



## CARRERA DE ESPECIALIZACIÓN EN INTELIGENCIA ARTIFICIAL

MEMORIA DEL TRABAJO FINAL

### **Clasificación de reclamos de usuario**

**Autor:**

**Ing. Lucas Rivela**

Director:

Dr. Lic. Rodrigo Cárdenas (FIUBA)

Jurados:

Nombre del jurado 1 (pertenencia)

Nombre del jurado 2 (pertenencia)

Nombre del jurado 3 (pertenencia)

*Este trabajo fue realizado en la Ciudad Autónoma de Buenos Aires,  
entre mayo de 2023 y septiembre de 2023.*



## *Resumen*

La presente memoria describe el diseño e implementación de un sistema de clasificación de reclamos desarrollado para Ualá. La solución permite optimizar el tiempo que lleva el proceso de clasificación mediante la asignación automática de categorías de primer y segundo nivel; y por otro lado, reducir la cantidad de personas necesarias para esta tarea.

Para poder realizar este trabajo se aplicaron conceptos de bases de datos, procesamiento del lenguaje natural y aprendizaje profundo para realizar la extracción de los datos, el procesamiento del texto, el entrenamiento de los modelos de IA y el despliegue de los mismos en un ambiente de desarrollo.



## *Agradecimientos*

A mi familia y amigos por apoyarme durante la realización de esta carrera.

A mis compañeros y profesores por el acompañamiento.

A mi director, Dr. Lic. Rodrigo Cárdenas por orientarme y guiarme en la realización de este trabajo.



# Índice general

<b>Resumen</b>	<b>I</b>
<b>1. Introducción general</b>	<b>1</b>
1.1. Introducción	1
1.1.1. Qué es la atención al cliente	1
1.1.2. Organización de un equipo de atención al cliente	1
1.1.3. El proceso de atención al cliente	2
1.2. Motivación	3
1.3. Estado del arte	4
1.3.1. Word Embeddings	4
1.3.2. Transformers	4
1.4. Objetivos y alcance	5
1.4.1. Objetivos	5
1.4.2. Alcance	6
<b>2. Introducción específica</b>	<b>7</b>
2.1. Requerimientos	7
2.2. Preprocesamiento del texto	8
2.3. Modelos de inteligencia artificial utilizados	8
2.4. Herramientas de software utilizadas	8
<b>3. Diseño e implementación</b>	<b>9</b>
3.1. Análisis del software	9
<b>4. Ensayos y resultados</b>	<b>11</b>
4.1. Pruebas funcionales del hardware	11
<b>5. Conclusiones</b>	<b>13</b>
5.1. Conclusiones generales	13
5.2. Próximos pasos	13
<b>Bibliografía</b>	<b>15</b>





# Índice de figuras

1.1. Organigrama de un área de atención al cliente <sup>1</sup> . . . . .	2
1.2. Evolución de las redes en cantidad de parámetros. <sup>2</sup> . . . . .	5
1.3. Diagrama general de la solución. . . . .	6



# Índice de tablas



# Capítulo 1

## Introducción general

En este capítulo se realiza una introducción al funcionamiento de un área de atención al cliente. Además, se menciona el estado del arte de los sistemas de procesamiento del lenguaje natural, y por último se explican los objetivos y alcances del presente trabajo.

### 1.1. Introducción

En esta sección se introduce un área típica de atención al cliente en una empresa.

#### 1.1.1. Qué es la atención al cliente

El área de atención al cliente se encarga de dar soporte al consumidor y tiene como objetivo resolver sus problemas [1].

A menudo se confunde el área de atención con el área de servicio al cliente. La principal diferencia radica en que el servicio es una función proactiva con la intención de anticiparse a las necesidades inmediatas y a largo plazo del cliente. La atención al cliente es una función reactiva que busca resolver los problemas que el cliente manifestó [2].

A continuación se listan las características principales del proceso [2]:

1. Inicio y duración: inicia cuando un cliente se pone en contacto con la empresa. Demora el tiempo que sea necesario hasta brindar una solución.
2. Objetivo: solucionar problemas que surjan del funcionamiento del producto o condiciones del servicio.
3. Actitud: reactiva.
4. Interacción: canales específicos. Por ejemplo, telefónicos, *email*, *chat*.
5. Participantes: cliente y representante del centro de atención. Raras veces entran en juego otros trabajadores de la empresa.

#### 1.1.2. Organización de un equipo de atención al cliente

A continuación se enumeran los principales roles que podemos encontrar en el área de atención al cliente:

1. Gerente de atención al cliente: tiene bajo su responsabilidad el cumplimiento de metas estratégicas. Gestionan el volumen de casos entrantes y comunican las tendencias a otros departamentos. [3]

2. Coordinador de atención al cliente: es quien está en el día a día en contacto con los agentes. Apoya la resolución de problemas y escala aquellos que requieren validaciones o decisiones que no estén a su alcance [4].
3. Analista de atención al cliente: canaliza las quejas, reclamos y sugerencias. Se encarga de proveer soporte a los usuarios. [4]

En algunas organizaciones, dependiendo de su tamaño, se pueden encontrar más o menos mandos intermedios entre el gerente y el coordinador. Una variante puede ser tener varios coordinadores respondiendo a un supervisor. Este supervisor puede estar a cargo de la atención de reclamos para ciertos productos de la empresa que están relacionados.

En la figura 1.1 podemos ver la principal tendencia que hay en cuanto a organización de equipos. La idea es que cada equipo sea especialista en un conjunto de casos.

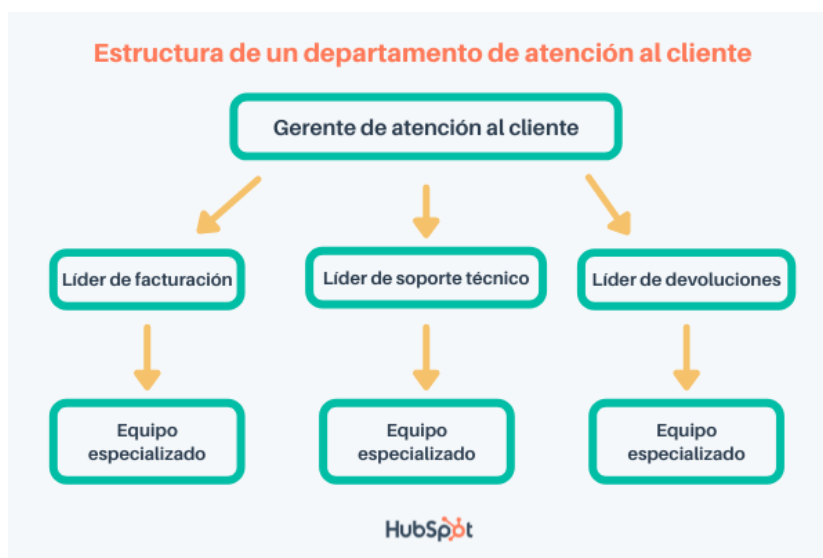


FIGURA 1.1. Organigrama de un área de atención al cliente<sup>1</sup>.

### 1.1.3. El proceso de atención al cliente

El proceso de atención al cliente consiste en una serie de pasos que se realizan para atender los reclamos y/o consultas que recibe la empresa [5].

A continuación se describen los pasos principales de un proceso de atención típico [6].

- Contacto y captura de la demanda del cliente: en esta etapa se recibe el mensaje del cliente por cualquiera de los canales establecidos y se procede a registrarlo en el sistema.
- Análisis y clasificación: se analiza la información recibida y se evalúa si es suficiente para proceder a la clasificación del caso. De lo contrario se vuelve a contactar al cliente para obtener más detalles.

<sup>1</sup>Imagen tomada de <https://blog.hubspot.es/service/que-es-atencion-al-cliente>

- Resolución: en esta etapa se busca dar una respuesta a la consulta o reclamo del cliente. En algunos casos puede involucrar varios equipos o personas, por lo que su duración es variable.
- Cierre: en esta etapa se presenta la solución al cliente y se le comunican los próximos pasos si los hubiera.

## 1.2. Motivación

Hoy en día al cliente se le hace muy fácil pasar a la competencia en caso de que la empresa no pueda cumplir con sus expectativas. Una mala CX (*Customer Experience*) puede generar un efecto de bola de nieve, ya que los usuarios que hayan tenido una mala experiencia, son mas propensos a comunicarlo con sus conocidos haciendo que la empresa no sólo pierda un cliente sino que además, se le dificulte expandir su mercado [7].

Según informes de *CX Trends* [8][9][10] y de Esteban Kolsky [11]:

- El 70 % de los consumidores gastará más en una empresa que ofrezca una buena CX.
- El 50 % de los clientes se pasaría a la competencia después de haber tenido tan solo una mala experiencia. Este valor sube a un 80 % si se les pregunta si hubieran tenido dos o más malas experiencias.
- El 72 % de los clientes compartiría una experiencia positiva con 6 o más personas.
- El 13 % de los clientes compartiría su experiencia con 15 o más personas si no está satisfecho.
- Sólo el 3.85 % de los clientes insatisfechos se lo comunican a la empresa.

Estos números muestran que la ausencia de quejas no es un signo de satisfacción. Por el contrario, es probable que los clientes hayan abandonado la empresa y estén compartiendo su insatisfacción con otras personas. También muestran la importancia de mantenerlos satisfechos, ya que son propensos a divulgarlo.

Más aún, entre las principales razones de una mala atención al cliente se encuentran tener tiempos de espera demasiado largos [12] y tener muchas derivaciones internas [13].

De todo esto se deduce que resulta muy importante que los procesos del área se realicen de forma eficiente. La utilización de modelos de IA para automatizar el proceso de clasificación de reclamos y consultas, ayuda a la empresa a responder a sus clientes con tiempos de respuesta menores y permite disponer de más analistas en la etapa de resolución y cierre. De esta manera, se genera una imagen positiva de la empresa que facilita la fidelización de clientes existentes y aumenta la incorporación de los nuevos.

### 1.3. Estado del arte

El PNL (Procesamiento del lenguaje natural) es un subcampo de la inteligencia artificial que busca enseñar a un programa informático a comprender, interpretar y generar texto en lenguaje humano. A principios del siglo XXI, el PNL experimentó un crecimiento significativo gracias a la aplicación de algoritmos de aprendizaje profundo y la disponibilidad de grandes corpus de texto [14].

#### 1.3.1. Word Embeddings

En 2003, Yoshua Bengio et al. entrenan la primer red neuronal orientada al lenguaje, utilizando vectores para representar palabras. Llamarían a este proceso “aprender una representación distribuida para cada palabra” [15].

En 2008 Ronan Collobert y Jason Weston introducen el concepto de *Word Embeddings* como una herramienta potente para tareas de PNL. Se distinguirían de sus antecesores mencionando que su objetivo era predecir la relevancia de una palabra dada la parte previa y posterior de la oración, a diferencia de el trabajo previo que buscaba predecir la probabilidad de una palabra dada la parte previa de una oración.

En 2013, Tomas Mikolov et al. publican su artículo dónde detallan las arquitecturas CBOW y *Skip-Gram*. Además liberan el primer modelo pre-entrenado llamado *Word2Vec* (basado en *Skip-Gram*) popularizando el uso de los *Word Embeddings* [16].

Posteriormente, en 2014, Pennington et al. publicarían *GloVe* como otro método de generación de *Word Embeddings* que utiliza una probabilidad de co-ocurrencia [17]. Con esta técnica, si dos palabras co-existen muchas veces, ambas palabras tienen una probabilidad alta de tener un mismo significado.

Los *Word Embeddings* se convirtieron en la herramienta principal dentro del PNL. Capturan el significado de una palabra y la traducen a una representación numérica que puede ser usada como entrada para las redes neuronales.

#### 1.3.2. Transformers

El avance del aprendizaje profundo, permitió a los investigadores desarrollar arquitecturas de redes neuronales más avanzadas y eficientes.

Sin embargo, no fue hasta 2017, cuando Vaswani et. al publican la utilización de un mecanismo de atención para desarrollar una nueva arquitectura que ellos llamarían *Transformers* que se marca un verdadero hito [18]. Las redes más usadas hoy en día están basadas en esta mejora.

Básicamente, en los *Transformers* se reemplazan las capas recurrentes que se venían utilizando hasta ese momento por “capas de atención” que codifican cada palabra en función del resto de la frase, permitiendo de esta forma, introducir el contexto en la representación matemática del texto. Por este motivo, sus *Embeddings* generados son denominados *Embeddings* contextuales.

Otra de las innovaciones introducidas es el uso de *Embeddings* posicionales. Con ellos se logra una mayor paralelización, ya que no es más necesario pasar una palabra a la vez. Agregando un valor secuencial con cada palabra, hacen posible pasarle a la red todas las palabras en simultáneo.



Estas redes fueron evolucionando en tamaño y complejidad conforme transcurría el tiempo. En la figura 1.2 se puede ver la evolución de la cantidad de parámetros de estas redes a lo largo de los años.

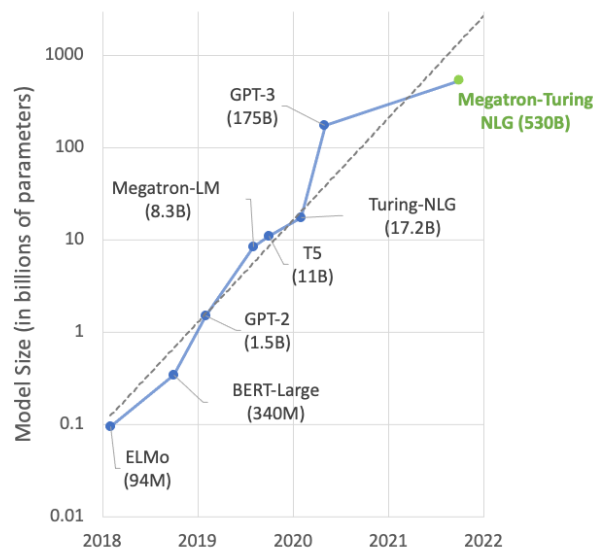


FIGURA 1.2. Evolución de las redes en cantidad de parámetros.<sup>2</sup>

Los modelos basados en *Transformers* se convirtieron en el modelo por defecto para tareas como traducción, clasificación de texto y resumen de textos por citar algunos ejemplos.

En el área de atención al cliente, los principales usos de estas redes son [19]:

- *Chatbots*: ayudan a responder consultas de forma inmediata evitando filas de espera.
- Análisis de sentimiento: para realizar análisis sobre comentarios y quejas de los clientes.
- Redireccionamiento de tickets: eliminan cuellos de botella en la parte de clasificación al redirigir el ticket directamente al área correspondiente.

## 1.4. Objetivos y alcance

### 1.4.1. Objetivos

El propósito de este trabajo fue el desarrollo de modelos de inteligencia artificial (IA) para mejorar y agilizar la atención al cliente. Para lograr este objetivo es necesario realizar la clasificación automática de los reclamos de usuario en categorías de primer y segundo nivel.

En la figura 1.3 se presenta un diagrama de alto nivel de la solución. Se observa que en primera instancia, los reclamos se reciben por correo y chat, y luego son registrados en Salesforce, el sistema de CRM que utiliza el área de atención al cliente. Estos datos se replican en BigQuery a través de un proceso orquestado por Apache Airflow.

<sup>2</sup>Imagen tomada de <https://developer.nvidia.com/blog/using-deepspeed-and-megatron-to-train-megatron-turing-nl>

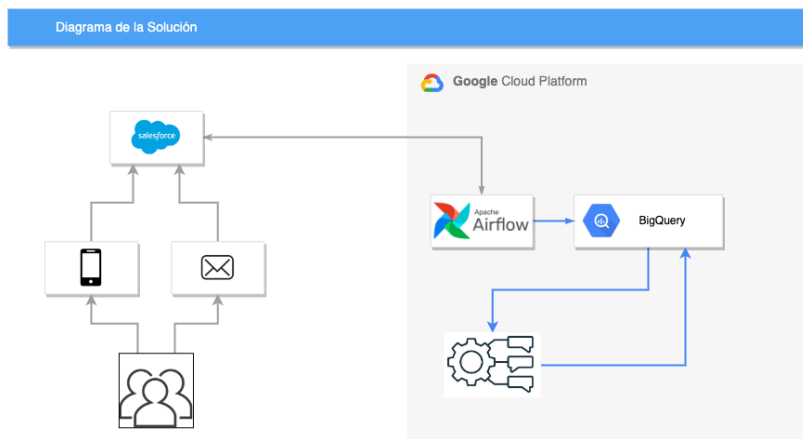


FIGURA 1.3. Diagrama general de la solución.

Los modelos de IA clasifican los casos que están en BigQuery y guardan la categoría de primer y segundo nivel en una tabla que luego el usuario podrá utilizar.

#### 1.4.2. Alcance

El alcance de este proyecto estuvo orientado a desarrollar un prototipo de solución de software que incluyó las siguientes actividades:

- Obtención de los datos: se corresponde con el análisis de las fuentes de datos disponibles, tanto para entrenar los modelos de IA como para la parte de predicción de nuevos casos.
- Análisis exploratorio de los datos: se corresponde con las actividades necesarias para generar nuevos *insights*, que sirvieron para guiar el desarrollo de los modelos.
- Modelado: se corresponde con la generación de variables a partir de los datos disponibles.
- Entrenamiento: se corresponde con el entrenamiento de los modelos a partir de las variables obtenidas. También incluye la selección del mejor modelo para cada clasificación.
- Despliegue: se corresponde con el diseño de la infraestructura para ejecutar los modelos de IA y su despliegue en un ambiente no productivo.
- Documentación: se corresponde con los documentos de soporte que explican los procesos de modelado, entrenamiento y despliegue.

El alcance de este trabajo no cubre:

- La adaptación de los modelos a nuevas categorías inexistentes en los años 2021 y 2022.
- El despliegue de los modelos en un ambiente productivo.
- El soporte de la infraestructura desplegada en el ambiente no productivo.

## Capítulo 2

# Introducción específica

En este capítulo se enumeran los requisitos que la solución debe cumplir y luego se describen principalmente las herramientas utilizadas durante el desarrollo para el entrenamiento y despliegue de los modelos para satisfacerlos.

### 2.1. Requerimientos

En esta sección se detallan la funcionalidad que el trabajo debía cubrir y las restricciones de implementación.

#### 1. Requerimientos funcionales

- a) El sistema debe poder detectar la categoría de un reclamo escrito en lenguaje natural.
- b) El sistema debe poder detectar la categoría de una consulta escrita en lenguaje natural.
- c) El usuario debe poder utilizar los resultados de la clasificación desde una base de datos.
- d) El proceso debe ser capaz de interpretar errores de ortografía.
- e) El proceso debe ser capaz de adaptarse a distinta cantidad de palabras en el mensaje.
- f) La solución debe ejecutarse en forma *batch*, corriendo diariamente y tomando los casos del día anterior.

#### 2. Requerimientos no funcionales

- a) El sistema debe estar desarrollado en lenguaje Python.
- b) El código debe ser versionado con Git.
- c) La solución debe estar desplegada sobre infraestructura de Google Cloud Platform.
- d) La salida de los modelos debe ser almacenada en BigQuery.
- e) El proceso debe ser ejecutado a través del orquestador Apache Airflow.

#### 3. Requerimientos de testing

- a) Se deben generar métricas de desempeño de los modelos con el dataset de entrenamiento y de prueba.

#### 4. Requerimientos de documentación

- a) Se debe confeccionar un documento con el diseño de la arquitectura de alto nivel.
- b) Se debe confeccionar un documento con el diseño de los modelos de IA.
- c) Se debe confeccionar un documento que especifique los datos que consumen los modelos y su origen.

## **2.2. Preprocesamiento del texto**

En esta sección se introducen las herramientas utilizadas para realizar el limpieza y preparación del texto que sirvió tanto para la parte de entrenamiento como para la parte de predicción en el despliegue.

### **2.2.1. NLTK**

## **2.3. Modelos de inteligencia artificial utilizados**

## **2.4. Herramientas de software utilizadas**

## Capítulo 3

# Diseño e implementación

### 3.1. Análisis del software

La idea de esta sección es resaltar los problemas encontrados, los criterios utilizados y la justificación de las decisiones que se hayan tomado.

Se puede agregar código o pseudocódigo dentro de un entorno `lstlisting` con el siguiente código:

```
\begin{lstlisting}[caption= "un epígrafe descriptivo"]
las líneas de código irían aquí...
\end{lstlisting}
```

A modo de ejemplo:

```
1 #define MAX_SENSOR_NUMBER 3
2 #define MAX_ALARM_NUMBER 6
3 #define MAX_ACTUATOR_NUMBER 6
4
5 uint32_t sensorValue[MAX_SENSOR_NUMBER];
6 FunctionalState alarmControl[MAX_ALARM_NUMBER]; //ENABLE or DISABLE
7 state_t alarmState[MAX_ALARM_NUMBER]; //ON or OFF
8 state_t actuatorState[MAX_ACTUATOR_NUMBER]; //ON or OFF
9
10 void vControl() {
11
12     initGlobalVariables();
13
14     period = 500 ms;
15
16     while(1) {
17
18         ticks = xTaskGetTickCount();
19
20         updateSensors();
21
22         updateAlarms();
23
24         controlActuators();
25
26         vTaskDelayUntil(&ticks, period);
27     }
28 }
```

CÓDIGO 3.1. Pseudocódigo del lazo principal de control.



## Capítulo 4

# Ensayos y resultados

### 4.1. Pruebas funcionales del hardware

La idea de esta sección es explicar cómo se hicieron los ensayos, qué resultados se obtuvieron y analizarlos.





## Capítulo 5

# Conclusiones

### 5.1. Conclusiones generales

La idea de esta sección es resaltar cuáles son los principales aportes del trabajo realizado y cómo se podría continuar. Debe ser especialmente breve y concisa. Es buena idea usar un listado para enumerar los logros obtenidos.

Algunas preguntas que pueden servir para completar este capítulo:

- ¿Cuál es el grado de cumplimiento de los requerimientos?
- ¿Cuán fielmente se pudo seguir la planificación original (cronograma incluido)?
- ¿Se manifestó algunos de los riesgos identificados en la planificación? ¿Fue efectivo el plan de mitigación? ¿Se debió aplicar alguna otra acción no contemplada previamente?
- Si se debieron hacer modificaciones a lo planificado ¿Cuáles fueron las causas y los efectos?
- ¿Qué técnicas resultaron útiles para el desarrollo del proyecto y cuáles no tanto?

### 5.2. Próximos pasos

Acá se indica cómo se podría continuar el trabajo más adelante.



# Bibliografía

- [1] Tiendanube. *Atención al cliente: qué es y claves para mejorar tu servicio.*  
<https://www.tiendanube.com/blog/que-es-servicio-atencion-cliente/>. Mar. de 2023. (Visitado 01-07-2023).
- [2] HubSpot. *¿En qué se diferencian el servicio al cliente y la atención al cliente?*  
<https://blog.hubspot.es/service/diferencia-servicio-cliente-y-atencion-cliente>. Ene. de 2023. (Visitado 01-07-2023).
- [3] HubSpot. *¿Qué hace un gerente de servicio al cliente?*  
<https://blog.hubspot.es/service/gerente-servicio-cliente>. Ene. de 2023. (Visitado 01-07-2023).
- [4] Zendesk. *Manual para un departamento de atención al cliente exitoso.*  
<https://www.zendesk.com.mx/blog/manual-de-funciones-de-servicio-al-cliente/>. Abr. de 2023. (Visitado 01-07-2023).
- [5] HubSpot. *Qué es el proceso de atención al cliente y cuáles son sus fases clave.*  
<https://blog.hubspot.es/service/proceso-atencion-cliente>. Abr. de 2023. (Visitado 01-07-2023).
- [6] Zendesk. *Entienda lo que son las fases del proceso de atención al cliente y en que puedes ayudar tener ese proceso interno en una empresa.*  
<https://www.zendesk.com.mx/blog/fases-del-proceso-de-atencion-al-cliente/>. Jun. de 2020. (Visitado 01-07-2023).
- [7] Smart Tribune. *The Importance of Customer Experience.*  
<https://blog.smart-tribune.com/en/importance-of-customer-experience>. Sep. de 2021. (Visitado 01-07-2023).
- [8] Zendesk. *CX Trends 2023.* <https://cxtrends.zendesk.com/mx>. Feb. de 2023. (Visitado 01-07-2023).
- [9] HubSpot. *Qué es la atención al cliente, elementos clave e importancia.*  
<https://blog.hubspot.es/service/que-es-atencion-al-cliente>. Abr. de 2023. (Visitado 01-07-2023).
- [10] Zendesk. *5 examples of bad customer service (and how to be great instead).*  
<https://www.zendesk.com/blog/what-is-bad-customer-service/>. Mayo de 2023. (Visitado 01-07-2023).
- [11] SuperOffice. *32 CUSTOMER EXPERIENCE STATISTICS YOU NEED TO KNOW FOR 2023.*  
<https://www.superoffice.com/blog/customer-experience-statistics/>. Feb. de 2023. (Visitado 01-07-2023).
- [12] HubSpot. *8 ejemplos de mal servicio al cliente (y cómo evitarlos).*  
<https://blog.hubspot.es/service/mal-servicio-cliente>. Ene. de 2023. (Visitado 02-07-2023).
- [13] Zendesk. *8 problemas comunes en servicio al cliente y cómo resolverlos.*  
<https://www.zendesk.com.mx/blog/problemas-comunes-con-clientes/>. Feb. de 2021. (Visitado 02-07-2023).
- [14] Piperlab. *Cómo hemos llegado hasta ChatGPT y el resto de LLMs.*  
<https://piperlab.es/2023/03/29/como-hemos-llegado-hasta-chatgpt-y-el-resto-de-llms/>. Mar. de 2023. (Visitado 02-07-2023).

- [15] Yoshua Bengio, Réjean Ducharme y Pascal Vincent. «A Neural Probabilistic Language Model». En: (2003).
- [16] Tomas Mikolov y col. «Efficient Estimation of Word Representations in Vector Space». En: (2013).
- [17] Jeffrey Pennington, Richard Socher y Christopher D. Manning. «GloVe: Global Vectors for Word Representation». En: (2014).
- [18] Ashish Vaswani y col. «Attention Is All You Need». En: (2017).
- [19] TechTarget. *5 examples of effective NLP in customer service*. <https://www.techtarget.com/searchenterpriseai/feature/5-examples-of-effective-NLP-in-customer-service>. Feb. de 2021. (Visitado 03-07-2023).