofrece una fácil incorporación a dispositivos móviles y páginas web, sin embargo, es la opción más costosa de todas y tienen un enfoque en donde toma información de

documentos para entregar información al usuario, por lo cual también se descarta ya que se busca algo con enfoque conversacional.

Amazon Lex es la plataforma quizá la más económica, después de Dialogflow ES que es gratuita, sin embargo, no se especializa en aplicaciones sociales, sino más bien en el ámbito de negocios, viajes entre otros y tampoco cuenta con una gran variedad de herramientas para mejorar la experiencia del Chatbot.

Por último, se tienen las dos plataformas derivadas de Dialogflow que es la ES y la versión CX, donde la versión ES, es la versión actual del Chatbot, la cual debido a que es gratuita tiene bastantes limitaciones, como lo es la falta de una interfaz gráfica, falta de giros conversacionales, a diferencia de su hermano mayor Dialogflow CX, que es una plataforma con una interfaz gráfica bastante intuitiva, también cuenta con la posibilidad de implementar giros conversacionales complejos, es pensada para abordar una mayor cantidad de casos de uso, tiene diferentes herramientas para poder testear de forma rápida al agente inteligente, respecto a los flujos implementados y así tener una mayor seguridad de su funcionamiento, para luego poder ser utilizado por los propios usuarios.

Por tal razón Dialogflow CX fue la plataforma seleccionada para realizar la nueva versión del Chatbot a desplegar en la plataforma Investigadores por naturaleza, que incorpora un enfoque conversacional, herramientas avanzadas para el desarrollo de Bots entre otros.

## **6. Recopilación de preguntas para la nueva versión del Chatbot**

De acuerdo con lo mencionado en el capítulo 4, uno de los requerimientos de diseño a cumplir, menciona que la versión a desarrollar debe incorporar como base de entrenamiento, todas las frases utilizadas en el Chatbot de la primera versión desarrollada.

La versión actualmente del Chatbot desplegada en la plataforma Investigadores por naturaleza, cuenta con 119 frases de entrenamiento, las cuales fueron recopiladas y que podrán ser consultadas en el Apéndice A, estas frases serán incorporadas en el entrenamiento preliminar de la segunda versión desarrollada como se explicará en el capítulo 8 y serán tomadas en cuenta para el diseño de los diagramas conversacionales próximos a presentar.

## **7. Casos de uso**

En este capítulo se busca presentar la identificación de casos de uso al lector, que fueron tomados en cuenta que para el posterior diseño de los diagramas conversacionales que se utilizaron para la implantación del Chatbot.

Aunque dentro de las restricciones mencionadas en capítulos anteriores, en que el caso de uso de enfoque sería el conversacional, es bastante importante identificar si existen otros casos de uso, que podrían ser tomados en cuenta para futuros trabajos que se hagan sobre el Chatbot desarrollado como se observa en la figura 6.

### **7.1. Rol de desconocido**

Para el caso del rol de usuario desconocido, este estaría en la posibilidad de registrarse a las bases de datos que conectarían con el sitio web transmedia, podría iniciar sesión o solamente interactuar con el chat conversacional del Bot.

### **7.2. Rol de estudiantes del semillero Radiogis**

En el caso del rol de estudiante del semillero Radiogis, únicamente tendrían acceso los estudiantes que perteneces a este semillero, en el cual una vez inician sesión, podrían realizar sugerencias de entrenamiento para que el desarrollador del Chatbot, el cual podría recolectar mediante una base de datos sugerencias de entrenamiento y posteriormente pueda realizar los respectivos ajustes y entrenar de forma interna, los usuarios también podrían realizar consultas

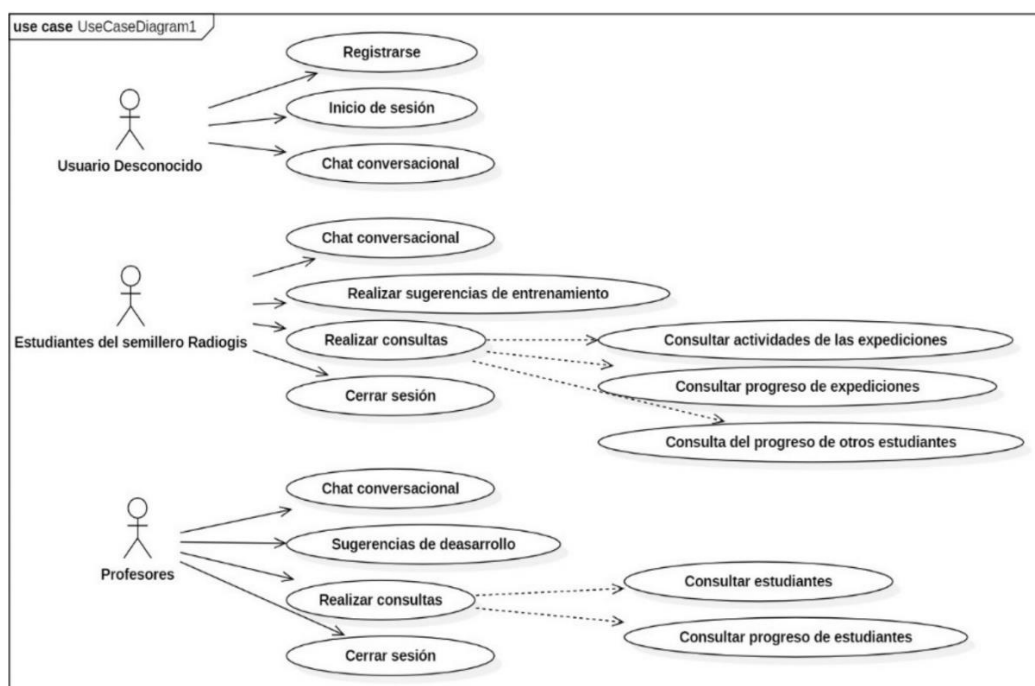
sobre las actividades y el progreso en las expediciones realizadas por el propio estudiante o de sus compañeros y finalmente los estudiantes tendrán también la posibilidad de interactuar con el chat conversacional o finalizar la sesión.

### 7.3. Rol de profesores

En el caso del rol de profesor, una vez que el usuario inicie sesión, este podrá acceder al chat conversacional, podrá realizar sugerencias respecto al Chatbot o al sitio web transmedia para que sean tomadas en cuenta por los desarrolladores, realizar consultas respecto a sus estudiantes y sus progresos o finalmente cerrar sesión.

#### Figura 6.

*Grafica de los casos de uso identificados*



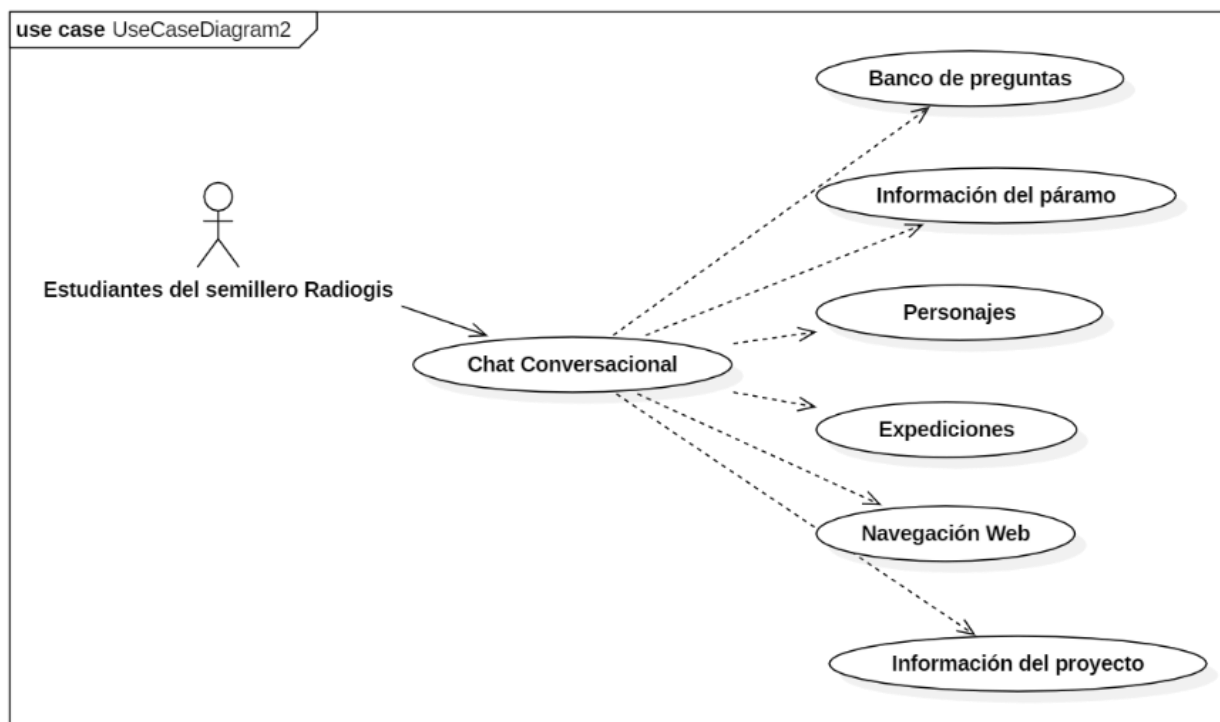
Como se pudo observar en la figura anterior, se puede observar que el caso de uso en común para todos es el chat conversacional, razón por la cual, el presente proyecto únicamente abarcará el diseño e implementación de este caso de uso.

Los demás casos de uso podrían ser añadidos al desarrollo en próximos aportes que den continuidad al actual proyecto, una vez sea finalizado y entregado al grupo de investigación.

Para poder desarrollar el caso de uso del chat conversacional, se toma como referencia toda la información recopilada de las frases de entrenamiento de la versión anterior, tal como se mencionó en el capítulo 6, la cual sirve como base para construir 6 categorías, en las cuales está el banco de preguntas con única respuesta, información referente al paramo, personajes del proyecto, sobre las expediciones, información y ayuda sobre la navegación en la página web transmedia e información general de todo el proyecto, tal como se presenta en la figura 7, que en el próximo capítulo se abarcará a más detalle.

**Figura 7.**

*Caso de uso - Chat conversacional*



## 8. Diseño de diagramas conversacionales

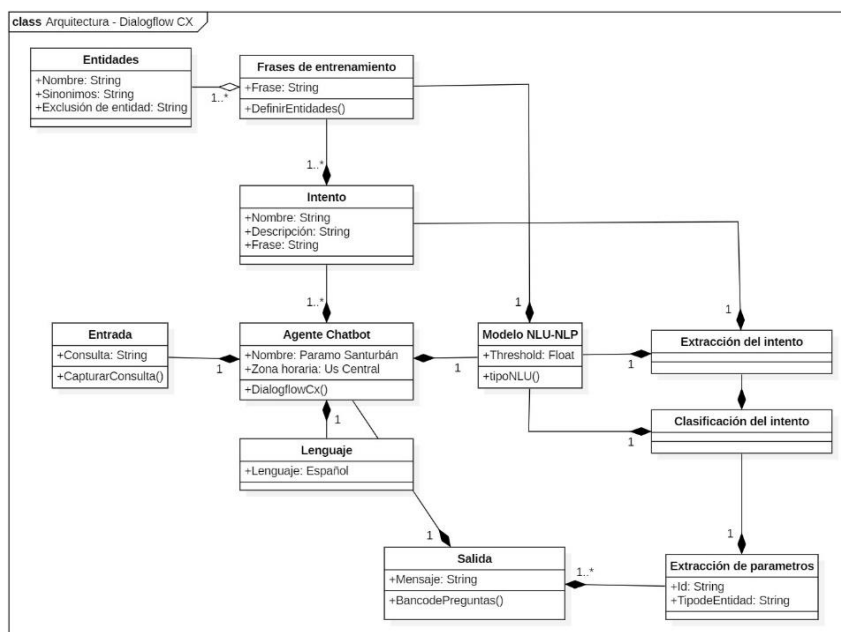
El presente capítulo tiene como objetivo presentar la implementación del caso de uso del chat conversacional presentado en diagramas conversacionales, los cuales estarán basados en los criterios mínimos en la representación de diagramas bajo el estándar UML.

### 8.1. Diagrama de la arquitectura del Chatbot

En la figura 8 se representa la arquitectura del Chatbot a implementar en Dialogflow CX, en el cual gracias a que se usa la norma UML, se pueden observar la dependencia que tienen los elementos respecto al Bot. En la arquitectura presentada, se observa como el agente está conformado por los intentos, las frases de entrenamiento que a su vez podrían conformarse por las entidades, por el lenguaje con el que se va a implementar el agente, el modelo NLU/NLP el cual es utilizado por la extracción y clasificación de intentos para extraer parámetros que generan los mensajes de salida.

**Figura 8.**

*Arquitectura del Chatbot basado en diagramas UML*









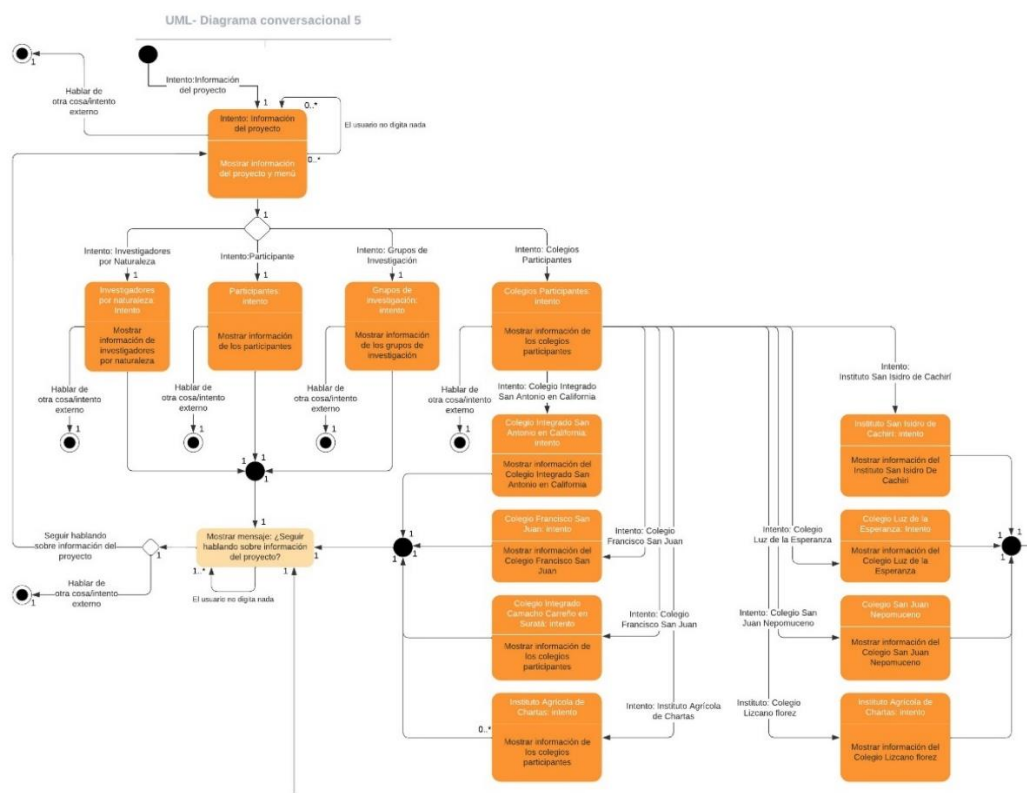


## 8.6. Diagrama conversacional 5

En la figura 13 se presenta el diagrama conversacional 5, en donde el caso de uso está relacionado a la información que el usuario desee saber sobre el proyecto, que a su vez se le presentarían una serie de sugerencias sobre Investigadores por naturaleza, Colegios participantes, Grupos de investigación e información de quienes pueden participar en el proyecto, en el cual el único que no tiene única respuesta es el relacionado a los colegios, ya que ahí se podrá pedir información a detalle de los colegios presentados en la figura 11 y al finalizar cada pregunta del diagrama conversacional, se le preguntará si desea seguir hablando sobre el proyecto o por si el contrario realiza otra pregunta no relacionada al diagrama podría irse a otro diagrama conversacional, banco de preguntas o finalizar la sesión.

**Figura 13.**

*Diagrama conversacional 5*



## 9. Implementación del Chatbot en Dialogflow CX

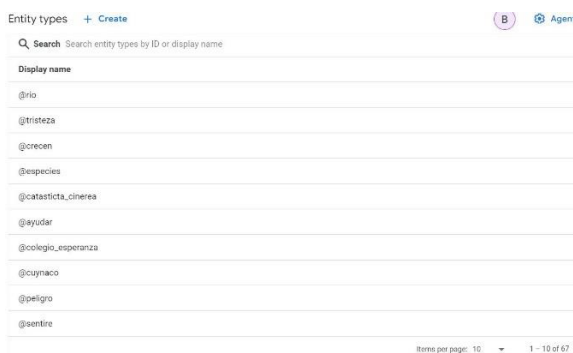
Este capítulo tiene como objetivo presentarle al lector, el proceso de implementación que se llevó a cabo al momento de la construcción del agente inteligente en la plataforma de Dialogflow CX, el cual fue guiado utilizando los diagramas conversacionales basados en UML, del capítulo anterior.

### 9.1. Creación de las entidades

Como se mencionó anteriormente, las entidades permiten extraer características específicas de la solicitud realizada por el usuario, con el fin de mejorar la detección. Para Chatbot desarrollado se creó una entidad para las palabras clave a identificar en un intento y así poder mejorar el entendimiento de una palabra cuando esta puede estar en otras preguntas similares pero en un contexto diferente, lo que puede ocasionar que el modelo se tienda a confundir, un ejemplo de esto pueden ser las preguntas que puede hacer el usuario sobre los personaje, las cuales son las mismas para cada uno, con la variación del nombre del personaje, por lo cual el agente puede tender a confundir las respuestas a entregar, en donde la solución a esto es implementar identidades como se observa en la figura 14.

#### Figura 14.

*Grafica de identidades integradas en Dialogflow CX*



Nota: Grafica obtenida de la plataforma de desarrollo Dialogflow CX

### 9.1.1. Creación y configuración de los intentos

Como primer paso para la creación del agente, se tendrán que crear todos los intentos, los cuales son los que llevarán a las preguntas con únicas respuestas o a los diagramas conversacionales, que después podrán asignarse de acuerdo a los diagramas diseñados en capítulos previos.

Como se presenta en la figura 15, al ingresar a la opción de intentos en la plataforma de Dialogflow, aparecerá en el interfaz, la posibilidad de agregar los intentos deseados. Para el caso del Chatbot a implementar, se crearon 179 intentos que representan las frases recopiladas en la versión preexistente y una serie de intentos que se adicionaron de acuerdo con los diagramas conversacionales diseñados.

**Figura 15. Grafica de los intentos integrados en Dialogflow CX**

The screenshot shows the 'Intents' page in Dialogflow CX. At the top, there are navigation links for 'Intents', '+ Create', 'Agent settings', and 'Test Agent'. Below the navigation is a search bar with the text 'Search Search intents by ID or display name'. The main content is a table listing various intents. The table has a 'Display name' column and a column with yellow warning triangles. The intents listed are: Default Welcome Intent, Default Negative Intent, el\_agua\_hierve\_mas\_rápido, novios, 5.2.8.Colegio\_lizcano\_flores, 1.2.animales, que\_vamos\_hacer, Estas\_ocupado, que\_te\_parezco, and 2.1.3.caracteristica\_canis. At the bottom right, there is a pagination control showing 'Items per page: 10' and '1 - 10 of 179'.

| Display name                 |   |
|------------------------------|---|
| Default Welcome Intent       | ▲ |
| Default Negative Intent      | ▲ |
| el_agua_hierve_mas_rápido    | ▲ |
| novios                       |   |
| 5.2.8.Colegio_lizcano_flores |   |
| 1.2.animales                 |   |
| que_vamos_hacer              | ▲ |
| Estas_ocupado                | ▲ |
| que_te_parezco               |   |
| 2.1.3.caracteristica_canis   |   |

Nota: Grafica obtenida de la plataforma de desarrollo Dialogflow CX

Como se presenta en la figura 16, para la configuración de cada intento se requiere principalmente del nombre y las frases de entrenamiento que ayudarán a la identificación de este, y adicional se podría agregar una leve descripción y una etiqueta.

Hay que tener en cuenta que estas frases de entrenamiento pueden agregarse con archivos externos, sin embargo, para el desarrollo todo fue incluido de forma manual. En este procedimiento de agregar las frases, en el caso de ser necesario se podrán ir vinculando los intentos ya creados anteriormente.

### Figura 16.

#### *Grafica de la configuración de un intento en Dialogflow Cx*

Intents represent something your users want to do during a conversation with your agent (for example, schedule an appointment). [Learn more](#)

Display name \*

1.Hablame\_sobre\_ti

Labels ?

Description

Training phrases

When a user says something similar to a training phrase, Dialogflow matches it to the intent. You don't have to create an exhaustive list. Dialogflow will fill out the list with similar expressions. To extract parameter values, use [annotations](#) with available [system](#) or [custom](#) entity types.

Type a training phrase and press 'Enter' Add

↑ 🗑

**Q Search** Search training phrases

| Training phrases           |
|----------------------------|
| Hola quien eres            |
| Hola, quien eres           |
| Hola quien eres?           |
| Hola, quien eres?          |
| <b>Parámo</b> de santurban |
| <b>Paramo</b> santurban    |
| <b>Paramo</b> de santurban |
| como te llamas             |
| cual <b>es</b> tu nombre   |
| santurban                  |

Items per page: 10 1 - 10 of 28 << < > >>

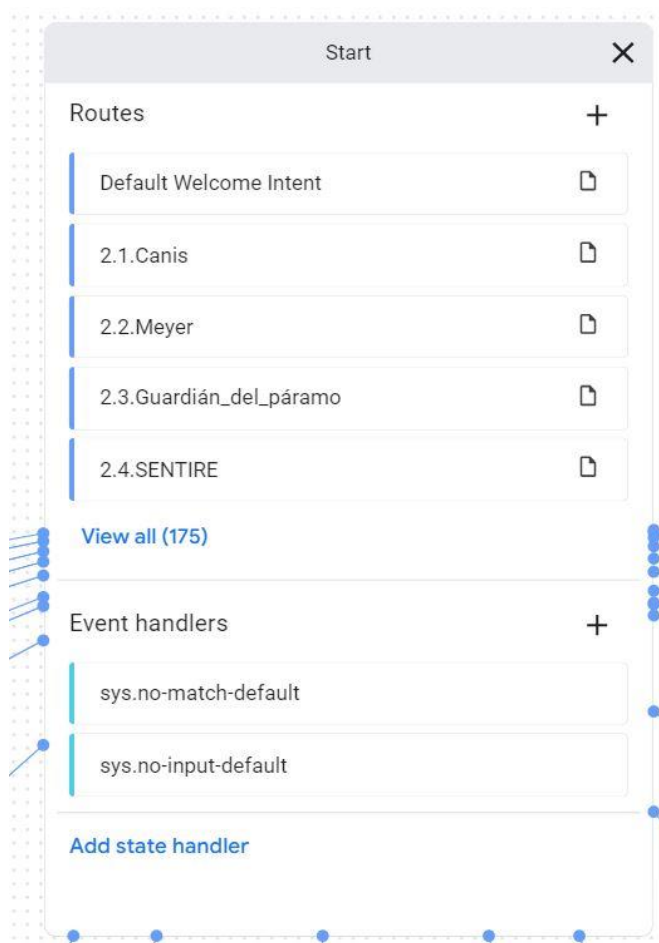
Nota: Grafica obtenida de la plataforma de desarrollo Dialogflow CX

Una vez se han agregado todos los intentos a utilizar en Dialogflow, se debe ir a la pestaña de Build, en donde se observarán todos los diagramas construidos de forma gráfica, que al entrar por primera vez solo se encontrará la página de Start, la cual dará paso a todas las páginas creadas, que en conjunto al configurarse darán lugar a los diagramas conversacionales deseados.

Como se observa en la figura 17, en la página de Start, se configuraron los 175 intentos, los cuales pertenecen al banco de preguntas con única respuesta y adicional, se configuraron las preguntas que llevarían a los diagramas conversacionales como se presentó en el capítulo anterior.

### Figura 17.

*Grafica de la página Start en la plataforma de Dialogflow CX*

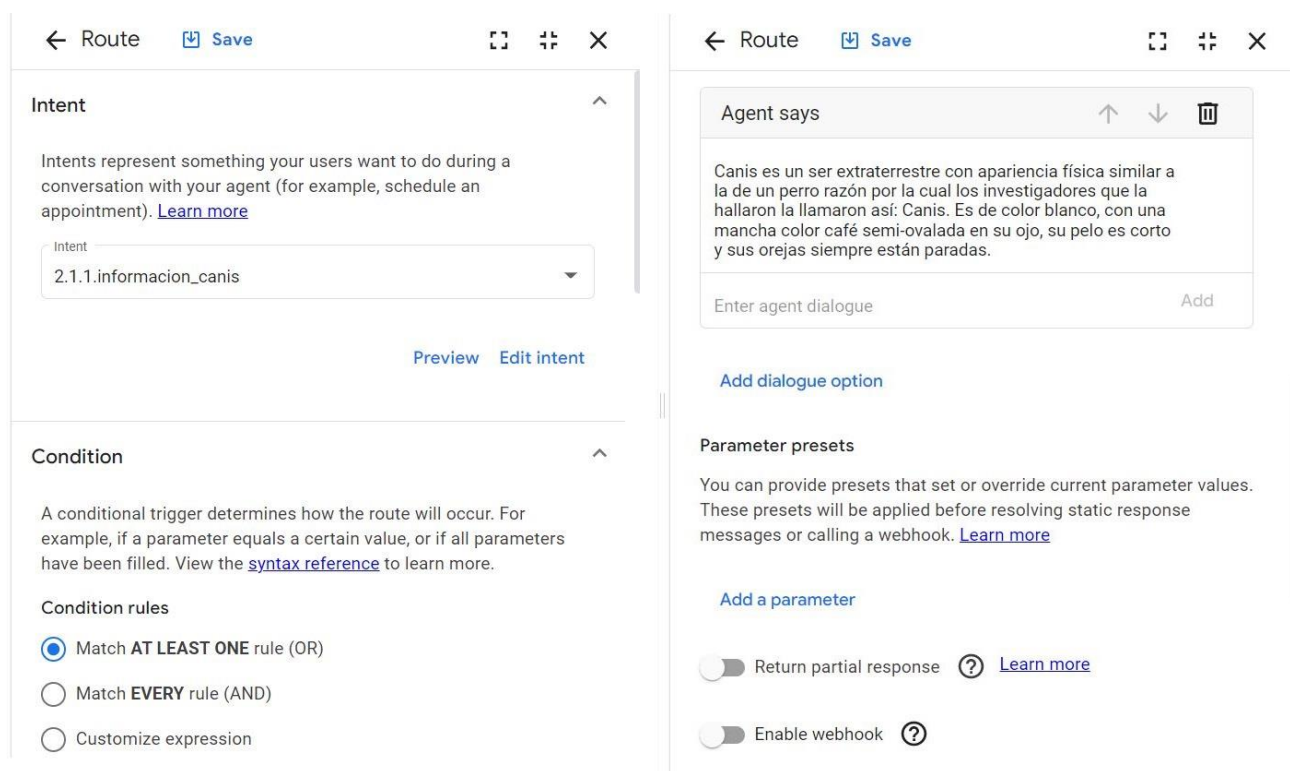


Nota: Grafica obtenida de la plataforma de desarrollo Dialogflow CX

En la figura 18, se puede observar el intento asociado a la información de uno de los personajes, Canis, en donde al asociar dicho intento a una página determinada, se le puede asociar la frase a mostrar, la extracción de algún parámetro, algún condicional basado en la extracción de dicho parámetro o redireccionar a otra página, ya que como se mencionó en capítulos anteriores su funcionamiento está basado en las máquinas de estado.

### Figura 18.

*Grafica de la configuración de un intento en Dialogflow Cx*



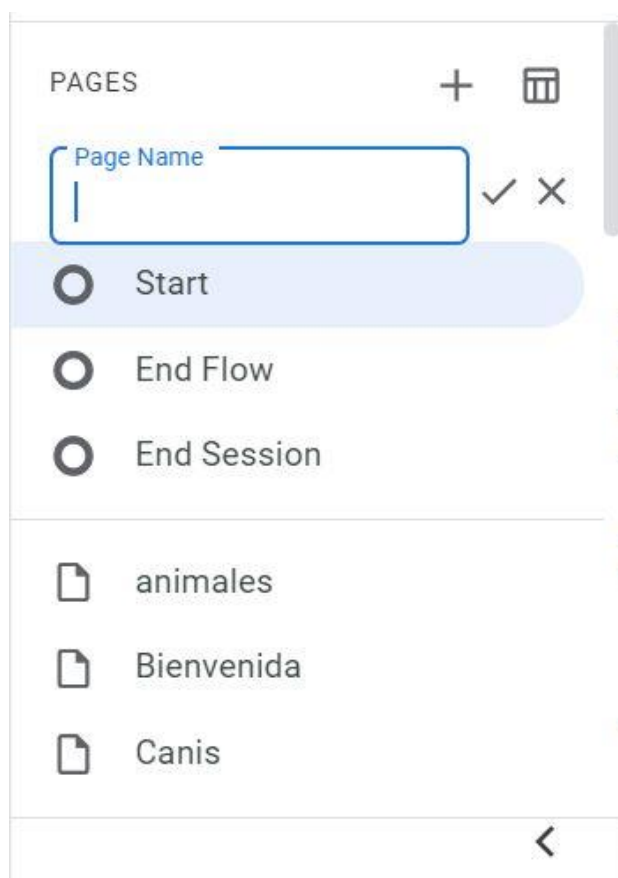
Nota: Grafica obtenida de la plataforma de desarrollo Dialogflow CX

Para poder asociar las diferentes paginas a los diferentes intentos configurados, se requiere crear para cada una de estas, ya sea para reunir múltiples intentos o exclusivamente uno, una página. Para esto se debe dirigir a la parte izquierda de la plataforma y dar en *add page*, como se observa en la figura 19, donde una vez creado se puede dirigir nuevamente a la página donde se

estaban configurando los intentos, y agregar la nueva página en la configuración de transiciones, al realizar esto, gráficamente se empieza a observar un diagrama con las respectivas transiciones.

### Figura 19.

*Grafica del interfaz de una página en Dialogflow CX*

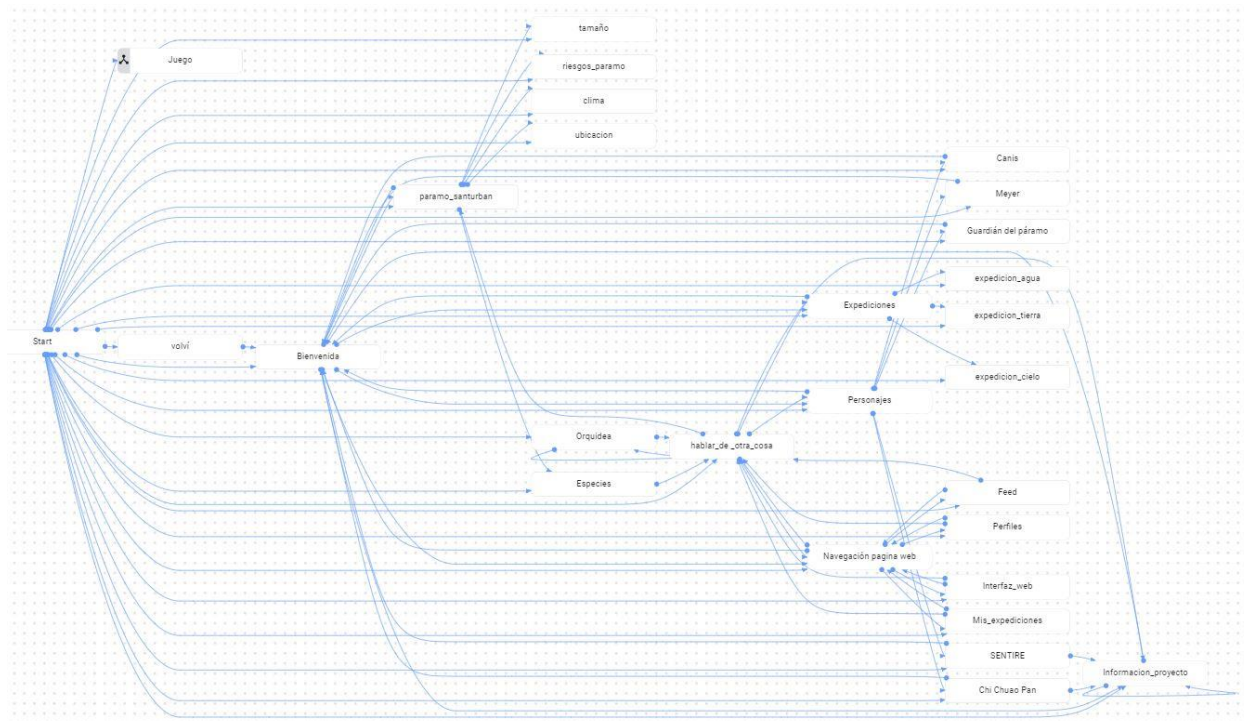


Nota: Grafica obtenida de la plataforma de desarrollo Dialogflow CX

En la figura 20, se puede observar el diagrama construido al finalizar toda la construcción del agente conversacional, en donde se realizaron todas las configuraciones de todos los intentos, las transiciones entre páginas, configuración de entidades y de los demás parámetros que lo conforman respecto con los diagramas conversacionales 1,2,3,4, y 5, mencionados en capítulos previos.

**Figura 20.**

*Grafica del interfaz grafico del Chatbot diseñado e implementado en Dialogflow Cx.*



## 10. Análisis de resultados

El presente capítulo tiene como objetivo presentar los resultados obtenidos en los diferentes tests llevados a cabo tanto como con la comunidad universitaria, como con los niños que participaron en las diferentes actividades para poder determinar las respectivas métricas de rendimiento del Chatbot.

### 10.1. Sondeo previo

Para poder determinar el rendimiento y recoger comentarios del Chatbot antes de testarse con los niños, se seleccionaron 10 personas pertenecientes a la comunidad universitaria, las cuales realizaron una serie de 166 preguntas cada uno, de las cuales 106 preguntas pertenecen al banco de preguntas y las 60 preguntas restantes, fueron repartidas para probarse dentro de los diferentes diagramas conversacionales presentados en capítulos anteriores, dicha recolección de información puede ser consultada en el Apéndice B.

Como se observa en la tabla 5, se utilizó una métrica de rendimiento de tasa de aciertos, la cual utiliza el número de respuestas correctas en la totalidad de usuarios, respecto a la cantidad total de preguntas realizadas por todos, (Carlos F. Gonzales, 2020), que aunque es una métrica de rendimiento básica, es suficiente para realizar el sondeo previo del rendimiento del Chatbot, en el cual se logró obtener una exactitud del 96.87% y una satisfacción por parte de los usuarios del 92,53%, adicionalmente se recogieron todos los comentarios del usuario para re entrenar el modelo y mejorar diferentes aspectos del Chatbot y prepararlo mejor para las próximas actividades.

**Tabla 5. Tabla de la métrica de rendimiento de la tasa de aciertos**

|           | Número de preguntas respondidas de forma correcta | Número de preguntas con la que está satisfecho el usuario |
|-----------|---|---|
| Usuario 1 | 160   | 122   |

|              |        |     |
|--------------|--------|-----|
| Usuario 2    | 162    | 162 |
| Usuario 3    | 161    | 161 |
| Usuario 4    | 159    | 159 |
| Usuario 5    | 160    | 133 |
| Usuario 6    | 159    | 159 |
| Usuario 7    | 163    | 160 |
| Usuario 8    | 162    | 160 |
| Usuario 9    | 159    | 158 |
| Usuario 10   | 163    | 162 |
| Exactitud    | 96,87% |     |
| Satisfacción | 92,53% |     |

## 10.2. Actividad para la identificación de necesidades de los niños

Una vez realizado el sondeo de rendimiento del Chatbot y realizadas las respectivas correcciones con base a los comentarios realizados por los que llevaron a cabo la actividad, se hizo un primer acercamiento con los niños del Colegio Lizcano Flores de Matanza, para poder identificar las necesidades de los niños referente a la temática realizada en el páramo de Santurbán, en el cual participaron 11 niños, los cuales fueron escogidos por el profesor Humberto Rodríguez – Ingeniero de Telecomunicaciones, el cual imparte clases de Tecnología e Informática en el

colegio Lizcano Flores, quien abrió pequeños espacios con tiempo limitado para llevar a cabo la charla introductoria a las actividades y a los tests a realizar por parte de los estudiantes.

Este primer acercamiento, se realizó de forma remota para mayor facilidad de los estudiantes y del docente, se les explicó la importancia que tienen estos Chatbot en la actualidad junto a los últimos desarrollos realizados en los diferentes campos del conocimiento y se finalizó presentando el trabajo realizado con la narrativa del páramo de Santurbán por parte de RadioGis, grupo de investigación perteneciente a la Universidad Industrial de Santander. Después de dicha presentación se llevó a cabo una actividad con el fin de identificar, que se podría agregar o corregir en el Chatbot referente al páramo, en la que los niños podían interactuar con él, de forma libre, preguntando lo que quisiesen referente a la temática del páramo de Santurbán y a su vez se iba tomando nota respecto a comentarios e interacciones.

De la actividad se pudo extraer que los niños querían saber sobre los frailejones, municipios cercanos, entre otros y que el Chatbot aún no incorporaba la información sobre las temáticas preguntadas, por lo cual dicha información se utilizó para entrenar nuevamente el Chatbot y así poder llevar a cabo el primer test de rendimiento de este, en los que participarían los mismos 11 niños.

También la actividad sirvió para darse cuenta de que ellos no podrían realizar un test tan extenso, como el realizado en el sondeo previo, donde cada persona respondió alrededor de 166 preguntas, ya que pasado cierto tiempo tendían a aburrirse o perder el interés, por lo cual el test a realizar sería diferente respecto al sondeo previo.

Para los dos Test de rendimiento realizados, se dividieron los 11 niños en dos grupos, 6 para cada test, asignando un estudiante en los dos grupos con el fin de que no volvieran responder las mismas preguntas de preferencia.

### **10.3. Test de rendimiento 1**

Para el primer test de rendimiento, los primeros 6 niños respondieron un total de 171 preguntas, las cuales se programaron en 6 sesiones, en donde por sesión, se repartieron todas las preguntas entre todos ellos, esto con el fin de que cada niño no respondiera la totalidad de las preguntas en una sola sesión, esto ya que como se pudo observar en el sondeo previo, que fue una actividad muy larga e iba a generar que se cansaran o aburrieran y generara un probable error en los resultados recogidos en los testeos a realizar. También es de resaltar que el testeo aplicado a los niños tiene como principal diferencia respecto al sondeo previo realizado, en que no se les presentó ningún tipo de sugerencia para realizar la pregunta, sino por el contrario, ellos tuvieron que formular las preguntas de acuerdo a la situación que se le iba presentado.

Para este primer test aplicado, se recogieron un total de 1026 preguntas y para poder determinar el rendimiento de Chatbot, esta información recogida se clasificó para armar dos matrices de confusión con las cuales se pudo determinar la exactitud y otras métricas de interés.

#### ***10.3.1. Matriz de confusión binaria***

Para determinar la exactitud del modelo NLU/NLP incorporado en el agente implementado, se utilizó de una matriz de confusión binaria conformada por dos clases, positiva y negativa, donde el modelo podrá predecir dentro de estas dos clases.

En el apéndice C, se podrán encontrar las 6 sesiones realizadas con los niños, en las cuales se recogió información sobre la pregunta realizada, la respuesta del Chatbot y luego de forma manual se completó con la información que debió responder el Chatbot, el intento de lo que debiera responder y del que respondió. De tal forma, se clasificaron todos los datos dentro de la matriz de confusión de forma manual. La primera matriz de confusión generada con la información recogida se podrá observar en la tabla 6.

**Tabla 6. Matriz de confusión binaria en el Test de rendimiento 1**

|           | Positivos | Negativos |
|-----------|-----------|-----------|
| Positivos | 881       | 133       |
| Negativos | 6         | 6         |

De la matriz de confusión presentada en la tabla 6, se puede extraer la información necesaria para determinar que el Chatbot tiene una exactitud del 86.45% que indica el porcentaje de predicciones correctas, el cual es mucho menor respecto al realizado en el sondeo previo, debido a que, para este Test, los niños debían formular por sí mismos la pregunta, por lo que se intuye que eso pudo haber influido en la exactitud determinada, sin embargo, la exactitud determinada cumple con una de las especificaciones de diseño a cumplir.

$$Exactitud = \frac{881 + 6}{881 + 6 + 133 + 6} * 100 = 86.45\%$$

$$Precisión = \frac{881}{881 + 6} * 100 = 99.32\%$$

$$Recall = \frac{881}{881 + 133} * 100 = 86.88\%$$

$$F1 - Score = \frac{881}{881 + \left(\frac{133 + 6}{2}\right)} * 100 = 92.68\%$$

Es importante mencionar, que, con la matriz de confusión presentada, también se pudo determinar una precisión del 99.32% que permite determinar el porcentaje de predicción de la totalidad de los datos que fueron indicados como correctos, un Recall del 86.88% que permite determinar el porcentaje de datos que realmente eran positivos y F1-score del 92.68% el cual es

un balance entre la precisión y el Recall en donde la métrica incrementa a medida que los falsos negativos y positivos disminuyen.

### ***10.3.2. Matriz de confusión multiclase***

A diferencia de la matriz de confusión binaria anteriormente utilizada para determinar las diferentes métricas de rendimiento del modelo, que solo tomaba a consideración que tanto se confundía en general, la segunda matriz de confusión multiclase determina las mismas métricas de rendimiento, con la diferencia que revisa que tanto se equivoca cada intento respecto a los demás, obteniéndose una matriz de confusión de 170x170 en la cual se clasificó toda la información recogida y se pudo determinar la precisión, Recall y F1-score de cada intento.

En el Apéndice C, en la página de la matriz de confusión multiclase, se puede observar que, en la tabla de métricas de rendimiento, hay múltiples intentos con métricas de precisión, Recall y F1-Score por debajo del 80%, sin embargo, sus porcentajes generales del modelo están por encima del 80%.

En la tabla 7, se puede observar que se obtuvo una exactitud del 86,45% el cual indica el porcentaje de respuestas correctas, una precisión del 89.99% el cual indica las respuestas correctas que indicó el modelo, un Recall del 86.67% el que indica las respuestas correctas reales del modelo y un F1-Score del 86.34% el cual indica un equilibrio entre Recall y precisión en donde toma a consideración los falsos positivos y los falsos negativos.

**Tabla 7. Resultados de la matriz de confusión multiclase en el Test de rendimiento 1**

| Métricas de rendimiento |        |
|-------------------------|--------|
| Exactitud               | 86.45% |
| Precisión               | 89.99% |

---

|          |        |
|----------|--------|
| Recall   | 86.67% |
| F1-Score | 86.34% |

---

#### **10.4. Test de rendimiento 2**

Para la segunda prueba de rendimiento, se contó con la ayuda de los 6 niños restantes, incluyendo adicional otro estudiante del test anterior, que respondieron las mismas 171 preguntas, las cuales se repartieron en 6 sesiones al igual que la actividad realizada en el primer test de rendimiento.

Para esta segunda prueba, se tomaron a consideración los errores obtenidos en cada intento y se aplicaron las respectivas correcciones para reentrenar el modelo de NLU que incorpora el Chatbot y al igual que el Test previo, se recogió un total de 1026 preguntas y para poder determinar su rendimiento y se utilizaron las dos matrices de confusión utilizadas en el Test de rendimiento anterior.

##### ***10.4.1. Matriz de confusión binaria***

En el Apéndice D, se podrán encontrar las 6 sesiones realizadas con los niños, en las cuales se recogió información sobre la pregunta realizada, la respuesta del Chatbot y luego de forma manual se completó al igual que los Test de rendimiento previos. La primera matriz de confusión binaria generada con la información recogida se observa en la tabla 8.

**Tabla 8. Matriz de confusión binaria en el Test de rendimiento 2**

|           | Positivos | Negativos |
|-----------|-----------|-----------|
| Positivos | 997       | 17        |
| Negativos | 8         | 4         |

De la matriz de confusión presentada en la tabla 8, se puede determinar que el Chatbot tiene una exactitud del 97.56% que indica el porcentaje de predicciones correctas, el cual es mucho mayor, respecto al obtenido en el Test de rendimiento 1, la cual cumple con una de las especificaciones de diseño a cumplir.

$$Exactitud = \frac{997 + 4}{997 + 4 + 17 + 8} * 100 = 97.56\%$$

$$Precisión = \frac{997}{997 + 8} * 100 = 99.2\%$$

$$Recall = \frac{997}{997 + 17} * 100 = 98.32\%$$

$$F1 - Score = \frac{997}{997 + \left(\frac{8 + 17}{2}\right)} * 100 = 98.76\%$$

Es importante mencionar, que, con la matriz de confusión presentada, también se pudo determinar una precisión del 99.2% que permite determinar el porcentaje de predicción de la totalidad de los datos que fueron indicados como correctos, un Recall del 98.32% que permite determinar el porcentaje de datos que realmente eran positivos y F1-score del 98.76% el cual es un balance entre la precisión y el Recall en donde la métrica incrementa a medida que los falsos negativos y positivos disminuyen.

### 10.4.2. Matriz de confusión multiclase

En el Apéndice D, en la página de la matriz de confusión multiclase se podrá encontrar a detalle la información clasificada y se podrán encontrar las métricas que se resumen en la tabla 9.

En la tabla 9, se puede observar que se obtuvo una exactitud del 97,66% el cual indica el porcentaje de respuestas correctas, una precisión del 98.38% el cual indica las respuestas correctas que indicó el modelo, un Recall del 97.99% el que indica las respuestas correctas reales del modelo y un F1-Score del 98.00% el cual indica un equilibrio entre Recall y precisión en donde toma a consideración los falsos positivos y los falsos negativos.

**Tabla 9. Resultados de la matriz de confusión multiclase en el Test de rendimiento 2**

| Métricas de rendimiento |        |
|-------------------------|--------|
| Exactitud               | 97.56% |
| Precisión               | 98.38% |
| Recall                  | 97.99% |
| F1-Score                | 98.00% |

## **11. Trabajos futuros**

Gracias a el Chatbot desarrollado en este proyecto de investigación, se deja una sólida base para que próximos investigadores puedan integrar el uso de bases de datos para incorporar los demás casos de uso expuestos en el capítulo 7 que podrían mejorar la experiencia de los estudiantes y docentes.

Adicionalmente los próximos investigadores que deseen dar continuidad al proyecto podría crear una estación meteorológica en el páramo de Santurbán que incorpore sensores de temperatura, humedad, Nivel PH, captura de imágenes entre otros, para el cual tendrán que diseñar un sistema embebido que puedan enlazarse con una base de datos que pueda ser consultada las 24 horas por los estudiantes y profesores a través del Chatbot.

## **12. Conclusiones**

El Chatbot implementado en Dialogflow CX, logra eliminar las limitaciones que tenía la versión preexistente, adicionalmente es una plataforma de bajo costo, que cuenta con modelos de inteligencia artificial como lo es el NLU/NLP, permite incorporar complejos flujos conversacionales y además cuenta con la posibilidad de conectar el agente con las demás herramientas con las que cuenta Google Cloud, las cuales podrían ser aprovechadas por los investigadores que den continuad al actual proyecto.

Se observa en la implementación, la importancia de contar con un diseño previo de los diagramas conversacionales, los cuales estan basados en la norma UML, ya que funcionan como un tipo de plano a seguir al momento de implementar.

En la actividad realizada con los 11 niños del colegio Lizcano Flores con el fin de identificar que necesidades adicionales habían respecto a la versión previa, la cual fue desplegada en la página de web del proyecto transmedia, se pudo identificar que al tener múltiples diagramas

conversacionales, permite que ellos pueden explorar el Chatbot aunque no sepan que preguntar, debido a que él va ofreciendo sugerencias para afianzar la interacción con ellos, además la actividad permitió identificar que no podían responder un test con demasiadas preguntas, lo que eventualmente iba a generar que aburrieran y cansaran, lo cual se utilizó para modificar los Test que realizaron posteriormente. También se pudo evidenciar un alto interés por la inteligencia artificial, lo cual fue lo que los motivó a participar en los posteriores testeos del Chatbot.

Se observó en el capítulo 10, en donde se determinó que la exactitud del modelo utilizado por el Chatbot fue del 96.87% en el sondeo previo, y del 86.45% en el primer Test de rendimiento realizado con los niños, el cual se concluye que se debe a la diferencia de edades.

Se pudo evidenciar que la exactitud obtenida al finalizar el Test de rendimiento 2, fue del 97.66% la cual presenta una mejora del 11.21% respecto al 86.45% de exactitud obtenida en el Test de rendimiento 1, la cual da cumplimiento a una de las especificaciones de diseño propuestas.

Se resalta la utilidad de utilizar una matriz de confusión multiclase para este tipo de Chatbots, respecto a la matriz de confusión binaria, esto tomando en cuenta que la matriz multiclase permite obtener a detalle, métricas como lo es la precisión, Recall y F1-Score respecto a todos los intentos incorporados, lo que permitió identificar rápidamente en donde centrar la atención para reentrenar el modelo y obtener mejores resultados, algo que no ocurre con la matriz binaria, ya que aunque permite calcular las mismas métricas, indica que tanto se está equivocando el modelo en general.

Finalmente se puede concluir que el Chatbot desarrollado logra apoyar las interacciones de los niños respecto a la página web transmedia, en donde se mantiene la narrativa creada en el páramo de Santurbán, que a su vez apoya la apropiación del conocimiento y generar interés respecto a uno de los ecosistemas más importantes como lo es el páramo de Santurbán y al tiempo

deja un precedente documentado en el diseño, implementación, entrenamiento y determinación de las métricas de rendimiento en lo que se puede considerar como el primer Chatbot con la narrativa de un páramo.

### **13. Recomendaciones**

Se le recomienda al grupo de investigación RadioGis, quién recibe el Chatbot desarrollado, realizar alianzas con la escuela de Biología, con el fin de poder recolectar una mayor cantidad de especies para mostrarle a los niños y que se mantenga actualizada la información.

## Referencias Bibliográficas

- Alejandro Caloguera. (2012). *Guía para la producción y distribución de contenidos TRANSMEDIA para múltiples plataformas* (Noviembre 2012, Santiago de Chile ed.). Cerebro Digital / Clickprint.
- Alexander von Humboldt. (2011). *El gran libro de los páramos*. Brigitte LG Baptiste.
- Amazon Web Service. (2022). *Amazon Lex V2 Developer Guide*.  
<https://aws.amazon.com/es/lex/?nc=sn&loc=0>.  
<https://docs.aws.amazon.com/lexv2/latest/dg/what-is.html>
- Arista-Jalife, A., Calderón-Auza, G., Fierro-Radilla, A., & Nakano, M. (2017). Clasificación de Imágenes Urbanas Aéreas: Comparación entre Descriptores de Bajo Nivel y Aprendizaje Profundo. *Información Tecnológica*, 28(3), 209-224. 10.4067/S0718-07642017000300021
- Carlos F. Gonzales. (2020). *10 métricas clave para evaluar el rendimiento de tu Chatbot*. Inbenta. <https://www.inbenta.com/es/blog/10-metricas-clave-para-evaluar-el-rendimiento-de-tu-chatbot/>
- Comunicaciones, V. (2022, Abril1,). Investigadores por Naturaleza continúa su travesía. Niños, niñas y jóvenes del Páramo de Santurbán son los protagonistas ? Universidad Industrial de

- Santander. <https://uis.edu.co/investigadores-por-naturaleza-continua-su-travesia-ninos-ninas-y-jovenes-del-paramo-de-santurban-son-los-protagonistas/>
- Das, S., & Kumar, E. (Dec 2018). Determining Accuracy of Chatbot by applying Algorithm Design and Defined process. Paper presented at the 1-6. 10.1109/CCAA.2018.8777715  
<https://ieeexplore.ieee.org/document/8777715>
- Garcia Brustenga, G., Fuertes Alpiste, M., & Molas Castells, N. (2018). *Briefing Paper: Chatbots in Education*. (). Barcelona: 10.7238/elc.chatbots.2018  
<http://hdl.handle.net/10609/80185>
- Géron, A. (2019). In Tache N. (Ed.), *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd Editio ed.). O'Reilly Media, Inc.
- Google Cloud. (2022a). *Dialogflow CX documentation* .  
<https://cloud.google.com/dialogflow/cx/docs?hl=es-419>
- Google Cloud. (2022b). *Dialogflow ES documentation*. Google Cloud.  
[https://cloud.google.com/dialogflow/es/docs?hl=en#documentation\\_contents](https://cloud.google.com/dialogflow/es/docs?hl=en#documentation_contents)
- Gupta, J., Singh, V., & Kumar, I. (March 2021). Florence- A Health Care Chatbot. Paper presented at the 504-508. 10.1109/ICACCS51430.2021.9442006  
<https://ieeexplore.ieee.org/document/9442006/>
- IBM. (2022). *Classic Watson Assistant*. <https://www.ibm.com/co-es>.  
<https://cloud.ibm.com/docs/assistant?topic=assistant-index>

Kasthuri, E., & Balaji, S. (February 2021). A Chatbot for Changing Lifestyle in Education. Paper presented at the 1317-1322. 10.1109/ICICV50876.2021.9388633

<https://ieeexplore.ieee.org/document/9388633/>

M. Heydarian, T. E. Doyle, & R. Samavi. (2022). *MLCM: Multi-Label Confusion Matrix* 10.1109/ACCESS.2022.3151048

Mabunda, K., & Ade-Ibijola, A. (November 2019). PathBot: An Intelligent Chatbot for Guiding Visitors and Locating Venues. Paper presented at the 160-168.

10.1109/ISCM147871.2019.9004411 <https://ieeexplore.ieee.org/document/9004411/>

Microsoft. (2022). *Azure Bot Service documentation*. MICROSOFT.

<https://docs.microsoft.com/en-us/documentation/>

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill Science/Engineering/Math.

Muhammad, A. F., Susanto, D., Alimudin, A., Adila, F., Assidiqi, M. H., & Nabhan, S.

(September 2020). Developing English Conversation Chatbot Using Dialogflow. Paper presented at the 468-475. 10.1109/IES50839.2020.9231659

<https://ieeexplore.ieee.org/document/9231659/>

Murali, V., Sarma, R. J., Sukanya, P. A., & Athri, P. (December 2018). ChEMBL Bot - A Chat Bot for ChEMBL database. Paper presented at the 1-6.

10.1109/ICINPRO43533.2018.9096710 <https://ieeexplore.ieee.org/document/9096710/>

*NLP vs NLU: What's The Difference?* BMC Blogs. <https://www.bmc.com/blogs/nlu-vs-nlp-natural-language-understanding-processing/>

P, M. G. C., Srivastava, A., Chakraborty, S., Ghosh, A., & Raj, H. (April 2021). Development of Information Technology Telecom Chatbot: An Artificial Intelligence and Machine Learning Approach. Paper presented at the 216-221. 10.1109/ICIEM51511.2021.9445354  
<https://ieeexplore.ieee.org/document/9445354/>

*Precise Semantics of UML State Machines (PSSM)* (2019).

Ricardo Borja-Robalino, Antonio Monleón-Getino, & José Rodellar. (2020). *Estandarización de métricas de rendimiento para clasificadores Machine y Deep Learning*

Vanguardia. (2021). *Un santandereano es el creador de la primera humana virtual colombiana*. [www.vanguardia.com](https://www.vanguardia.com). <https://www.vanguardia.com/area-metropolitana/bucaramanga/un-santandereano-es-el-creador-de-la-primera-humana-virtual-colombiana-JL4580920>

Yoshua Bengio, Ian Goodfello, Aaron Courville. (2016). In 2016), The MIT Press; Illustrated edición (18 Noviembre (Ed.), *Deep Learning*

Zuleika, M., Díaz, J., Alexander, F., Sepúlveda, S., Camilo, C., Otálora, A., Doctorado En Ciencias Química, Bastidas, David, C., & Parra, C. (2022). *Herramienta de software para la extracción automática de parámetros de desempeño a partir de publicaciones científicas de celdas solares de perovskita*

## **Apéndices**

**Apéndice A.** Recopilación de frases de entrenamiento de la versión preexistente

**Apéndice B.** Sondeo previo

**Apéndice C.** Test de rendimiento 1

**Apéndice D.** Test de rendimiento 2

Los apéndices están adjuntos y puede visualizarlos en la base de datos de la biblioteca UIS