



Universidad de San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ciencias y Sistemas

**DISEÑO DE LA INVESTIGACIÓN: LINEAMIENTOS Y DISEÑO DE INTÉRPRETE DEL  
IDIOMA KAQCHIKEL AL ESPAÑOL UTILIZANDO *MACHINE LEARNING***

**Karen Estefany Recinos Mejía**

Asesorado por el M.A. Ing. Edgar Rubén Sabán Raxón

Guatemala, noviembre de 2024

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**DISEÑO DE LA INVESTIGACIÓN: LINEAMIENTOS Y DISEÑO DE INTÉRPRETE DEL  
IDIOMA K'AT'IKEL AL ESPAÑOL UTILIZANDO *MACHINE LEARNING***

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA  
FACULTAD DE INGENIERÍA  
POR

**KAREN ESTEFANY RECINOS MEJÍA**

ASESORADO POR M.A. ING. EDGAR RUBÉN SABÁN RAXÓN

AL CONFERÍRSELE EL TÍTULO DE

**INGENIERA EN CIENCIAS Y SISTEMAS**

GUATEMALA, NOVIEMBRE DE 2024

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA  
FACULTAD DE INGENIERÍA



**NÓMINA DE JUNTA DIRECTIVA**

DECANO	Ing. José Francisco Gómez Rivera (a. i.)
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Ing. Kevin Vladimir Armando Cruz Lorente
VOCAL V	Ing. Fernando José Pas González
SECRETARIA	Ing. Hugo Humberto Rivera Pérez

**TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO**

DECANO	Ing. Aurelia Anabela Cordova Estrada
EXAMINADOR	Ing. José Alfredo González
EXAMINADOR	Ing. Miguel Ángel Cancinos
EXAMINADOR	Ing. Óscar Alejandro Paz
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

## **HONORABLE TRIBUNAL EXAMINADOR**

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

**DISEÑO DE LA INVESTIGACIÓN: LINEAMIENTOS Y DISEÑO DE INTÉRPRETE DEL  
IDIOMA K'AT'IKEL AL ESPAÑOL UTILIZANDO *MACHINE LEARNING***

Tema que me fuera asignado por la Dirección de la Escuela de Estudios de Posgrado, con fecha septiembre de 2024.

A handwritten signature in black ink, appearing to read 'KAREN ESTEFANY RECINOS MEJIA', written over a horizontal line.

**Karen Estefany Recinos Mejía**



EEPFI-PP-5174-2024

Guatemala, 28 de septiembre de 2024

Director  
Carlos Gustavo Alonzo  
Escuela De Ingenieria En Sistemas  
Presente.

Estimado Carlos Gustavo Alonzo

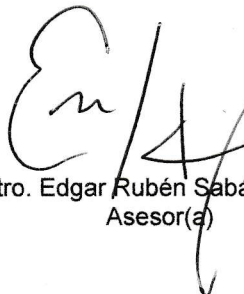
Reciba un cordial saludo de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería.

El propósito de la presente es para informarle que se ha revisado y aprobado el Diseño de Investigación titulado: **DISEÑO DE LA INVESTIGACIÓN LINEAMIENTOS Y DISEÑO DE INTERPRETE DEL IDIOMA K'QCHIKEL AL ESPAÑOL UTILIZANDO MACHINE LEARNING**, el cual se enmarca en la línea de investigación: **Área de Innovación - Dispositivos y sistemas para el desarrollo de sistemas inteligentes.**, presentado por la estudiante **Karen Estefany Recinos Mejia** carné número **201313924**, quien optó por la modalidad del "PROCESO DE GRADUACIÓN DE LOS ESTUDIANTES DE LA FACULTAD DE INGENIERÍA OPCIÓN ESTUDIOS DE POSTGRADO". Previo a culminar sus estudios en la Maestría en Artes en Tecnologías De La Inf. Y La Comunicacion.

Y habiendo cumplido y aprobado con los requisitos establecidos en el normativo de este Proceso de Graduación en el Punto 6.2, aprobado por la Junta Directiva de la Facultad de Ingeniería en el Punto Décimo, Inciso 10.2 del Acta 28-2011 de fecha 19 de septiembre de 2011, firmo y sello la presente para el trámite correspondiente de graduación de Pregrado.

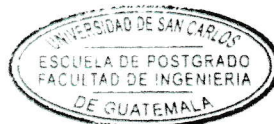
Atentamente,


*"Id y Enseñad a Todos"*

  
Mtro. Edgar Rubén Sabán Raxón  
Asesor(a)

**Edgar Rubén Sabán Raxón**  
Ingeniero en Ciencias y Sistemas  
Colegiado No 10834

  
Mtro. Marlon Antonio Pérez Turk  
Coordinador(a) de Maestría



  
Mtra. Aurelia Anabela Cordova Estrada  
Directora  
Escuela de Estudios de Postgrado  
Facultad de Ingeniería





EEP-EICS-5036-2024

El Director de la Escuela De Ingenieria En Sistemas de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del Asesor, el visto bueno del Coordinador y Director de la Escuela de Estudios de Postgrado, del Diseño de Investigación en la modalidad Estudios de Pregrado y Postgrado titulado: **DISEÑO DE LA INVESTIGACIÓN LINEAMIENTOS Y DISEÑO DE INTERPRETE DEL IDIOMA KAQCHIKEL AL ESPAÑOL UTILIZANDO MACHINE LEARNING** , presentado por el estudiante universitario **Karen Estefany Recinos Mejia**, procedo con el Aval del mismo, ya que cumple con los requisitos normados por la Facultad de Ingeniería en esta modalidad.

ID Y ENSEÑAD A TODOS

Mtro. Carlos Gustavo Alonzo  
Director  
Escuela De Ingenieria En Sistemas

Guatemala, septiembre de 2024



**USAC**  
TRICENTENARIA  
Universidad de San Carlos de Guatemala

Decanato  
Facultad de Ingeniería

24189101- 24189102

LNG.DECANATO.OIE.716.2024

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería en Ciencias y Sistemas, al Trabajo de Graduación titulado: **DISEÑO DE LA INVESTIGACIÓN: LINEAMIENTOS Y DISEÑO DE INTERPRETE DEL IDIOMA K'ATCHEKEL AL ESPAÑOL UTILIZANDO MACHINE LEARNING**, presentado por: **Karen Estefany Recinos Mejía** después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:

Ing. José Francisco Gómez Rivera  
Decano a.i.



Guatemala, noviembre de 2024

Para verificar validez de documento ingrese a <https://www.ingenieria.usac.edu.gt/firma-electronica/consultar-documento>

Tipo de documento: Correlativo para orden de impresión Año: 2024 Correlativo: 716 CUI: 2139027450101

Escuelas: Ingeniería Civil, Ingeniería Mecánica Industrial, Ingeniería Química, Ingeniería Mecánica Eléctrica, - Escuela de Ciencias, Regional de Ingeniería Sanitaria y Recursos Hidráulicos (ERIS). Postgrado Maestría en Sistemas Mención Ingeniería Vial. Carreras: Ingeniería Mecánica, Ingeniería Electrónica, Ingeniería en Ciencias y Sistemas. Licenciatura en Matemática. Licenciatura en Física. Centro de Estudios Superiores de Energía y Minas (CESEM). Guatemala, Ciudad

## **ACTO QUE DEDICO A:**

<b>Dios</b>	Por estar a mi lado siempre y guiarme en mi vida.
<b>Mis padres</b>	Blanca Mejía y Manuel Recinos, por el apoyo incondicional que me han dado y estar siempre apoyándome mis decisiones en todo momento.
<b>Mi esposo</b>	Edgar Calderón, por ser mi inspiración, apoyarme y estar a mi lado siempre, animándome a seguir adelante.
<b>Mis hermanos</b>	Christopher, Erick y Héctor Recinos, por el apoyo incondicional que me han dado.
<b>Mis abuelos</b>	Juana Salas y Hector Mejía, que me observan desde el cielo dándome su bendición para seguir adelante.
<b>Mi familia</b>	Mis tíos y primos, quienes siempre están apoyándome en todo momento.



## **AGRADECIMIENTOS A:**

<b>Universidad de San Carlos de Guatemala</b>	Por permitirme estudiar y sacar mis estudios de pregrado y posgrado.
<b>Mi esposo</b>	Edgar Calderón, quien me anima siempre a seguir adelante dándome todo su cariño, conocimiento y apoyo.
<b>Mi mamá</b>	Blanca Mejía, quien siempre me anima para que siga estudiando y para salir adelante en mis estudios, dándome su cariño y apoyo.
<b>Mi asesor</b>	Edgar Sabán, por brindarme su apoyo y conocimiento para la realización de mi trabajo de tesis.
<b>Mis profesores</b>	Por brindarme sus conocimientos durante toda mi carrera estudiantil.

## ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES .....	V
LISTA DE SÍMBOLOS .....	VII
GLOSARIO .....	IX
1. INTRODUCCIÓN .....	1
2. ANTECEDENTES .....	3
3. PLANTEAMIENTO DEL PROBLEMA .....	9
3.1. Pregunta central .....	11
3.1. Preguntas auxiliares .....	12
4. JUSTIFICACIÓN .....	13
5. OBJETIVOS .....	17
5.1. General .....	17
5.2. Específicos .....	17
6. NECESIDADES A CUBRIR Y ESQUEMA DE SOLUCIÓN.....	19
7. ALCANCES .....	23
7.1. Alcance investigativo .....	23
7.2. Alcance técnico .....	23
7.3. Alcance de resultados .....	24

8.	MARCO TEÓRICO .....	25
8.1.	Traducción automática mediante aprendizaje profundo .....	25
8.1.1.	Traducción palabra por palabra .....	26
8.1.2.	Traducción basada en reglas.....	27
8.1.3.	Traducción automática basada en estadística.....	28
8.2.	Redes neuronales recurrentes .....	31
8.2.1.	Red neuronal recurrente .....	33
8.2.2.	Red neuronal recursiva.....	35
8.2.3.	Red neuronal recurrente recursiva .....	36
8.3.	Redes neuronales profundas .....	38
8.3.1.	Aprendizaje profundo.....	39
8.3.2.	Modelos de lenguaje neuronal.....	42
8.4.	Aprendizaje profundo en traducción de máquina .....	43
8.4.1.	DNN en el proceso de traducción .....	44
8.4.1.1.	Alineación de palabras.....	44
8.4.1.2.	Selección y reordenamiento de reglas.....	45
8.4.1.3.	Modelo del lenguaje.....	46
8.4.1.4.	Traducción conjunta.....	47
8.5.	Traducción automática estadística (SMT) .....	47
8.5.1.	Enfoque de transferencia.....	51
8.5.2.	Enfoque directo.....	52
8.5.3.	Enfoque de traducción estadística.....	53
8.6.	Aprendizaje dual para la traducción automática neuronal .....	55
8.6.1.	Traducción automática neuronal.....	55
8.6.2.	Aprendizaje dual .....	57
9.	PROPUESTA DE ÍNDICE DE CONTENIDOS .....	61

10.	METODOLOGÍA.....	65
10.1.	Características del estudio.....	65
10.1.1.	Enfoque mixto .....	65
10.1.2.	Alcance explicativo.....	67
10.1.3.	Diseño experimental .....	67
10.2.	Unidades de análisis.....	68
10.3.	Variables.....	69
10.4.	Fases de estudio .....	71
11.	TÉCNICAS DE ANÁLISIS DE LA INFORMACIÓN .....	73
12.	CRONOGRAMA.....	75
13.	FACTIBILIDAD DEL ESTUDIO .....	79
13.1.	Factibilidad técnica .....	79
13.2.	Factibilidad financiera .....	80
13.3.	Fuentes de financiamiento.....	82
	REFERENCIAS .....	85



## ÍNDICE DE ILUSTRACIONES

### FIGURAS

<b>Figura 1.</b>	Entrenamiento del modelo predictivo.....	20
<b>Figura 2.</b>	Diagrama de salida de la interpretación de los idiomas kaqchikel al español .....	21
<b>Figura 3.</b>	Ejemplo del sistema de traducción palabra por palabra .....	27
<b>Figura 4.</b>	Ejemplo del sistema de traducción basado en reglas.....	28
<b>Figura 5.</b>	Ejemplo de la oración que se divide en palabras significativas ...	29
<b>Figura 6.</b>	Ejemplo de la traducción de palabras significativas al idioma destino .....	30
<b>Figura 7.</b>	Diagrama de bloques de RNN .....	32
<b>Figura 8.</b>	Estructura de un RNN desplegado en la red .....	32
<b>Figura 9.</b>	Red neuronal recurrente .....	34
<b>Figura 10.</b>	Red neuronal recursiva.....	35
<b>Figura 11.</b>	Red neuronal recurrente recursiva .....	37
<b>Figura 12.</b>	Entrenamiento e implementación de redes neuronales. ....	38
<b>Figura 13.</b>	Vista general de SMT .....	48
<b>Figura 14.</b>	Arquitecturas de los modelos (a) SMT y (b) NMT .....	50
<b>Figura 15.</b>	Ejemplo de sentencias con el enfoque de transferencia.....	51
<b>Figura 16.</b>	Ejemplo del enfoque directo de Ohidujjaman y colaboradores (2021) .....	52
<b>Figura 17.</b>	Representación de probabilidades $f$ y $e$ .....	53
<b>Figura 18.</b>	Ejemplo del enfoque de traducción estadística.....	54
<b>Figura 19.</b>	Algoritmo para el aprendizaje dual .....	59
<b>Figura 20.</b>	Cronograma de actividades de la investigación.....	77

## **TABLAS**

<b>Tabla 1.</b>	DNN en traducción automática.....	43
<b>Tabla 2.</b>	Variables en estudio .....	69
<b>Tabla 3.</b>	Cronograma propuesto para la investigación .....	75
<b>Tabla 4.</b>	Factibilidad financiera del estudio con costos en quetzales .....	81

## LISTA DE SÍMBOLOS

Símbolo	Significado
%	Porcentaje
Q	Quetzal (moneda de Guatemala)





## **GLOSARIO**

<b>Objetivo</b>	Finalidad que se pretende alcanzar con la realización de un proceso o actividad que conlleva una serie de pasos ordenados.
<b>Traducción</b>	En lingüística, trasladar el significado de un mensaje de un idioma a otro, con el objetivo de amplificar su comprensión.



# 1. INTRODUCCIÓN

Este trabajo de investigación está conformado por dos partes, la primera describirá el planteamiento del problema: resolver las brechas lingüísticas que tiene el idioma kaqchikel al español, utilizando herramientas tecnológicas de aprendizaje automático enfocadas específicamente en la interpretación de voz a texto del idioma kaqchikel al español. También incluirá los objetivos y un resumen del marco metodológico donde se definirá qué tipo de investigación se trabajará para la realización del sistema. La segunda parte estará dividida en 6 capítulos, donde se describirá la solución que tendrá la investigación al problema establecido.

En el capítulo uno se describirá antecedentes de la utilización de la tecnología de *machine learning* que se han realizado, para facilitar la comunicación entre idiomas. En el segundo capítulo se detallará la justificación por la cual se realizará la investigación, cómo puede beneficiar a que la comunicación fluya de forma más rápida y precisa utilizando herramientas tecnológicas para resolver las brechas lingüísticas que se tienen en la comunicación de diferentes idiomas.

En el capítulo tres se describirán los alcances de la investigación que se tendrán para la implementación del sistema tecnológico que ayude a la solución del problema. En el capítulo cuatro es donde se describirá el marco teórico, las teorías, métodos y metodologías que se han establecido para la creación de sistemas de traducción con aprendizaje automático para el procesamiento del lenguaje natural en idiomas.

Seguidamente, en el capítulo cinco, se describirá la presentación de los resultados de la investigación: Se describirá cómo estará implementada la solución del problema, detallando los diseños de la base de datos y del sistema; Se hará una descripción de la recolección de los datos para la implementación de la red neuronal que generará el modelo de traducción para los idiomas del kaqchikel al español; funcionamiento general del sistema; las pruebas de traducción de palabras y oraciones de *speech to text*, utilizando el modelo generado de la traducción de los idiomas kaqchikel al español y una descripción del prototipo a implementar para la solución del problema.

Por último, en el capítulo seis es donde se discutirán los resultados obtenidos de la implementación del prototipo de interpretación del idioma maya kaqchikel al español, analizando los resultados de las traducciones echas con el prototipo implementado, comparando los resultados de las pruebas con los objetivos, dando una descripción de estos y definiendo trabajos futuros que se puede realizar después de la investigación.

## **2. ANTECEDENTES**

La comunicación es la forma esencial que las personas utilizan diariamente para poder comunicar sus inquietudes, expresiones, intereses o hablar con sus familiares, entre otros. Dependiendo de donde estén ubicados los grupos de personas pueden hablar determinados idiomas o dialectos, los cuales pueden llegar a tener brechas de comunicación entre quienes no hablan en su idioma materno y quieren comunicarse entre sí.

Hoy en día la comunicación ha ido evolucionando por medio de las tecnologías de la información, ya no es como se conocía antes. La comunicación siempre fue de forma presencial, luego por medios como la imprenta, la radio, la televisión y vías telefónicas. Llegó la era digital y la comunicación tuvo una revolución con la utilización del Internet, donde ya personas de todo el mundo se podían comunicar. La comunicación se volvió más fácil y de rápido acceso para todos, utilizando medios como correos electrónicos, mensajes de texto, redes sociales, realidad virtual, realidad aumentada y la aparición de la inteligencia artificial.

Özkaya (2023) describe que la integración de la inteligencia artificial (IA) y el aprendizaje automático (ML) en los servicios lingüísticos ha marcado el comienzo de una nueva era de comunicación rápida, que trasciende las barreras lingüísticas y hace que el contenido multilingüe sea más accesible. Estas tecnologías ya no se limitan a la mera automatización, sino que han evolucionado hasta convertirse en herramientas sofisticadas que contribuyen significativamente a la calidad, velocidad y diversidad de las tareas relacionadas con el lenguaje.

El aprendizaje automático puede ayudar a interpretar los idiomas, utilizando las redes neuronales para su aprendizaje, para que sea capaz de traducir lo que las personas quieren expresar en su idioma materno al de las personas que se les quiere transmitir el mensaje, y así solucionar esas brechas lingüísticas, como se indica en el artículo de Özkaya (2023) sobre cómo puede ayudar la inteligencia artificial en la interpretación de los idiomas.

Fiorini *et al.* (2023) indican que el inglés se considera en general la lengua franca de la comunicación académica. Un uso tan generalizado tiene ciertamente la ventaja de facilitar los intercambios en un panorama de investigación cada vez más internacionalizado.

Sin embargo, este dominio lingüístico también genera desigualdades entre los investigadores y limita la difusión del conocimiento científico dentro de las comunidades de habla no inglesa. Aunque la traducción podría identificarse rápidamente como la solución, la comunicación académica ha estado históricamente marcada por una escasez de recursos humanos y financieros para apoyar los procesos de traducción tradicionales (Fiorini *et al.*, 2023).

El objetivo que presentan Fiorini *et al.* (2023) es un enfoque multiusuario para la evaluación de la traducción automática para su uso en la comunicación académica, presentando una metodología de evaluación y ajuste diseñada para cumplir con las necesidades de los diferentes usuarios objetivo (traductores, investigadores, lectores). Dado el enfoque de la conferencia, el artículo describe con más detalle la metodología de evaluación relacionada con la persona traductor. El documento también incluye conclusiones preliminares generales e información sobre el trabajo de evaluación en curso.

Ohidujjaman *et al.* (2021) hablan del intercambio estratégico de conocimientos sin fronteras y el desarrollo de la comunicación en que interactúan los dialectos, además de otros factores importantes como la educación, la medicina, los negocios, la investigación y otros están ampliamente difundidos en todo el mundo en función de diversas jergas. La expresión bilingüe o multilingüe es el estándar de contar con recursos lingüísticos nuevos o desconocidos junto con sus recursos.

El esfuerzo inicial del estudio de Ohidujjaman *et al.* (2021) es el de implementar los enfoques TA (traducción automática) para el procesamiento del idioma inglés al bengalí y viceversa. El énfasis del estudio es que se identifican las distintas ambigüedades junto con sus mejores soluciones.

Ciertos enfoques de traducción automática, como la traducción palabra a palabra, directa, de transferencia, interlingua, basada en corpus y estadística, están sobreviviendo y pocos de ellos se implementan en este procesamiento inteligente del lenguaje natural (SNLP) para enviar el idioma de origen al idioma de destino (Ohidujjaman *et al.*, 2021).

Ramírez *et al.* (2021) realizan el proyecto MultiTraiNMT Erasmus+, que tiene como objetivo desarrollar un plan de estudios abierto e innovador en traducción automática neuronal (NMT) para estudiantes de idiomas y traductores como ciudadanos multilingües. Estos autores describen que la traducción automática podría ayudar a abordar el desajuste entre el objetivo deseado de la UE de tener ciudadanos multilingües que hablen al menos dos idiomas extranjeros y la situación actual en la que los ciudadanos generalmente están muy por debajo de este objetivo.



Ramírez *et al.* (2021) describen que la traducción automática se considera un recurso que puede ayudar a los ciudadanos en su intento de adquirir y desarrollar habilidades lingüísticas si se les capacita de manera informada y crítica. Para ello comparten materiales de capacitación que consisten en un libro de texto de acceso abierto, una aplicación web de NMT de código abierto llamada MutNMT con fines de capacitación y correspondientes actividades.

Tapo *et al.* (2020) indican que los idiomas de bajos recursos presentan desafíos únicos para la traducción automática (neural). Abordan el caso del bambara, una lengua mande para la cual los datos de entrenamiento son escasos y que requiere cantidades significativas de preprocesamiento. Más que la situación lingüística del bambara en sí, el contexto sociocultural en el que viven los hablantes de bambara plantea desafíos para el procesamiento automatizado de esta lengua.

Tapo *et al.* (2020) presentan el primer conjunto de datos paralelos para la traducción automática de Bambara hacia y desde el inglés y el francés, así como los primeros resultados comparativos sobre la traducción automática hacia y desde bambara. Estos autores reflexionan sobre los desafíos al trabajar con lenguajes de bajos recursos y proponen estrategias para hacer frente a la escasez de datos en la traducción automática (TA) de bajos recursos.

Ohidujjaman *et al.* (2021), Ramírez *et al.* (2021) y Tapo *et al.* (2020) presentan en sus trabajos de investigación métodos y herramientas para implementar inteligencia artificial con aprendizaje automático, generando una guía de cómo pueden ayudar para la interpretación del idioma maya kaqchikel al español, traduciendo de forma más clara y rápida.

En Guatemala no se presentan evidencias de que existan trabajos realizados de interpretación del idioma kaqchikel al español utilizando *machine learning* con *speech to text*, solo un traductor de palabras escritas en español al kaqchikel desarrollado por Mynor Xico en su página web.



### **3. PLANTEAMIENTO DEL PROBLEMA**

Guatemala es un país multiétnico, pluricultural y multilingüe en que se hablan distintos idiomas mayas, aproximadamente alrededor de 22, los cuales son transmitidos por generaciones. Dependiendo del lugar, las comunidades pueden hablar uno o incluso más idiomas mayas y, al momento de que estas personas migran del departamento en que se encuentren, no pueden comunicarse con las demás personas que no hablan su idioma materno, provocando grandes brechas lingüísticas.

Es probable que estas brechas lingüísticas se deben a la falta de traductores jurados de los idiomas mayas en el país, ya que en Guatemala se vive en una sociedad en que los idiomas mayas no son aceptados por todos los ciudadanos, desde la imposición del español como lengua oficial durante la colonización, provocando una baja cantidad de iniciativas, tanto privadas como públicas, que tengan como fin la conservación y reducción de la brecha lingüística.

Otros factores que influyen en la brecha lingüística son la limitación de materiales didácticos y recursos lingüísticos de los idiomas mayas, pues una gran cantidad se perdió durante la colonización. Todo esto dificulta la enseñanza y el aprendizaje de los idiomas mayas, y a su vez provoca que existan pocos sistemas tecnológicos relacionados.

Con el crecimiento de la población en Guatemala, la escasez de traductores jurados de los idiomas mayas para procesos legales se hace más notable, provocando que la población guatemalteca que no habla español como

su idioma materno no pueda comunicarse con la misma fluidez que los demás, impidiéndoles agilizar los procesos como el acceso a la educación, colocar una denuncia, recibir una adecuada atención médica o ejercer sus derechos y obligaciones que tienen como ciudadanos en el país, entre otras cosas.

Por eso existe una necesidad de traductores jurados que puedan traducir de forma rápida y precisa los idiomas mayas al español, para que así los ciudadanos guatemaltecos puedan expresarse en cualquier lugar en su idioma materno sin problema de las brechas lingüísticas.

Hoy en día existen pocas herramientas tecnológicas, como diccionarios de palabras de kaqchikel al español y viceversa, algunos más avanzados pueden traducir oraciones simples, todos estos son con traducciones de texto, pero no existe evidencia de que puedan traducir el idioma maya kaqchikel de voz a texto.

La falta de estas herramientas tecnológicas para la interpretación de los idiomas mayas, entre ellos el kaqchikel, es notable, ya que hoy en día existen herramientas capaces de traducir a muchos idiomas oficiales de todo el mundo, y la mayoría de los ciudadanos con acceso a estas tecnologías no están interesados por conservar los idiomas mayas ni por romper las brechas lingüísticas, al considerar los idiomas mayas como antiguos y alejados por completo de la tecnología.

Derivado a que el idioma kaqchikel es muy antiguo, muchas palabras que se utilizan hoy en día para describir el entorno no tienen una traducción específica al idioma kaqchikel, provocando que las personas que lo hablan lo mezclen con el idioma español.

Otros de los problemas que se tiene con las brechas de comunicación de los idiomas mayas al español es que muchos guatemaltecos no logran tener un buen desarrollo integral, como una buena educación o un buen trabajo, debido a que la mayoría de los procesos legales son realizados en el idioma español.

Al suceder que una persona que hable un idioma maya quiera colocar un negocio fuera de su comunidad, en una de las ciudades, o estudiar en algún otro lugar donde no hablen su idioma materno, provoca que muchos decidan quedarse en su comunidad o realizar trabajos que no requieran de mayor comunicación con otras personas que no hablen su idioma, generando también una desigualdad social y económica, ya que se sectoriza los trabajos a los cuales tienen acceso. Muchos de estos problemas se deben a la falta de herramientas tecnológicas para que sea posible la interpretación de los idiomas mayas y no se vuelvan excluyentes entre las comunidades.

Las razones indicadas anteriormente, y considerando que el idioma kaqchikel es el idioma maya más hablado en el país, y es el que contiene más programas de aprendizaje, lo convierten en candidato para lograr que la inteligencia artificial pueda aprender a traducirlo correctamente al español, implementando una red neuronal de *machine learning* capaz de lograr traducir oraciones orales sencillas. Derivado del problema, se plantean las siguientes preguntas:

### **3.1.1. Pregunta central**

¿Cómo implementar un prototipo tecnológico de traducción con *machine learning* que ayude a la comunidad hablante del idioma maya Kaqchikel a comunicarse de forma precisa y rápida con las personas hablantes del idioma español para resolver las brechas lingüísticas de entendimiento de los idiomas?

### 3.1.2. Preguntas auxiliares

- ¿Cómo se obtendrá la recolección de los datos para la construcción del *data set* que tendrá el sistema de traductor del idioma maya kaqchikel al español?
- ¿Qué modelo de procesamiento del lenguaje natural de *machine learning* se utilizará para el traductor del idioma kaqchikel al español?
- ¿Cómo implementar un prototipo de página web capaz de traducir oraciones simples de *speech to text* de los idiomas kaqchikel al español?

## 4. JUSTIFICACIÓN

El presente trabajo tiene como propósito contribuir en las áreas de dispositivos y sistemas para el desarrollo de sistemas que apliquen inteligencia artificial y *machine learning*, que ayude a las tecnologías de la información y la comunicación para fortalecer la visión multicultural de Guatemala y para el apoyo al desarrollo humano nacional.

Se sabe que la comunicación es la forma esencial que utilizan las personas para interactuar. Guatemala se considera un país multilingüe en el que se habla aproximadamente 22 idiomas mayas, dependiendo de dónde se encuentren ubicados los ciudadanos pueden incluso hablar más de un idioma maya como su idioma materno y, en el momento de que estas personas migran de su ubicación hacia otro departamento, o hacia la ciudad y quieren comunicarse con las demás personas que solo hablan el español, se provocan grandes brechas lingüísticas por la falta de entendimiento.

La falta de traductores jurados de los idiomas mayas hace que gran parte de la población guatemalteca no pueda llegar a comunicarse de forma fluida con los que no hablan su mismo idioma, ya que en Guatemala el idioma oficial es el español. Puede que estas personas no reciban adecuada atención por la brecha lingüística que se tiene entre los idiomas. Por ello es importante realizar un sistema tecnológico que ayude a interpretar lo que desean comunicar de una forma más rápida y precisa. El idioma maya con mayor población hablante y que tiene programas de aprendizaje es el kaqchikel, que es un candidato para poder realizar un intérprete de este al español.



El desarrollo de un sistema tecnológico de traducción básico con *machine learning* y redes neuronales que interpreten el idioma kaqchikel al español puede optimizar los procesos de traducción manual, reduciendo así el tiempo y el costo asociados a la traducción tradicional y, en lugares en que no se cuente con traductores para el idioma, puede ser una alternativa más rápida para traducir lo que se quiere comunicar al español.

Un sistema tecnológico de traducción básico automatizado puede ampliar el acceso a la información y los recursos en línea para las personas que no dominan el idioma, tanto los que hablan idiomas mayas como español. Esto puede ser especialmente beneficioso en áreas como la educación, la salud y el gobierno, donde una información precisa y actualizada es crucial para la toma de decisiones y el bienestar de las personas.

También un sistema tecnológico de traducción básico preciso y eficiente puede eliminar las barreras y brechas lingüísticas del idioma que dificultan la comunicación entre personas de diferentes culturas y regiones, ya que no se cuenta con un sistema tecnológico específico que pueda traducir de forma precisa lo que se quiere comunicar.

Realizar el intérprete del idioma kaqchikel al español permitiría una mayor colaboración nacional e internacional en proyectos de investigación, desarrollo tecnológico y transferencia de conocimiento a las demás personas de la sociedad que quieran expresar sus conocimientos y aportar sus ideas con los demás, fortaleciendo la visión multicultural de Guatemala y apoyando al desarrollo humano nacional.

Con el sistema tecnológico de traducción básica se puede promover la inclusión social, al permitir que las personas con diferentes idiomas participen plenamente en la sociedad. Esto puede contribuir a reducir la discriminación y fomentar la igualdad de oportunidades para todos, como también puede facilitar el comercio nacional e internacional al permitir una comunicación fluida entre empresarios, compradores y vendedores de diferentes regiones y países. Esto puede generar nuevas oportunidades comerciales, impulsar el crecimiento económico y fomentar la cooperación nacional e internacional.

Esta investigación puede guiar a otras personas a que puedan implementar un sistema capaz de interpretar diferentes idiomas mayas o dialectos que se hablen en el mundo y necesiten el apoyo para traducir de los mismos al español.



## 5. OBJETIVOS

### 5.1. General

Implementar un prototipo tecnológico de traducción con *machine learning* que ayude a la comunidad hablante del idioma maya kaqchikel a comunicarse de forma precisa y rápida con las personas hablantes del idioma español, para resolver las brechas lingüísticas de entendimiento de los dos idiomas.

### 5.2. Específicos

1. Describir y diseñar cómo se recolectará los datos para la construcción del *data set* que tendrá el sistema de traductor del idioma maya kaqchikel al español.
2. Definir el modelo de procesamiento de lenguaje natural de *machine learning* que utilizará el traductor del idioma kaqchikel al español.
3. Implementar un prototipo de página web capaz de traducir oraciones simples de *speech to text* de los idiomas kaqchikel al español.



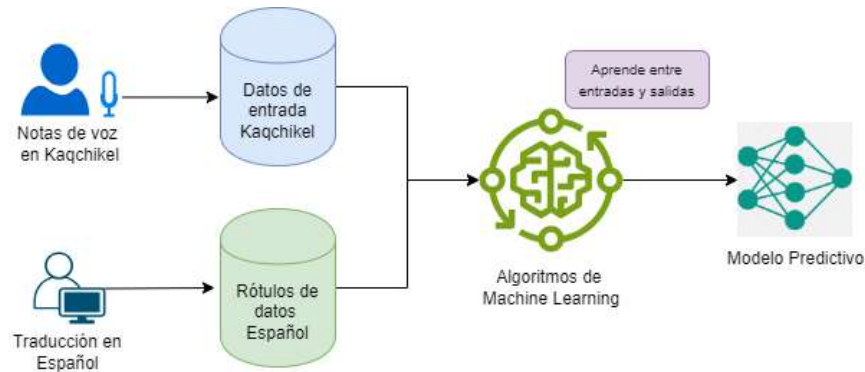
## **6. NECESIDADES A CUBRIR Y ESQUEMA DE SOLUCIÓN**

Para abordar el problema de las brechas lingüísticas del idioma maya kaqchikel al español se propone un sistema tecnológico de traducción básico que sea capaz de traducir oraciones básicas habladas de kaqchikel al español, por medio de aprendizaje automático, que se alimentará con una serie de muestras a la red neuronal que aprende la forma de traducir el idioma kaqchikel al español.

El proceso, como lo muestra la Figura 1, consiste en que se tendrá como información de entrada palabras y oraciones habladas en kaqchikel y la traducción al español de forma escrita, estas serán guardadas en una base de datos para su interpretación por medio de algoritmos del aprendizaje automático y el procesamiento de lenguaje natural, generando así el modelo predictivo de los datos aprendidos.

**Figura 1.**

*Entrenamiento del modelo predictivo*



*Nota.* Entrenamiento del modelo predictivo de los idiomas kaqchikel al español. Elaboración propia, realizado en draw.io.

En la rotulación de los datos se ingresarán las traducciones de las palabras y oraciones solicitadas en español de las notas de voz que se obtuvieron de las diferentes personas a colaborar con el sistema. Teniendo ya los datos de las traducciones de español y kaqchikel se ingresan como entradas al algoritmo de *machine learning*.

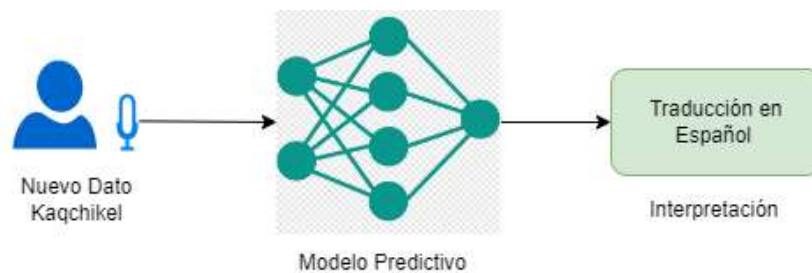
El algoritmo de *machine learning* consiste en una red neuronal que se diseñará para interpretar las distintas pronunciaciones que se hablan en kaqchikel y que corresponden a las traducciones específicas del español. Una vez entrenado el algoritmo de aprendizaje automático, se genera el modelo predictivo, que consiste en una red neuronal que se encontrará lista para la interpretación del idioma kaqchikel al español.

Ya teniendo el modelo predictivo, se puede utilizar para la realización de pruebas con nuevos datos para interpretar lo que se quiere decir en kaqchikel ya traducido al español, con las oraciones básicas y palabras con las que fue

entrenado el modelo, generando así la salida de la traducción escrita del español, como se muestra en la Figura 2.

**Figura 2.**

*Diagrama de salida de la interpretación de los idiomas kaqchikel al español*



*Nota.* Diagrama de la interpretación de los idiomas kaqchikel al español. Elaboración propia, realizado en draw.io.

Para mostrar las respuestas de la traducción que genera el modelo predictivo y la utilización de este, se realizará una interfaz de usuario de página web que sea intuitiva y fácil de utilizar, donde se realizarán las pruebas de traducción del idioma kaqchikel al español.

Para la recolección de los datos se realizará una serie de entrevistas a varias personas que puedan traducir palabras y oraciones del español al kaqchikel, grabando las notas de voz de cada una de las pronunciaciones en una base de datos con la traducción al español.





## **7. ALCANCES**

### **7.1. Alcance investigativo**

Se tomarán en consideración investigaciones, artículos, libros y tesis que utilizan *machine learning* para traducir idiomas con diferentes redes neuronales, tomando en cuenta lo siguiente:

- Qué lineamientos se necesitan para la recolección de los datos e implementación del sistema.
- Qué tipo de red neuronal se utilizó para el aprendizaje de los idiomas.
- Nivel de profundidad de la red neuronal con que cuenta el sistema de traducción de los idiomas.
- Cómo fue implementado, con qué algoritmos y qué métodos se utilizaron para los sistemas de traducción.
- Número de retroalimentación para la red neuronal.
- Qué tan preciso se realizan las traducciones.

### **7.2. Alcance técnico**

Se implementará un algoritmo de red neuronal que sea capaz de traducir frases simples de kaqchikel hablado con nota de voz y se interpretará al idioma español de forma escrita. Para eso se utilizará *speech to text* para el prototipo de traducción automática de los idiomas kaqchikel al español.

Esta investigación se desarrollará con las siguientes especificaciones técnicas:

- Base de datos de los idiomas con su traducción.
- Tipo de red neuronal.
- Profundidad de la red neuronal y número de capas.
- Tensor Flow con *speech to text* para las notas de voz.
- De forma local o plataformas en la nube como Amazon Web Services (AWS) de forma gratuita.
- Aprendizaje a partir de pocos ejemplos de traducción.
- Preprocesamiento de los datos en la base de datos de los idiomas.

### **7.3. Alcance de resultados**

A continuación, se detallarán los resultados que tendrá la investigación del prototipo y los lineamientos que tiene el intérprete del idioma kaqchikel al español.

- La construcción de un *data set* del idioma kaqchikel al español.
- Un modelo de traducción básico con interfaz de usuario, con las capacidades de traducir frases y oraciones cortas entre los idiomas kaqchikel al español.

## 8. MARCO TEÓRICO

### 8.1. Traducción automática mediante aprendizaje profundo

El aprendizaje profundo es un enfoque utilizado recientemente para la traducción automática. A diferencia de la traducción automática tradicional, la traducción automática neuronal es una mejor opción para una traducción más precisa y también proporciona un mejor rendimiento. La red neuronal profunda (*Deep Neural Network* - DNN), se puede utilizar para mejorar los sistemas tradicionales con el fin de hacerlos más eficientes (Singh, *et al.*, 2017).

Singh *et al.* (2017) indican que se necesitan diferentes técnicas y bibliotecas de aprendizaje profundo para desarrollar un mejor sistema de traducción automática, como las redes neuronales recurrentes (*Recurrent Neural Network* - RNN), memoria a corto y a largo plazo (*Long Short-term Memory* – LSTM), entre otros, para entender el sistema que convertirá la oración del idioma de origen al idioma de destino. Adaptar las redes adecuadas y las estrategias de aprendizaje profundo es una buena opción porque ajusta el sistema para maximizar la precisión del sistema de traducción en comparación con otros.

Ali *et al.* (2021) indican que la traducción automática (*Machine Translation* – MT) traduce la oración del idioma de origen en una oración de lenguaje específica. Así la aplicación de MT en varios idiomas se ha convertido en parte de la vida diaria. Básicamente, los sistemas de traducción automático analizan la oración ingresada y construyen su estructura gramatical para generar la traducción a la estructura del idioma de destino.

Ali *et al.*, (2021) proponen un estudio en el que establecieron un sistema eficiente que analizará, comprenderá y generará lenguajes que los humanos usan de forma natural. El sistema MT propuesto contiene cuatro partes principales: (1) análisis, (2) codificación, (3) decodificación y (4) generación. En la parte de análisis la oración fuente se divide en palabras significativas y cada palabra se verifica léxicamente. En la parte de codificación se utilizan dos técnicas de aprendizaje profundo: redes neuronales recurrentes (RNN) y codificaciones.

Se envía una palabra a la vez a los RNN y genera un conjunto de números utilizando algoritmos de codificación. En la parte de decodificación, se entrena a otro RNN para genera la palabra en el idioma de destino a partir del idioma de origen. En la parte de generación, esas palabras se clasifican en categorías sintácticas basadas en verbo, adverbio, sustantivo, pronombre, entre otros. Esas palabras pueden intercambiarse sus posiciones y combinarse con otras palabras para crear oraciones significativas en el idioma objetivo. Entonces es posible generar el conjunto de números a partir de oraciones en el idioma de origen usando RNN y el algoritmo de codificación, y así resultará más fácil traducir la oración al idioma de destino (Ali, *et al.*, 2021).

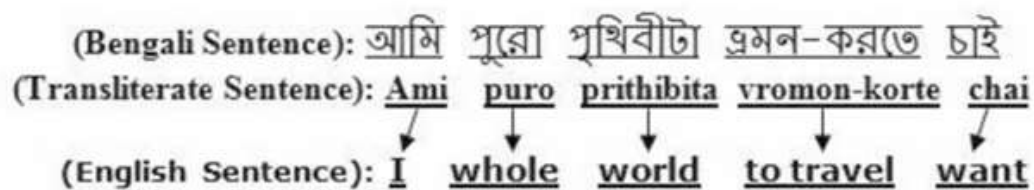
#### **8.1.1. Traducción palabra por palabra**

Ali *et al.* (2021) indican que este es el proceso de traducción más fácil y sencillo. El concepto principal de este sistema es remplazar cada palabra de una oración con la palabra traducida en el idioma de destino. La figura 3 muestra un ejemplo de traducción del bengalí al inglés utilizando la técnica palabra por palabra. El proceso de implementación de esta técnica es sencillo porque requiere un diccionario para la traducción de palabras. Pero la precisión

resultante de esta técnica es pobre, porque no le importa ninguna regla gramatical ni el orden de las palabras del idioma objetivo. En la Figura 3 se muestra un ejemplo del sistema de traducción palabra por palabra.

### Figura 3.

*Ejemplo del sistema de traducción palabra por palabra*



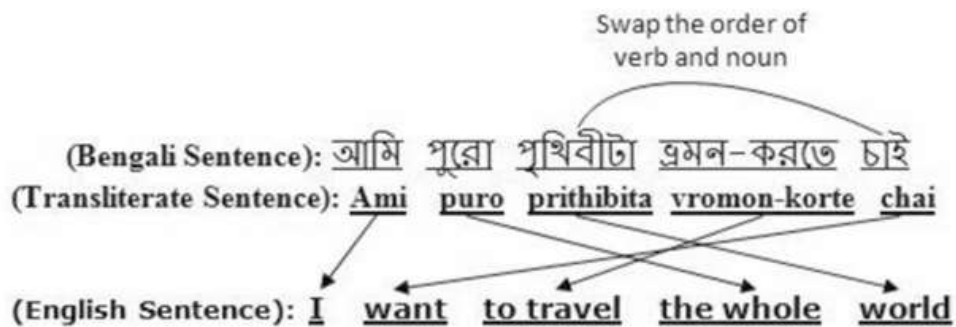
*Nota.* Sistema palabra por palabra. Obtenido de M. Ali *et al.* (2021). Machine translation using deep learning for universal networking language based on their structure. *International Journal of Machine Learning and Cybernetics*, 12(8), p. 2365. <https://link.springer.com/article/10.1007/s13042-021-01317-5>

#### 8.1.2. Traducción basada en reglas

Ali *et al.* (2021) indican que para que el sistema anterior sea más preciso, se deben agregar gramáticas específicas del idioma y un sistema de traducción basado en reglas. Por ejemplo, el orden de los verbos y los sustantivos puede intercambiarse porque en el idioma bengalí los verbos generalmente van después del sustantivo según la estructura sujeto-objeto-verbo (SOV), a diferencia del inglés, donde se sigue la estructura SVO. Y a veces dos palabras pueden traducirse en una sola palabra. La figura 4 muestra un ejemplo de traducción del bengalí al inglés utilizando gramática y contexto. Si se agregan más gramáticas y reglas, el sistema será más eficiente.

**Figura 4.**

*Ejemplo del sistema de traducción basado en reglas*



*Nota.* Sistema basado en reglas. Obtenido de M. Ali *et al.* (2021). Machine translation using deep learning for universal networking language based on their structure. *International Journal of Machine Learning and Cybernetics*, 12(8), p. 2366. <https://link.springer.com/article/10.1007/s13042-021-01317-5>

### 8.1.3. Traducción automática basada en estadística

Ali *et al.* (2021) indican, dado que los sistemas de traducción basados en reglas contienen muchas limitaciones, que se desarrollaron nuevos sistemas de traducción. En lugar de gramáticas y reglas de oración, el nuevo sistema de traducción utiliza estadísticas y probabilidades para traducir una oración.

Ali *et al.* (2021) indican que este sistema de traducción automática estadística requiere una gran cantidad de datos de entrenamiento. Lo interesante de dos sistemas de traducción automática estadística es que, aunque es poco probable, dos sistemas de traducción anteriores no generan una sola traducción para una frase. En lugar de generar una sola traducción generan todas las traducciones posibles y las colocan en términos de rango.

Finalmente, genera la frase traducida más baja como resultado, donde todo el proceso se ha completado en tres pasos que son:

Paso 1: se divide la oración en palabras significativas. En el primer paso, se divide la oración original en palabras simples y significativas que puedan traducirse fácilmente. La figura 5 muestra cómo la oración original se divide en palabras simples con significado. Aquí cada palabra subrayada es una palabra significativa (Ali *et al.*, 2021).

### Figura 5.

*Ejemplo de la oración que se divide en palabras significativas*

(Bengali Sentence): আমি পুরো পৃথিবীটা ভ্রমণ-করতে চাই  
(Transliterate Sentence): Ami puro prithibita vromon-korte chai

*Nota.* Oración en palabras significativas. Obtenido de M. Ali *et al.* (2021). Machine translation using deep learning for universal networking language based on their structure. *International Journal of Machine Learning and Cybernetics*, 12(8), p. 2367.  
<https://link.springer.com/article/10.1007/s13042-021-01317-5>

Paso 2: traducir cada palabra significativa al idioma de destino de todas las formas posibles. En el segundo paso, se descubre como los humanos tradujeron esas palabras de todas las formas posibles. No solo el simple diccionario de traducción interviene en la búsqueda de estas palabras, sino también la forma en que los humanos reales traducen estas mismas palabras (Ali *et al.*, 2021).



Este proceso es de gran ayuda para descubrir todas las formas posibles de traducir cada palabra. La Figura 6 muestra todas las formas posibles de traducir cada palabra de una oración. Dado que cada palabra contiene muchas palabras traducidas en el idioma de destino, algunas palabras traducidas son más significativas y frecuentes que otras. Se ha proporcionado una puntuación para cada palabra traducida en función de la frecuencia con la que aparecen estas traducciones en los datos de entrenamiento (Ali *et al.*, 2021).

Paso 3: hacer todas las oraciones posibles y seleccionar la adecuada. Finalmente, el tercer paso utiliza todas las combinaciones posibles de palabras traducidas para encontrar la oración posible. Se pueden generar más de 2000 oraciones diferentes solo a partir de las palabras traducidas que se muestran en la figura 6 (Ali *et al.*, 2021).

## Figura 6.

*Ejemplo de la traducción de palabras significativas al idioma de destino*



*Nota.* Traducción de palabras al idioma de destino. Obtenido de M. Ali *et al.* (2021). Machine translation using deep learning for universal networking language based on their structure. *International Journal of Machine Learning and Cybernetics*, 12(8), p. 2368. <https://link.springer.com/article/10.1007/s13042-021-01317-5>

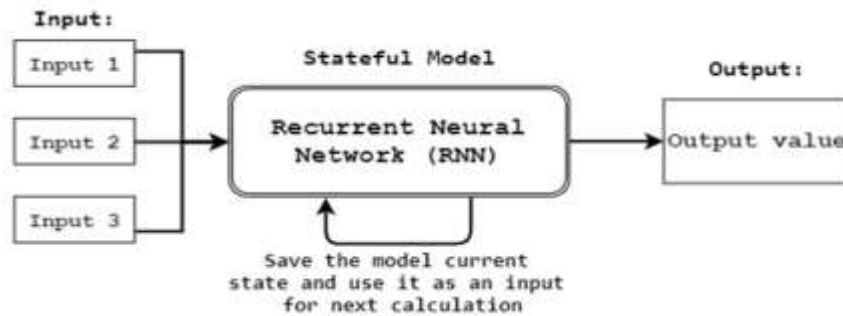
## **8.2. Redes neuronales recurrentes**

Ali *et al.* (2021) indican que al poder utilizar la información secuencial en el cálculo es la idea principal de los RNN. En una red neuronal, todas las entradas y salidas son independientes entre sí. Pero esto no es eficiente para la mayoría de los sistemas. Cuando se desea construir un sistema que pueda predecir la siguiente palabra que se va a escribir, en ese caso es necesario saber la palabra actual, de lo contrario, al sistema le resultará difícil predecir la siguiente palabra. En este tipo de aplicaciones, las redes neuronales no pueden funcionar bien.

Para resolver estos problemas se requiere un sistema que pueda utilizar el resultado de los estados anteriores para los cálculos del siguiente estado. RNN es una versión ligeramente actualizada de una red neuronal, donde el resultado del estado anterior es una de las entradas para el siguiente estado. A diferencia de las redes neuronales, la siguiente salida depende de la salida anterior en las RNN. En la siguiente página, la Figura 7 muestra el diagrama de bloques de RNN y la figura 8 muestra la estructura de una RNN despegada en una red:

**Figura 7.**

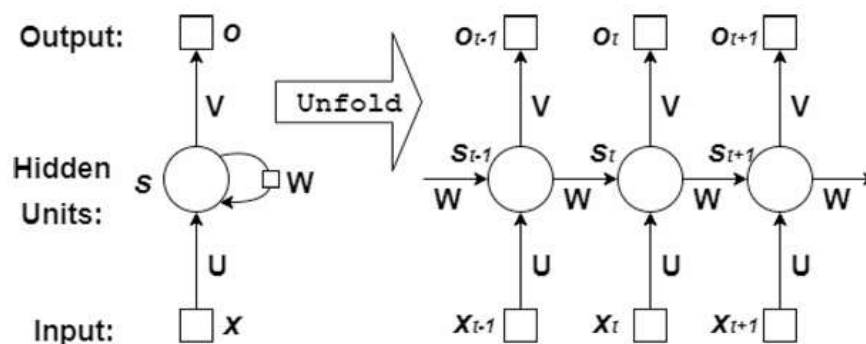
*Diagrama de bloques de RNN*



*Nota.* Diagrama de bloques de RNN. Obtenido de M. Ali *et al.* (2021). Machine translation using deep learning for universal networking language based on their structure. *International Journal of Machine Learning and Cybernetics*, 12(8), p. 2369. <https://link.springer.com/article/10.1007/s13042-021-01317-5>

**Figura 8.**

*Estructura de un RNN desplegado en la red*



*Nota.* RNN desplegado en la red. Obtenido de M. Ali *et al.* (2021). Machine translation using deep learning for universal networking language based on their structure. *International Journal of Machine Learning and Cybernetics*, 12(8), p. 2370. <https://link.springer.com/article/10.1007/s13042-021-01317-5>

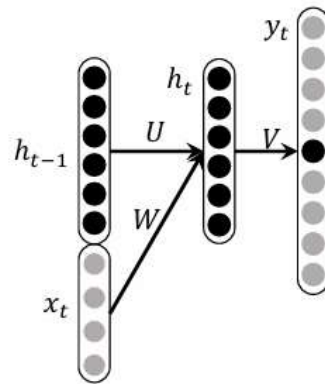
### **8.2.1. Red neuronal recurrente**

Liu *et al.* (2014) indicaron que las redes neuronales recurrentes se suelen utilizar para el procesamiento de secuencias, como los modelos de lenguaje. Los métodos de procesamiento de secuencias comúnmente utilizados, como el modelo oculto de Markov (*Hidden Markov Model* - HMM) y el modelo de lenguaje de n-gramas, solo utilizan un historial limitado para la predicción.

En HMM el estado anterior se utiliza como historial, y para el modelo de lenguaje de n-gramas (por ejemplo, n es igual a 3), el historial son las dos palabras anteriores. Se propone que la red neuronal recurrente utilice información histórica ilimitada y tenga conexiones recurrentes en estados ocultos, de modo que la información historia se pueda utilizar de forma circular dentro de la red durante un tiempo arbitrariamente largo.

**Figura 9.**

*Red neuronal recurrente*



*Nota.* Red neuronal recurrente. Obtenido de S. Liu, N. Yang, M. Li, M. & M. Zhou, (2014). A recursive recurrent neural network for statistical machine translation. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. (p. 1491) Long Papers. <https://aclanthology.org/P14-1140.pdf>

Como se muestra en la Figura 9, la red contiene tres capas, una capa de entrada, una capa oculta y una capa de salida. La capa de entrada es una concatenación de  $h_{t-1}$  y  $x_t$ , donde  $h_{t-1}$  es un vector de valor real, que es la información histórica desde el momento 0 hasta  $t - 1$ .  $x_t$  es la incrustación de la palabra de entrada en el momento  $t$ . La incrustación de palabras  $x_t$  está integrada con el historial anterior  $h_{t-1}$  para generar la capa oculta actual, que es un nuevo vector de historial  $h_t$  (Liu, *et al.*, 2014).

Con base en  $h_t$ , podemos predecir la probabilidad de la siguiente palabra, que forma la capa de salida  $y_t$ . El nuevo historial  $h_t$  se utiliza para la predicción futura y se actualiza con nueva información de la incrustación de palabras  $x_t$  de forma recurrente (Liu, *et al.*, 2014).

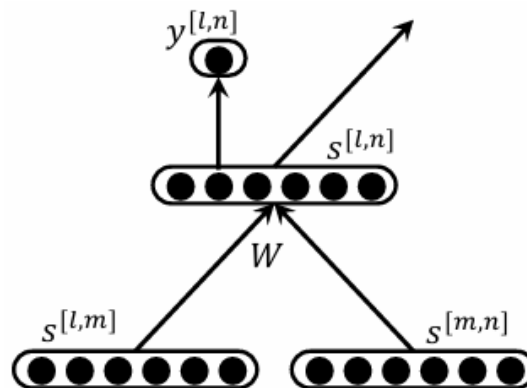
### 8.2.2. Red neuronal recursiva

Liu *et al.* (2014) indicaron que además de la estructura secuencial anterior, la estructura de árbol también suele construirse en diversas tareas del procesamiento del lenguaje natural (*Natural Language Processing* - NLP), como el análisis y la decodificación SMT. Para generar una estructura de árbol, se introducen redes neuronales recursivas para el análisis del lenguaje natural.

Al igual que con las redes neuronales recurrentes, las redes neuronales recursivas también pueden utilizar información histórica ilimitada del subárbol enraizado en el nodo actual. Las redes neuronales recursivas binarias comúnmente utilizadas generan la representación del nodo padre, con las representaciones de dos nodos hijos como entrada.

**Figura 10.**

*Red neuronal recursiva*



*Nota.* Red neuronal recursiva. Obtenido de S. Liu, N. Yang, M. Li, M. & M. Zhou, (2014). A recursive recurrent neural network for statistical machine translation. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. (p. 1492) Long Papers. <https://aclanthology.org/P14-1140.pdf>

Como se muestra en la Figura 10,  $s^{[l,m]}$  y  $s^{[m,n]}$  son las representaciones de los nodos secundarios y se concatenan en un vector para ser la entrada de la red,  $s^{[l,m]}$  es la representación generada del nodo padre,  $y^{[l,n]}$  es la puntuación de confianza de que tan plausible es que se crea el nodo padre. Por otro lado,  $L$ ,  $m$ ,  $n$ , son los índices de la cadena. Por ejemplo, para el análisis del lenguaje natural,  $s^{[l,m]}$  es la representación del nodo padre, que podría ser un nodo NP o VP, y también es la representación de todo el subárbol que cubre de  $L$  a  $n$  (Liu, *et al.*, 2014).

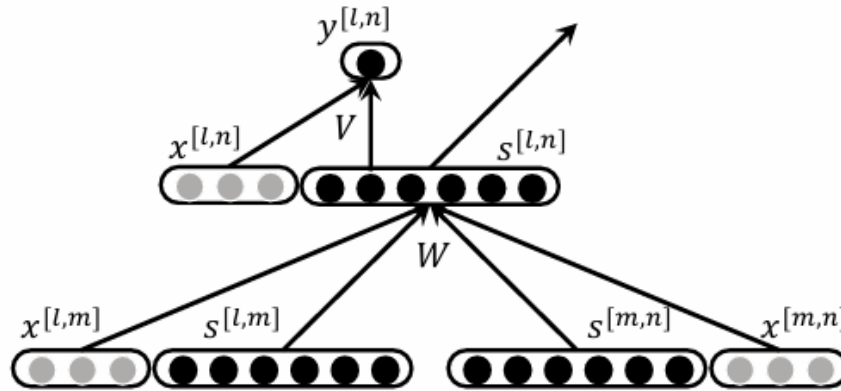
### 8.2.3. Red neuronal recurrente recursiva

Liu *et al.* (2014) indican que la incrustación de palabras  $x_t$  se integran como nueva información de entrada en redes neuronales recurrentes para cada predicción, pero en las redes neuronales recursivas no se utiliza información de entrada adicional excepto los dos vectores de representación de los nodos secundarios.

Sin embargo, es crucial cierta información global que no puede ser generada por las representaciones secundarias, es crucial para el rendimiento de SMT, como el score del modelo de lenguaje y la puntuación del modelo de distorsión. Para integrar dicha información global y también mantener la capacidad de generar una estructura de árbol, combinamos la red neuronal recurrente y la red neuronal recursiva para formar una red neuronal recurrente recursiva (R2NN) (Liu *et al.*, 2014).

**Figura 11.**

*Red neuronal recurrente recursiva*



*Nota.* Red neuronal recurrente recursiva. Obtenido de S. Liu, N. Yang, M. Li, M. & M. Zhou, (2014). A recursive recurrent neural network for statistical machine translation. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. (p. 1493) Long Papers. <https://aclanthology.org/P14-1140.pdf>

Como se muestra en la Figura 11, con base en la red recursiva, se agregan tres vectores de entrada  $x^{[l,m]}$  para el nodo hijo  $[l,m]$ ,  $x^{[m,n]}$  para el nodo hijo  $[m,n]$ , y  $x^{[l,n]}$  para el nodo padre  $[l,n]$ . Reciben el nombre de vectores de entrada recurrentes, ya que se toman prestados de redes neuronales recurrentes. Los dos vectores de entrada recurrente  $x^{[l,m]}$  y  $x^{[m,n]}$  se concatenan como la entrada de la red, con las representaciones originales del nodo hijo  $s^{[l,m]}$  y  $s^{[m,n]}$ . El vector de entrada recurrente  $x^{[l,n]}$  se concatenan con la representación del nodo padre  $s^{[l,n]}$  para calcular el puntaje de confianza  $y^{[l,n]}$  (Liu *et al.*, 2014).

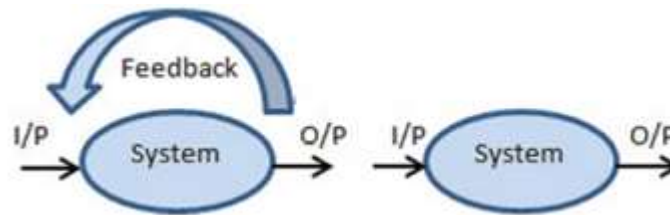


### 8.3. Redes neuronales profundas

Las redes neuronales con más de una capa oculta se conocen como redes neuronales profundas (DNN). Estas redes primero entran en la fase de formación y luego se implementan para resolver el problema. La estructura y el proceso de formación de DNN dependen de la tarea asignada, como indican Singh *et al.* (2017) en la Figura 12:

**Figura 12.**

*Entrenamiento e implementación de redes neuronales*



*Nota.* Entrenamiento e implementación de redes neuronales. Obtenido de S. Singh *et al.* (2017). Machine translation using deep learning: An overview. In: *2017 International conference on computer, communications and electronics (comptelix)*, (p. 162). IEEE. [10.1109/COMPTELIX.2017.8003957](https://doi.org/10.1109/COMPTELIX.2017.8003957)

Zhang y Zong (2015) indican que la traducción automática (*Machine Translation* - MT) tiene como objetivo encontrar, para la fase del idioma de origen, la fase más probable en el idioma de destino que comparta el significado más similar. Básicamente, la MT es una tarea de predicción de secuencia a secuencia. Donde ofrece una descripción general completa de las aplicaciones de DNN en MT desde dos puntos de vista: aplicación indirecta, que intenta mejorar los sistemas de MT estándar, y aplicación directa, que adopta DNN para diseñar un modelo de MT puramente neuronal.

La aplicación indirecta diseña nuevas funciones con DNN en el marco de los sistemas MT estándar, que constan de múltiples submodelos (como la selección de traducción y los modelos de idioma). Por ejemplo, los DNN se pueden aprovechar para representar el idioma de origen, la semántica del contexto y predecir mejor los candidatos a traducción (Zhang y Zong, 2015).

La aplicación directa considera la MT como una tarea de predicción de secuencia a secuencia y, sin utilizar ninguna información de los sistemas de MT estándar, diseña dos redes neuronales profundas: un codificador, que aprende representaciones continuas de oraciones en el idioma de origen, y un decodificador, que genera la oración en el idioma de destino con la representación de la oración de origen (Zhang y Zong, 2015).

### **8.3.1. Aprendizaje profundo**

El aprendizaje profundo es una técnica nueva, ampliamente utilizada en diferentes aplicaciones de aprendizaje automático. Permite que el sistema aprenda como un humano y mejore la eficiencia con capacitación.

Los métodos de aprendizaje profundo tienen la capacidad de representar características mediante el uso de aprendizaje supervisado/no supervisado; incluso existen capas superiores y más abstractas. Aprendizaje profundo utilizado actualmente en aplicaciones de imágenes, análisis *big data*, reconocimiento de voz, traducción automática (Singh, *et al.*, 2017).

Yang y Yue (2020) indican que el aprendizaje profundo (*Deep Learning* - DL) se origina en el campo del aprendizaje automático y tiene como objetivo construir y simular la red neuronal profunda (*Deep Neural Network* – DNN) del cerebro humano para su análisis y aprendizaje. Las DNN han mostrado

ventajas sobresalientes en la resolución de algunos problemas complejos, porque puede simular las neuronas del cerebro humano para llevar a cabo una transmisión profunda de múltiples capas para interpretar datos, lo cual se ha verificado en el campo del reconocimiento de voz.

Wen y Yue (2021) indican que la idea básica del aprendizaje profundo es utilizar hábilmente el aprendizaje no supervisado para entrenar y ajustar cada capa de la red. Hay dos pasos para desarrollar el aprendizaje profundo: primero, construir neuronas de solo una capa para cada capa de la red y limitar la cantidad de tiempos de entrenamiento para las redes de una sola capa.

Segundo, después de completar el paso 1, se utilizó el algoritmo de Wake-Sleep para optimizar y ajustar todo tipo de parámetros. Los pesos de cada capa de la red neuronal (excepto la capa superior de la red neuronal) se procesan en dos direcciones para convertirlo en un modelo gráfico. Luego se utiliza el algoritmo Wake-Sleep para ajustar los pesos de diferentes direcciones (arriba y abajo), de modo que los pesos de diferentes direcciones puedan estar en equilibrio. Entre ellos el algoritmo Wake-Sleep se compone de Wake y Sleep.

La cognición de la tabla Wake puede obtener el estado del nodo de cada capa de acuerdo con el peso cognitivo y las características externas y externas, y al mismo tiempo, el peso generado de cada capa se puede modificar adecuadamente mediante el descenso de gradiente. Se generará una parte de la tabla de Sleep y se puede obtener el estado subyacente según el peso generado y las características de nivel superior, y se puede modificar el peso cognitivo según sea necesario (Wen y Yue, 2021).

Wen y Yue (2021) indican el proceso de capacitación de *Deep Learning* se divide en dos partes: primero, comenzar desde abajo y entrenar capa por capa. En otras palabras, los parámetros de cada capa se entrenan sin supervisión por medio de datos de calibración. Específicamente, los parámetros de la primera capa se encuentran con datos calibrados para obtener la capa oculta correspondientemente.

Bajo la restricción de la capacidad del modelo, la estructura de datos es fácil de distinguir y aprender por el modelo, para lograr la adquisición de las características correspondientes. Al aprender la capa N-1, la entrada de la capa N-1 se pueden representar por la salida de N-1. En este punto los parámetros de cada capa se pueden obtener entrenando la capa N-1 (Wen y Yue, 2021).

Segundo, empezar desde abajo, entrenar capa por capa y ajustar los parámetros del modelo apropiadamente. Específicamente, sobre la base del primero paso, los parámetros del modelo en cada capa se optimizan y se ajustan con aprendizaje supervisado (Wen y Yue, 2021).

Esto es diferente del proceso de inicialización de la red neuronal anterior, porque los parámetros de inicialización de DL obtenidos después del primer paso de aprendizaje de la estructura de datos de entrada no pertenecen a la inicialización aleatoria, por lo que el valor inicial puede considerarse como el valor optimo global. Se puede ver que el efecto inicial DL mejora en gran medida a través del primer paso de aprendizaje de características (Wen y Yue, 2021).

### 8.3.2. Modelos de lenguaje neuronal

Luong, Kayser y Manning (2015) indican que las redes neuronales profundas (DNN) han tenido éxito en el aprendizaje de funciones más complejas que en las superficiales y se han destacado en muchas tareas desafiantes, como el habla y la visión. Estos resultados han despertado interés en la aplicación de las DNNs también a los problemas de procesamiento del lenguaje natural, específicamente en la traducción automática (TA). Recientemente se han trabajado mucho en el uso de modelos neuronales del lenguaje (NLM) para mejorar en la calidad de traducción.

Luong, Kayser y Manning (2015) indican que los modelos de lenguaje natural son fundamentalmente redes de propagación hacia adelante, pero no necesariamente limitadas a una sola capa oculta. Como cualquier otro modelo de lenguaje, los NLM especifican una distribución  $p(w | c)$ , para predecir la siguiente palabra  $w$  dado un contexto  $c$ . El primer paso es buscar incrustaciones de palabras en el contexto y concatenarlas para formar una entrada  $h^{(0)}$ , para la primera capa oculta. Luego se construyen repetidamente representaciones ocultas de la siguiente manera, para  $l = 1, \dots, n$ :

$$h^{(l)} = f(w^{(l)}h^{(l-1)} + b^{(l)})$$

Donde  $f$  es una función no lineal como  $\tan h$ . La distribución predictiva,  $p(w | c)$ , se deriva entonces utilizando el softmax estándar:

$$s = W^{(sm)}h^{(n)} + b^{(sm)}$$
$$p(w|c) = \frac{\exp(s_w)}{\sum_{w \in V} \exp(s_w)}$$

En cuanto al objetivo, la forma típica de entrenar NLM es maximizar la probabilidad de los datos de entrenamiento o, equivalentemente, minimizar el objetivo de entropía cruzada de la siguiente forma:  $\sum_{(c,w) \in T} -\log p(w|c)$  (Luong, Kayser y Manning, 2015).

#### 8.4. Aprendizaje profundo en traducción de máquina

La idea principal detrás de esto es desarrollar un sistema que funcione como traductor. Con la ayuda de la historia y experiencias pasadas, una red neuronal profunda entrenada traduce las oraciones sin utilizar una gran base de datos de reglas. La traducción automática incluye otros procesos relacionados, como la alineación de palabras, las reglas de reordenamiento, el modelo del lenguaje, entre otros. Cada proceso del procesamiento de texto tiene soluciones apropiadas, como se muestra en la Tabla 1:

**Tabla 1.**

*DNN en traducción automática*

Procesamiento de Texto	Soluciones DNN		
Alineación de palabras	FNN	RNN	
Selección de reglas de traducción	FNN	RAE	CNN
Reordenamiento y predicción de estructura	FNN	RAE	CNN
Modelo de lenguaje	RAE	Recurrent NN (LSTM, GRU)	Recursive NN
Predicción de traducción conjunta	FNN	RNN	CNN

*Nota.* DNN en traducción automática. Obtenido de S. Singh *et al.* (2017). Machine translation using deep learning: An overview. In: *2017 International conference on computer, communications and electronics (comptelix)*, (p. 163). IEEE. [10.1109/COMPTELIX.2017.8003957](https://doi.org/10.1109/COMPTELIX.2017.8003957)

#### 8.4.1. DNN en el proceso de traducción

Singh *et al.* (2017) describieron que después del procesamiento (segmentación de oraciones, tokenización, entre otros), el proceso de traducción comienza con la alineación de palabras, seguida del reordenamiento y el modelaje del lenguaje, el cual presenta lo indicado a continuación.

##### 8.4.1.1. Alineación de palabras

En la alineación de palabras, la entrada al sistema es un par de oraciones paralelas, y la salida es un par de palabras que están más relacionadas entre sí. Supongamos que se tienen la oración fuente  $S = S_1, S_2 \dots, S_n$  y la oración objetivo  $T = T_1, T_2 \dots, T_n$ , entonces  $A$  es el conjunto que denota la correspondencia de palabras entre oraciones bilingües  $A = \{(i, j), 1 \leq i \leq n, 1 \leq j \leq n\}$  Aquí  $(i, j)$  denota el par  $(S_i, T_j)$  que son traducciones entre sí (Singh, *et al.*, 2017).

La red neuronal de avance (*Feed Forward Neural Network* – FNN) se puede utilizar para la tarea de alineación de palabras, pero se ha demostrado que la red neuronal recurrente (*Recurrent Neural Network* - RNN) es una mejor opción, ya que mantiene el historial y predice la próxima alineación precisa sobre la base del historial previo de alineaciones ( $A_x$  basado en el historial previo de alineaciones  $A_1^{x-1}$ ) (Singh, *et al.*, 2017).

La incrustación de palabras es un concepto clave utilizado en el aprendizaje profundo para encontrar el valor vectorial de las palabras. La incrustación de palabras es una representación vectorial de espacio continuo y tiene la capacidad de capturar la característica semántica y sintáctica de la palabra correspondiente. Se necesita un corpus grande para el entrenamiento,

ya que puede capturar información necesaria para fines de traducción. El vector de palabras se usa como entrada de la red neuronal profunda. Una herramienta popular Word2vec está disponible para genera el vector (Singh, *et al.*, 2017).

Las implementaciones de RNN para la tarea de alineaciones de palabras no solo aprenden la incrustación de palabras bilingües, sino que también adquieren la similitud entre palabras para utilizar la amplia información contextual de manera muy efectiva.

#### **8.4.1.2. Selección y reordenamiento de reglas**

Una vez que se realiza el proceso de alineación, el proceso de traducción conduce a la fase de selección/extracción de reglas. Aquí las reglas se seleccionan/extraen en función de la alineación de palabras y luego el modelo de reordenamiento se entrena con un texto bilingüe alineado con palabras (Singh, *et al.*, 2017).

El codificador automático recursivo se entrena con ejemplos de reordenamiento que ya se generan a partir de oraciones bilingües alineadas con palabras. RAE es la suficientemente capaz de capturar el conocimiento de la información del orden de palabras de la fase (Singh, *et al.*, 2017).

El siguiente paso es ordenar y predecir la estructura de la oración. La combinación de red neuronal recursiva y red neuronal recurrente (R2NN) constituyen una buena idea para ejecutar esto. Las dos preocupaciones principales aquí son 1) cuáles dos candidatos se compusieron primero y 2) en qué orden se comprendían. Para trabajar con la estructura del árbol, la red neuronal recursiva es la mejor opción, pero si se usa RNN con ella, entonces integran sus capacidades, ya que la RNN mantendrá el historial que será útil



para el modelado del lenguaje y la red neuronal recursiva será útil para generar la estructura del árbol de manera ascendente. De forma semisupervisada el aprendizaje se utilizará para el entrenamiento. R2NN es una combinación no lineal (Singh, *et al.*, 2017).

#### 8.4.1.3. Modelo del lenguaje

Singh *et al.* (2017) indican que la red neuronal recurrente se puede utilizar para aprender este modelo en el espacio continuo. En este modelo la concatenación de vectores de palabras se introduce en la capa de entrada y la capa oculta para encontrar la probabilidad de  $T_n$  en función de  $T_1^{n-1}$ . La red recurrente se puede diseñar para el modelo del lenguaje porque funciona muy bien en tareas de aprendizaje secuencia a secuencia.

Singh *et al.* (2017) proporcionan la secuencia de entrada  $(S_1, \dots, S_n)$ , y en función de la secuencia, predecirá la secuencia de salida  $(T_1, \dots, T_n)$ . Los vectores de entrada ingresados a la red uno por uno, se concatenan con el historial previo en las capas ocultas y luego se calcula la salida en cada paso. El cálculo de la red neuronal recurrente se puede explicar con las siguientes ecuaciones:

$$h_n = \text{sigm}(W^{hs} S_n + W^{hh} h_{n-1})$$

$$T_n = W^{th} h_n$$

Dos RNN son requeridas, una para el proceso de codificación y otra para el de decodificación. Si  $(S, T)$  es el par de oraciones de origen y destino, entonces  $S_1, S_2, \dots, S_n = \text{codificador}(S_1, S_2, \dots, S_n)$  mediante la regla de la cadena, la probabilidad de condición se puede calcular como:

$$P(S|T) = P ( T|S_1, S_2, \dots, S_n )$$

El decodificador es la combinación de una red neuronal recurrente con una capa Softmax. Es difícil entrenar RNN debido a las dependencias de largo plazo. Las redes LSTM evitan los problemas que se producen con las RNN. Utilizan un algoritmo de retropropagación a través del tiempo para aprender los parámetros del modelo (Singh, *et al.*, 2017).

#### **8.4.1.4. Traducción conjunta**

Se utiliza un modelo de traducción y lenguaje conjunto para predecir la palabra de destino con la ayuda de un historial ilimitado de palabras de origen y destino. RNN es la mejor red para esto. FNN y CNN solo se ocupan del aprendizaje mediante redes, pero RNN mantiene la secuencia independientemente de si la traducción se genera de izquierda a derecha o de derecha a izquierda (Singh, *et al.*, 2017).

### **8.5. Traducción automática estadística (SMT)**

La traducción automática estadística (SMT) es un conjunto de reglas que se utilizan para traducir el idioma de origen al idioma de destino. Se utilizan algoritmos de aprendizaje automático supervisados y no supervisados para construir los modelos estadísticos. Los modelos SMT, como el modelo de transferencia, el modelo basado en corpus y el modelo interlingua, son familiares para los desarrolladores e investigadores. La idea de SMT fue lanzada por primera vez por Warren Weaver en 1949 (Ohidujjaman *et al.*, 2021).

Después de lanzar el concepto de SMT, la mayoría de los investigadores están tratando de desarrollar el fenómeno. El diagrama de flujo de trabajo es como tomar el idioma de origen y traducirlo al idioma de destino y viceversa, como se muestra en la Figura 13 (Ohidujjaman *et al.*, 2021).

**Figura 13.**

*Vista general de SMT*



*Nota.* Vista general de SMT. Obtenido de F. Ohidujjaman, S. Faysal, M. Sumon & N. Huda (2021). Automatic machine translation for bangla and english resolving ambiguities. In: *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques*. (p. 27) ICREST. [doi:10.1109/ICREST51555.2021.9331085](https://doi.org/10.1109/ICREST51555.2021.9331085). <https://ieeexplore.ieee.org/document/9331085>

Liu *et al.* (2021) representan un modelo de traducción automática estadística en el cual supone un par de oraciones  $x = \{x_1, \dots, x_i, \dots, x_I\}$  e  $y = \{y_1, \dots, y_j, \dots, y_J\}$  están en el lado de origen y destino, respectivamente.  $x_i$  es la  $i$ -ésima palabra de  $x$  e  $y_j$  es la  $j$ -ésima palabra de  $y$ .  $I$  y  $J$  son longitudes de  $x$  e  $y$ , que pueden ser diferentes. Con base en la teoría de la decisión de Bayes, formulan SMT como:

$$\hat{y} = \arg \max p(y|x) \propto \arg \max p(x|y) p(y)$$

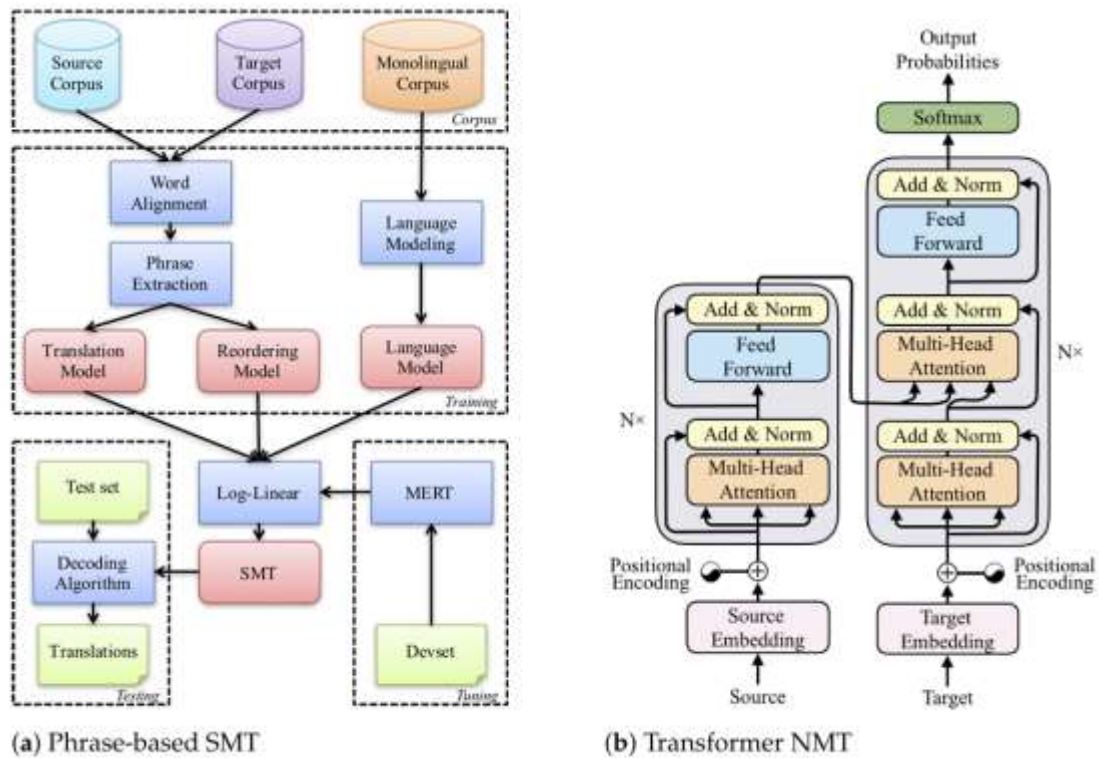
Donde  $\hat{y}$  denota el resultado de traducción con la mayor probabilidad de traducción. El problema de traducción se factoriza en  $p(x|y)$  y  $p(y)$ , que representa la probabilidad de traducción inversa y la probabilidad del modelo de lenguaje, respectivamente. El denominador  $p(x)$  se ignora, ya que permanece constante para una oración fuente dada  $x$ . La ventaja de esta descomposición es que podemos aprender probabilidades separadas para calcular  $\hat{y}$  (Liu *et al.*, 2021).

Liu *et al.* (2021) indican un modelo Log lineal, que incorpora diferentes características que contienen información de las oraciones de origen y de destino en el modelo, además de los modelos de lenguaje y traducción del enfoque del canal ruidoso original. La Figura 14 (a) describe las arquitecturas de SMT basada en frases, que consta de varios componentes como:

- Las palabras dentro del corpus paralelo se alinean y luego se extraen los pares de frases en función de los resultados de alineación de palabras.
- El modelo de traducción y el modelo de reordenamiento lexicalizado se pueden aprender usando frases alineadas.
- Se puede construir un modelo de lenguaje de n-gramas utilizando una gran cantidad de oraciones monolingües en el idioma de destino.
- Estos modelos están optimizados bajo el marco Log lineal para maximizar el rendimiento utilizando un conjunto de desarrollo.
- Con los parámetros de peso optimizados de las características de los modelos, finalmente es posible traducir el conjunto de pruebas y la puntuación de evaluación indica el rendimiento de todo el sistema.

**Figura 14.**

*Arquitecturas de los modelos (a) SMT y (b) NMT*



*Nota.* Modelos SMT y NMT. Obtenido de S. Liu, Y. Sun, & L. Wang (2021). Recent advances in dialogue machine translation. *Information*, 12(11), p. 484. <https://www.mdpi.com/2078-2489/12/11/484>

Ohidujjaman *et al.* (2021) indican que los modelos de traducción automática comunes son: el basado en palabra a palabra, el enfoque directo, el enfoque de transferencia, el basado en corpus, la interlingua y el enfoque de traducción estadística.

### 8.5.1. Enfoque de transferencia

El enfoque de transferencia involucra tres etapas como análisis transferencia y generación. Un ejemplo: buscar de la lengua de destino mediante el enfoque de transferencia. A continuación, se presenta el ejemplo en la Figura 15:

**Figura 15.**

*Ejemplo de sentencias con el enfoque de transferencia*

sentence: “তাহারা মাঠে খেলিতেছে ”

TABLE- I: BILINGUAL DICTIONARY

Bangla	English
তাহারা	They
মাঠ	field
খেলা	play

Grammar for source language: [subject] + [object] + [verb];  
[subject]= তাহারা, [object]= মাঠে, [verb] = খেলিতেছে

Grammar for target language: [subject] + [verb] + [object];  
[subject] = They, [object] = field, [verb] = play

Analysis stage:

তাহারা মাঠে খেলিতেছে = [subject] + [object] + [verb]

Transfer stage:

[subject] [object] [verb] = [subject] [verb] [object]

=> তাহারা মাঠে খেলিতেছে = তাহারা খেলিতেছে মাঠে

Generation stage:

[subject]+[object]+[verb] = [subject] + [verb] + [object]

=>তাহারা মাঠে খেলিতেছে = তাহারা খেলিতেছে মাঠে

[subject] + [verb] + [object] = [They] + [playing] + [field+in]

They [are playing] [the in field]

insertion ← → reflection

Hence the result = They are playing in the field.

*Nota.* Sentencias con enfoque de transferencia. Obtenido de F. Ohidujjaman, S. Faysal, M. Sumon & N. Huda (2021). Automatic machine translation for bangla and english resolving ambiguities. In: *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques*. (p. 28) ICREST. [doi:10.1109/ICREST51555.2021.9331085](https://doi.org/10.1109/ICREST51555.2021.9331085). <https://ieeexplore.ieee.org/document/9331085>

### 8.5.2. Enfoque directo

El enfoque directo se basa en el análisis morfológico de oraciones. Se centra en la estructura interna de las palabras y la formación de palabras. Es necesario identificar palabras en textos para determinar sus propiedades semánticas y sintácticas en el procesamiento del lenguaje natural (PLN) y la traducción automática (TA). Para analizar la estructura y formación de las palabras, se centran en las reglas y regulaciones del análisis morfológico. A continuación, se muestra un ejemplo en la Figura 16:

**Figura 16.**

*Ejemplo del enfoque directo de Ohidujjaman y colaboradores (2021)*

Source sentence "We will demand peace in the country"

Table- II: Bilingual Dictionary

Bangla	English
আমরা	We
দাবি	demand
শান্তি	peace
দেশ	country

Solution steps:

step1: Morphological analysis:

<We> <will demand> <peace> <in the country>

step2: Identify constituents:

<We> <demand + future> <peace> <in the country>

step3: Reorder according to target language:

<We> <in the country> <peace> <demand + future>

step4: Look up in the bilingual dictionary:

<আমরা> <দেশ + এ> <শান্তি> <দাবি + করব>

step5: Inflect the target language words appropriately:

আমরা দেশে শান্তি দাবি করব।

*Nota.* Ejemplo de enfoque directo. Obtenido de F. Ohidujjaman, S. Faysal, M. Sumon & N. Huda (2021). Automatic machine translation for bangla and english resolving ambiguities. In: *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques.* (p. 29) ICREST.

[doi:10.1109/ICREST51555.2021.9331085](https://doi.org/10.1109/ICREST51555.2021.9331085).

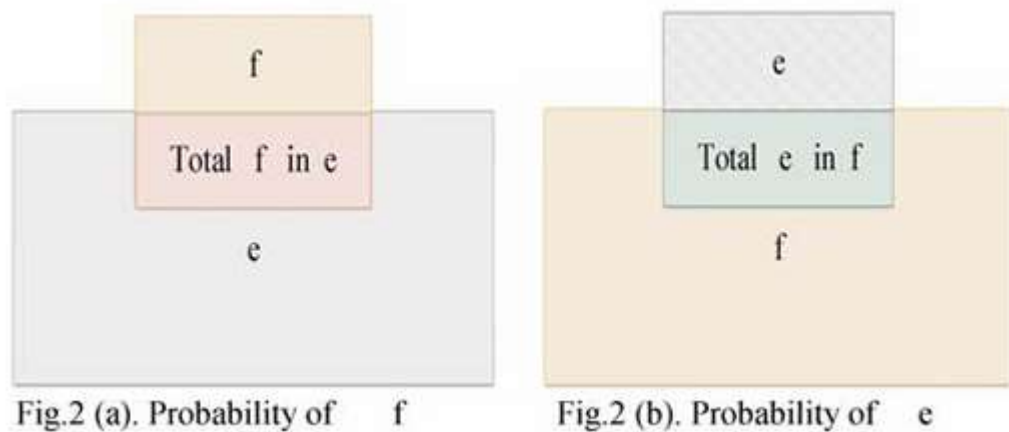
<https://ieeexplore.ieee.org/document/9331085>

### 8.5.3. Enfoque de traducción estadística

El modelo estadístico de traducción evalúa el mejor resultado a partir de los datos probables previos a través del idioma y las probabilidades de traducción. Suponiendo que el idioma extranjero =  $f$  (idioma de origen) y el idioma de destino =  $e$  (inglés).

**Figura 17.**

*Representación de probabilidades  $f$  y  $e$*



*Nota.* Probabilidades  $f$  y  $e$ . Obtenido de F. Ohidujaman, S. Faysal, M. Sumon & N. Huda (2021). Automatic machine translation for bangla and english resolving ambiguities. In: *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques*. (p. 30) ICREST. [doi:10.1109/ICREST51555.2021.9331085](https://doi.org/10.1109/ICREST51555.2021.9331085).  
<https://ieeexplore.ieee.org/document/9331085>

Suponiendo que hay “ $n$ ” oraciones en los idiomas de destino “ $e$ ”, para los idiomas de origen “ $f$ ”, entonces se tiene la Figura 17:



$$= \max \left\{ p \left( \frac{f}{e_1} \right) p(e_1), p \left( \frac{f}{e_2} \right) p(e_2), \dots, p \left( \frac{f}{e_n} \right) p(e_n) \right\}$$

$$= \max \left\{ p \left( \frac{f}{e_i} \right) p(e_i) \right\}$$

Por lo tanto  $p \left( \frac{e}{f} \right) = \max \left\{ p \left( \frac{f}{e_i} \right) p(e_i) \right\}$

[ donde  $i = 1, 2, 3, \dots, n$  y  $\hat{e} \in \{e_1, e_2, \dots, e_n\}$  ]

En la Figura 18 se presenta un ejemplo en donde aplica el enfoque de traducción estadística:

### Figura 18.

*Ejemplo del enfoque de traducción estadística*

An example: The source language sentence is such as লোকটি গরিব কিন্তু সৎ. Given the possible target sentence with translation and language probabilities:  
 $e_1$  = The man is poor and honest  
 $e_2$  = The man is poor but honest  
 $e_3$  = The man is poor so honest

Table- III: Translation and language probabilities

Target Sentences	Translation Probabilities	Language Probabilities
$e_1$	0.76	0.25
$e_2$	0.99	0.67
$e_3$	0.67	0.31

Finding the target language using statistical machine translation. It is known that,

$$p \left( \frac{f}{e} \right) = \max \left\{ p \left( \frac{f}{e_1} \right) p(e_1), p \left( \frac{f}{e_2} \right) p(e_2), p \left( \frac{f}{e_3} \right) p(e_3) \right\}$$

where  $p(f/e_1)=0.76$ ,  $p(e_1)=0.25$ ,  $p(f/e_2)=0.99$ ,  $p(e_2)=0.67$ ,  $p(f/e_3)=0.67$ ,  $p(e_3)=0.31$ . Therefore,  $P(\hat{e}/f)=\max\{0.19, 0.66, 0.21\} = 0.66 = e_2$ . Hence, the target language = The man is poor but honest.

*Nota.* Enfoque de traducción estadística. Obtenido de F. Ohidujjaman, S. Faysal, M. Sumon & N. Huda (2021). Automatic machine translation for bangla and english resolving ambiguities. In: 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques.

(p. 31) ICREST. [doi:10.1109/ICREST51555.2021.9331085](https://doi.org/10.1109/ICREST51555.2021.9331085).  
<https://ieeexplore.ieee.org/document/9331085>

## **8.6. Aprendizaje dual para la traducción automática neuronal**

He *et al.* (2016) propusieron un mecanismo de aprendizaje dual que puede aprovechar los datos monolingües (tanto en el idioma de origen como en el de destino) de una manera más efectiva. Al utilizar este mecanismo propuesto, estos datos monolingües pueden desempeñar un papel similar a los datos bilingües paralelos y reducir significativamente el requisito de datos bilingües paralelos durante el proceso de capacitación. Específicamente, el mecanismo de aprendizaje dual para traducción automática (MT).

### **8.6.1. Traducción automática neuronal**

Singh *et al.* (2017) indican que la traducción automática es un método para convertir la oración fuente de un lenguaje natural a otro lenguaje natural con la ayuda de sistemas computarizados y no es necesaria asistencia humana.

Hay diferentes enfoques disponibles para crear este tipo de sistemas, pero se necesita una técnica más sólida para crear un sistema mejor que los sistemas existentes. Una red bien entrenada conduce el sistema hacia su objetivo, que es generar un sistema de traducción más eficiente que sea capaz de proporcionar una buena precisión (Singh *et al.*, 2017).

He *et al.* (2016) indican que los sistemas de traducción automática neuronal generalmente se implementan con un marco codificador decodificador basado en red neuronal recurrente (RNN). Dicho marco aprende un mapeo probabilístico  $P(y|x)$  de una oración en el idioma de origen  $x = \{x_1, x_2, \dots, x_{T_x}\}$  a una oración en el idioma de destino  $y = \{y_1, y_2, \dots, y_{T_y}\}$  en los cuales  $x_i$  e  $y_t$  son las palabras  $i$ -th y  $t$ -th para las oraciones “x” e “y”, respectivamente.

Para ser más concreto el estudio, el codificador de NMT lee la oración fuente “x” y genera estados ocultos de  $T_x$  mediante un RNN.

$$h_i = f(h_{i-1}, x_i)$$

En ese caso  $h_i$  es el estado oculto en el momento i, y la función f es la unidad recurrente, como la unidad de memoria a corto plazo (LSTM) o la unidad recurrente cerrada (GRU). Luego, el decodificador de NMT calcula la probabilidad condicional de cada palabra objetivo  $y_t$  dadas sus palabras anteriores  $y_{<t}$ , así como la oración fuente, es decir  $P(y_t | y_{<t}, x)$ , que luego se usa para especificar  $P(y|x)$  según la regla de la cadena de probabilidad  $P(y_t | y_{<t}, x)$  viene dado como:

$$P(y_t | y_{<t}, x) \propto \exp(y_t; r_t, c_t)$$

$$r_t = g(r_{t-1}, y_{t-1}, c_t)$$

$$c_t = q(r_{t-1}, h_1, \dots, h_{T_x},)$$

Donde  $r_t$  es el estado oculto de RNN del decodificador en el momento t, calculado de manera similar por un LSTM o GRU,  $c_t$  denota la información contextual al generar la palabra  $y_t$  según los diferentes estados ocultos del decodificador.  $c_t$  puede ser una señal “global” que resume la oración x por ejemplo  $c_1 = \dots = c_{T_y} = h_{T_x}$ , o una señal “local” implementada por un mecanismo (He *et al.*, 2016).

Liu *et al.* (2021) indican que, en los últimos años, NMT ha logrado avances significativos hacia la construcción y utilización de una única red neuronal grande para manejar toda la tarea de traducción. Un modelo NMT

estándar optimiza directamente la probabilidad condicional de una oración objetivo  $y = y_1, \dots, y_J$  dada su correspondiente oración fuente  $x = x_1, \dots, x_I$ .

$$P(y|x; \theta) = \prod_{j=1}^J P(y_j|y_{<j}, x; \theta)$$

Donde  $\theta$  es un conjunto de parámetros del modelo e  $y_{<j}$  denota la traducción parcial. La probabilidad  $P(y|x; \theta)$  se define en el marco codificador-decodificador basado en redes neuronales, donde el codificador resume la oración fuente en una secuencia de representaciones  $H = H_1, \dots, H_I$  con  $H \in R^{I \times d}$ , y el decodificador genera palabras objetivo basadas en las representaciones.

Normalmente, este marco se puede implementar como una red neuronal recurrente (RNN), una red neuronal convolucional (CNN) y un transformador. El transformador se ha convertido en el paradigma NMT dominante entre los diferentes modelos, como se muestra en la figura 5 (Liu *et al.*, 2021).

### 8.6.2. Aprendizaje dual

He *et al.* (2016) presentan el mecanismo de aprendizaje dual para la traducción automática neuronal. Al notar que la TA (siempre) puede ocurrir en direcciones duales, primero diseña un juego de dos agentes con un paso de traducción hacia adelante y un paso de traducción hacia atrás, que puede proporcionar retroalimentación de calidad a los modelos de traducción dual, incluso usando datos monolingües únicamente. Luego proponen un algoritmo de aprendizaje dual, llamado DualNMT, para mejorar los dos modelos de

traducción en función de la retroalimentación de calidad proporcionada en el juego.

He *et al.* (2016) consideran dos corpus monolingües:  $D_a$  y  $D_b$ , que contienen oraciones de las lenguas A y B, respectivamente. Se debe tener en cuenta que estos corpus no están necesariamente alineados entre sí y es posible que incluso no tengan ninguna relación temática que los identifique.

Suponiendo que se tiene dos modelos de traducción (débiles) que pueden traducir oraciones de A a B y en verso, el objetivo es mejorar la precisión de los dos modelos mediante el uso de corpus monolingües en lugar de corpus paralelos. La idea básica es aprovechar la dualidad de los dos modelos de traducción. A partir de una oración en cualquier dato monolingüe, primero la traducción al otro idioma y luego la traducción hacia atrás al idioma original.

Al evaluar los resultados de esta traducción de dos saltos, se tendrá una idea de la calidad de los dos modelos de traducción y será posible mejorarlos en consecuencia. Este proceso se puede repetir durante muchas rondas hasta que ambos modelos de traducción converjan.

He *et al.* (2016) proponen el siguiente algoritmo para el aprendizaje dual, como se indica en la Figura 19:

**Figura 19.**

*Algoritmo para el aprendizaje dual*

---

**Algorithm 1** The dual-learning algorithm

---

- 1: **Input:** Monolingual corpora  $D_A$  and  $D_B$ , initial translation models  $\Theta_{AB}$  and  $\Theta_{BA}$ , language models  $LM_A$  and  $LM_B$ ,  $\alpha$ , beam search size  $K$ , learning rates  $\gamma_{1,t}, \gamma_{2,t}$ .
  - 2: **repeat**
  - 3:    $t = t + 1$ .
  - 4:   Sample sentence  $s_A$  and  $s_B$  from  $D_A$  and  $D_B$  respectively.
  - 5:   Set  $s = s_A$ .  $\triangleright$  Model update for the game beginning from A.
  - 6:   Generate  $K$  sentences  $s_{mid,1}, \dots, s_{mid,K}$  using beam search according to translation model  $P(\cdot|s; \Theta_{AB})$ .
  - 7:   **for**  $k = 1, \dots, K$  **do**
  - 8:     Set the language-model reward for the  $k$ th sampled sentence as  $r_{1,k} = LM_B(s_{mid,k})$ .
  - 9:     Set the communication reward for the  $k$ th sampled sentence as  $r_{2,k} = \log P(s|s_{mid,k}; \Theta_{BA})$ .
  - 10:    Set the total reward of the  $k$ th sample as  $r_k = \alpha r_{1,k} + (1 - \alpha) r_{2,k}$ .
  - 11:   **end for**
  - 12:   Compute the stochastic gradient of  $\Theta_{AB}$ :
 
$$\nabla_{\Theta_{AB}} \hat{E}[r] = \frac{1}{K} \sum_{k=1}^K [r_k \nabla_{\Theta_{AB}} \log P(s_{mid,k}|s; \Theta_{AB})].$$
  - 13:   Compute the stochastic gradient of  $\Theta_{BA}$ :
 
$$\nabla_{\Theta_{BA}} \hat{E}[r] = \frac{1}{K} \sum_{k=1}^K [(1 - \alpha) \nabla_{\Theta_{BA}} \log P(s|s_{mid,k}; \Theta_{BA})].$$
  - 14:   Model updates:
 
$$\Theta_{AB} \leftarrow \Theta_{AB} + \gamma_{1,t} \nabla_{\Theta_{AB}} \hat{E}[r], \Theta_{BA} \leftarrow \Theta_{BA} + \gamma_{2,t} \nabla_{\Theta_{BA}} \hat{E}[r].$$
  - 15:   Set  $s = s_B$ .  $\triangleright$  Model update for the game beginning from B.
  - 16:   Go through line 6 to line 14 symmetrically.
  - 17: **until** convergence
- 

*Nota.* Algoritmo para aprendizaje dual. Obtenido de D. He *et al.* (2016). Dual learning for machine translation. *Advances in neural information processing systems*, 29(4), p. 5. <https://proceedings.neurips.cc/paper/2016/hash/5b69b9cb83065d403869739ae7f0995e-Abstract.html>



## **9. PROPUESTA DE ÍNDICE DE CONTENIDOS**

La propuesta del índice de contenido para esta investigación está conformada por el siguiente esquema, en que se organizará el informe final y se construye de acuerdo con las indicaciones específicas. Su carácter es preliminar, pues podrá ampliarse, reducirse o modificarse en el momento de elaborar el informe final de la investigación

ÍNDICE DE ILUSTRACIONES

LISTA DE SÍMBOLOS

GLOSARIO

RESUMEN

PLANTEAMIENTO DEL PROBLEMA Y FORMULACIÓN DE PREGUNTAS

ORIENTADORAS

OBJETIVOS

RESUMEN DE MARCO METODOLÓGICO

INTRODUCCIÓN

1. ANTECEDENTES

2. JUSTIFICACIÓN

3. ALCANCES

3.1. Investigativos

3.2. Técnicos

3.3. Resultados



## 4. MARCO TEÓRICO

### 4.1 Traducción automática mediante aprendizaje profundo

#### 4.1.1 Traducción palabra por palabra

#### 4.1.2 Traducción basada en reglas

#### 4.1.3 Traducción automática basada en estadística

### 4.2 Redes neuronales recurrentes

#### 4.2.1 Red neuronal recurrente

#### 4.2.2 Red neuronal recursiva

#### 4.2.3 Red neuronal recurrente recursiva

### 4.3 Redes neuronales profundas

#### 4.3.1 Aprendizaje profundo

#### 4.3.2 Modelos de lenguaje neuronal

### 4.4 Aprendizaje profundo en traducción de máquina

#### 4.4.1 DNN en el proceso de traducción

##### 4.4.1.1 Alineación de palabras

##### 4.4.1.2 Selección y reordenamiento de reglas

##### 4.4.1.3 Modelo del lenguaje

##### 4.4.1.4 Traducción conjunta

### 4.5 Traducción automática estadística (SMT)

#### 4.5.1 Enfoque de transferencia

#### 4.5.2 Enfoque directo

#### 4.5.3 Enfoque de traducción estadística

### 4.6 Aprendizaje dual para la traducción automática neuronal

#### 4.6.1 Traducción automática neuronal

#### 4.6.2 Aprendizaje dual

## 5. PRESENTACIÓN DE RESULTADOS

### 5.1 Descripción general

### 5.2 Definición de la base de datos

- 5.2.1 Selección y diseño de recolección de los datos
    - 5.2.2 Estructuración de los datos
  - 5.3 Modelo de la solución
    - 5.3.1 Selección de metodología
    - 5.3.2 Descripción de tipo de red neuronal
    - 5.3.3 Implementación de la red neuronal
  - 5.4 Pruebas del uso del modelo
  - 5.5 Prototipo del sistema
    - 5.5.1 Especificaciones funcionales
    - 5.5.2 Módulos del sistema
    - 5.5.3 Diagramas de uso
6. DISCUSIÓN DE RESULTADOS
- 6.1 Análisis de las traducciones
  - 6.2 Comparación de resultados de las pruebas
  - 6.3 Descripción de trabajos futuros

CONCLUSIONES

RECOMENDACIONES

REFERENCIAS

ANEXOS



## **10. METODOLOGÍA**

En esta sección se definirá la metodología de la investigación a realizar, se incluye el diseño, el tipo de estudio, los alcances, las variables, las fases y los resultados esperados. La metodología a la cual esta investigación se adapta es un diseño experimental donde se realizará un prototipo de traducción del idioma kaqchikel al español, utilizando el aprendizaje automático para resolver las brechas de comunicación que se tienen durante la interacción de hablantes de los idiomas kaqchikel y español, sin traductores presentes.

### **10.1. Características del estudio**

En el primer apartado de este décimo capítulo del diseño de investigación se presenta, mediante el desarrollo de diversos incisos, todo el detalle de las características del estudio.

#### **10.1.1. Enfoque mixto**

El enfoque del estudio propuesto para la traducción básica del idioma kaqchikel al español utilizando aprendizaje automático es mixto y se cuenta también con enfoques cuantitativo y cualitativo. A continuación se describirá cada uno de estos aspectos.

El enfoque cuantitativo se basa en métricas numéricas que permiten medir objetivamente la calidad de las traducciones generadas del kaqchikel al español. Estas métricas ayudarán a comparar diferentes modelos, identificar

áreas de mejora y evaluar el impacto de cambios en el sistema. Se tienen las siguientes métricas.

- Precisión léxica: mide la frecuencia con que las palabras individuales en la traducción coinciden con las palabras en la traducción de referencia.
- Coherencia: evalúa la conexión lógica entre las diferentes partes de la traducción.
- BLEU, METEOR, ROUGE, TER: estas métricas son las más comunes para evaluar la calidad de las traducciones generadas por máquinas.
- Tiempo de respuesta: mide la velocidad con la que el sistema produce una traducción.

El enfoque cualitativo se basa en la evaluación subjetiva de la calidad de la traducción por parte de la percepción humana. Este enfoque permite identificar aspectos más sutiles de la traducción que las métricas cuantitativas pueden no capturar, entre las cuales se encuentran:

- Concisión: mide la capacidad del sistema para producir traducciones concisas y directas.
- Dominio específico: evalúa la capacidad del sistema para manejar terminología y conceptos específicos de un dominio determinado (por ejemplo, médico o legal).
- Evaluación humana: se realiza a través de juicios humanos para evaluar aspectos como la comprensibilidad, la relevancia y la calidad general de la traducción.

### **10.1.2. Alcance explicativo**

El alcance es explicativo, dado que el sistema de traducción debe predecir qué es lo que se quiere comunicar con respecto a las oraciones de entrada en el idioma origen al destino, donde se verificará qué tan precisa es la traducción del prototipo con respecto al idioma origen. Para eso se implementará una red neuronal de aprendizaje automático que tenga como entrada palabras e oraciones del idioma kaqchikel y su significado en español, lo cual se entrenará para obtener un modelo en capas y predecir la traducción del idioma kaqchikel al español.

### **10.1.3. Diseño experimental**

El diseño adoptado será experimental. Se verificará qué porcentaje de traducción puede el sistema traducir de forma correcta (de las palabras en kaqchikel al español), para eso se recolectarán los datos de los idiomas como entradas para la red neuronal, donde aprenderá automáticamente la traducción correspondiente del idioma origen al destino. El diseño experimental constará de los siguientes pasos:

- Primero: identificar todos los factores que pueden incidir en el experimento

En esta parte se establecen los datos necesarios para crear una red neuronal de aprendizaje automático con las variables de entrada, que serán los audios del idioma kaqchikel con su respectiva traducción al español de palabras y oraciones establecidas previamente, para crear el modelo del intérprete del idioma kaqchikel al español.

- Seleccionar o diseñar instrumentos de recolección de información

Aquí se diseñará las encuestas con las palabras y oraciones que tendrá el modelo del intérprete del idioma kaqchikel al español y cómo procesar los datos para que sean utilizados por el modelo de aprendizaje automático de estos idiomas.

- Seleccionar el diseño experimental apropiado

Se escogerá un tipo de red de aprendizaje profundo acorde a lo investigado en el marco teórico, se implementará su diseño y arquitectura para que el modelo del intérprete sea capaz de traducir audios en el idioma kaqchikel al español. Este modelo aprenderá con los datos recolectados de los idiomas para así generar el modelo de capas para interpretar estos idiomas.

- Seleccionar una muestra de estos grupos o clases

En esta sección se seleccionarán los datos y la realización de las pruebas con el modelo obtenido, lo cual da la red neuronal de aprendizaje profundo, para obtener los promedios de traducción de palabras correctas para la interpretación de los resultados y conclusiones para el intérprete del idioma kaqchikel al español.

## **10.2. Unidades de análisis**

La población que se utilizará en el estudio serán las personas que hablan el idioma kaqchikel y quieran comunicarse con las personas que hablan el español. Esta población se encuentra dividida en subpoblaciones dadas por grupos de personas departamentales que hablan idiomas mayas,

especialmente el kaqchikel, a cuyos hablantes está dirigida esta investigación, y a personas que pueden traducir los dos idiomas: el kaqchikel y el español. Se extraerán muestras de personas que entiendan ambos idiomas para la recolección de los datos de traducción y su verificación sobre si el modelo logra traducir correctamente la oración de origen.

### 10.3. Variables

Las variables que se tomará en cuenta para el estudio del prototipo de la traducción de los idiomas kaqchikel al español se describen en la Tabla 2 que se presenta a continuación:

**Tabla 2.**

*Variables en estudio*

Variable	Definición teórica	Definición operativa
<b>% de exactitud de la traducción</b>	Es una métrica que busca cuantificar cuán fielmente una traducción automática refleja el significado original de un texto.	Se realizarán diferentes pruebas de traducción de las palabras traducidas del kaqchikel al español y se tomará el promedio de la traducción exacta de las mismas.
<b>% de consistencias de datos</b>	Es la medida en la que los elementos traducidos de un conjunto de datos mantienen una coherencia terminológica y estilística a lo largo de todo el corpus.	Se realizarán diferentes pruebas de traducción de las palabras traducidas del kaqchikel al español y se tomará el promedio de la traducción, verificando si son consistentes o coherentes a la



Continuación de la Tabla 2.

		traducción original.
<b>Número de capas optimas</b>	Es la medida en la que indica qué tan óptima es una red neuronal, dependiendo del tamaño de corpus lingüístico, la complejidad de los idiomas y la arquitectura de la red neuronal.	Se realizará pruebas con diferentes números de capas de entrenamiento en la red neuronal.
<b>Número de rondas de entrenamiento</b>	Se refiere a la cantidad de veces que un modelo de aprendizaje automático procesa el conjunto de datos de entrenamiento durante el aprendizaje.	Se realizará pruebas con diferentes números de rondas de entrenamiento en la red neuronal.
<b>Rapidez de traducción</b>	Se refiere a la rapidez con la que un sistema de traducción automática puede procesar y traducir un texto de un idioma a otro. Se mide típicamente en palabras por segundo (pps) o caracteres por segundo (cps), y representa el rendimiento del sistema en términos de tiempo de procesamiento.	Se realizarán pruebas con las diferentes redes neuronales sobre qué tan rápido se traducen las palabras en kaqchikel al español, obtenidas del promedio de estas.

*Nota.* Variables del estudio para el intérprete del idioma kaqchikel al español utilizando aprendizaje automático. Elaboración propia, realizado con Microsoft Word.

#### **10.4. Fases del estudio**

Para la investigación del prototipo de intérprete del idioma maya kaqchikel al español utilizando aprendizaje automático se plantea las siguientes fases, en las cuales se indican las técnicas que se aplicarán y las actividades que se realizarán, describiendo cada una de ellas a continuación.

- Fase 1: revisión de literatura

En esta fase se revisará diferentes documentos como artículos, tesis y libros en que se plantean teorías o métodos de cómo se recolectan e implementan redes neuronales de aprendizaje automático para la interpretación de idiomas, realizando experimentos para ver cómo están conformadas las redes neuronales que se plantean con sus respectivos resultados.

- Fase 2: recolección de los datos

Se define cómo se realizará la recolección de los datos para la implementación de la investigación, en la cual se recolectarán las muestras de los audios del idioma maya kaqchikel con su respectiva traducción al idioma español y se realizarán encuestas a diferentes tipos de personas que puedan hablar el idioma kaqchikel y puedan traducir lo que se les indica en la encuesta, gravando cada una de las pronunciaciones de las palabras u oraciones establecidas. Estos serán los datos que servirán para la implementación de la investigación y que serán utilizados en la siguiente fase.

- Fase 3: implementación de la red neuronal

Describirá el proceso de cómo se realizará la red neuronal, su implementación, arquitectura y diseño, utilizando los datos de entrada recolectados en la fase 2. Esta red aprenderá cada una de las traducciones correspondientes de los idiomas, creando así el modelo que permitirá interpretar lo que se habla en el idioma kaqchikel al español.

- Fase 4: entrenamiento y pruebas de la red neuronal

En esta fase se utilizarán las variables del número de rondas de entrenamiento y el número de capas óptimas que tendrá el modelo para realizar diferentes experimentos donde se describirán los resultados que se tendrán al realizar las pruebas de traducción de los idiomas kaqchikel al español.

- Fase 5: análisis de resultados

Se describirán los resultados obtenidos de las pruebas realizadas en la fase anterior para encontrar el porcentaje de exactitud de la traducción, el porcentaje de consistencias de datos y la velocidad de traducción, sacando los promedios de las pruebas realizadas que tiene el modelo para la interpretación del idioma kaqchikel al español.

- Fase 6: informe final

En esta sección se llevará a cabo un informe final, colocando las conclusiones, resultados y recomendaciones que se debe tener para esta investigación, de acuerdo con los objetivos establecidos y verificando que sí se llegan a cumplir.

## **11. TÉCNICAS DE ANÁLISIS DE LA INFORMACIÓN**

En esta investigación, para la recolección de los datos que necesitará el sistema tecnológico de traducción básico del intérprete del idioma maya kaqchikel al español, se realizarán encuestas estructuradas entrevistando a las personas que hablan kaqchikel y español, donde se les indicará que proporcionen las traducciones de palabras y oraciones en kaqchikel y se grabará cada una de las pronunciaciones, para obtener los datos de entrada a la red neuronal de aprendizaje automático.

Esto generará el modelo del intérprete de los idiomas kaqchikel al español, como también pueden ser obtenidos los datos de entrada al sistema. Con la recolección de datos secundarios también se investigará si se cuenta con traducciones del idioma kaqchikel para utilizar en el sistema.

Para las técnicas de análisis de la información se recolectarán los datos de las pruebas experimentales que se realizarán con el modelo generado de la red neuronal de aprendizaje de los idiomas kaqchikel al español, realizando técnicas de estadística descriptiva y analizando los promedios y métricas que se tienen del porcentaje de las palabras traducidas correctamente del idioma origen al destino, concluyendo si el modelo cumple con los objetivos establecidos, e identificando mejoras que se pueden realizar en el modelo.

Entre las técnicas para el análisis de la información de la investigación del intérprete del idioma kaqchikel al español con las que se evaluará el sistema con las variables de la investigación se encuentran las siguientes:

- Métricas de evaluación estándar: como BLEU, que es una métrica ampliamente utilizada en el campo del Procesamiento del Lenguaje Natural (PLN). Esta métrica compara una traducción candidata con una o más traducciones de referencia, para determinar qué tan similar es la traducción candidata en comparación con las de referencia, para evaluar la calidad de las traducciones.
- Evaluación humana: se solicita a personas que hablen los idiomas kaqchikel y español que evalúen la fluidez, naturalidad y precisión de las traducciones que genere el modelo.
- Calidad de los datos: los datos de entrenamiento para el modelo deben ser limpios, consistentes y representativos del dominio.
- Calidad de las traducciones: donde se evaluarán las variables de % de la consistencia de los datos y el % de exactitud de la traducción, utilizando la entropía de Shannon, comparando la entropía de un texto fuente con la entropía de su traducción, y así se puede evaluar la fluidez y naturalidad de la traducción. Una traducción con una entropía similar a la del texto original sugiere una mayor calidad.
- Para el número de rondas de entrenamiento y capas óptimas de la red neuronal se utilizará *early stopping*: donde se monitorea continuamente estableciendo un criterio de parada, como un número máximo de rondas de entrenamiento y capas de la red neuronal, si ya no mejora en la métrica de evaluación, y si el entrenamiento se detiene cuando se cumple este criterio.

## 12. CRONOGRAMA

Se presenta en la Tabla 3 el cronograma propuesto para la investigación. La tabla está organizada en días hábiles de la semana de lunes a viernes, tomando en cuenta los días festivos o de asueto, donde se realizará cada una de las actividades de la investigación en semanas, conformando un total de 180 días, que forman 6 meses de días hábiles. Esta investigación se iniciará a trabajar a principio del mes de noviembre del 2024 y se finalizará a finales del mes de julio del 2025.

**Tabla 3.**

*Cronograma propuesto para la investigación*

<b>Actividades</b>	<b>Días</b>	<b>Inicio</b>	<b>Finaliza</b>
<b>Fase 1: Revisión de literatura</b>	<b>25</b>	<b>4/11/2024</b>	<b>6/12/2024</b>
Revisión y selección de documentos literarios acorde a la investigación	10	4/11/2024	15/11/2024
Lectura y análisis de los documentos literarios	10	18/11/2024	29/11/2024
Lectura de ensayos comparativas de trabajos similares en traducción de idiomas	5	2/12/2024	6/12/2024
<b>Fase 2: Recolección de datos</b>	<b>35</b>	<b>9/12/2024</b>	<b>31/01/2025</b>
Definición de lineamientos para la recolección de los datos	5	9/12/2024	13/12/2024
Diseño de encuestas de recolección de datos	5	16/12/2024	20/12/2024
Búsqueda de datos existentes en kaqchikel al español	5	23/12/2024	3/01/2025
Recolección de los datos	15	6/01/2025	24/01/2025
Estructuración de los datos	5	27/01/2025	31/01/2025
<b>Fase 3: Implementación de red neuronal</b>	<b>60</b>	<b>3/02/2025</b>	<b>29/04/2025</b>
Diseño de la red neuronal	5	3/02/2025	7/02/2025
Arquitectura del sistema con la red neuronal	5	10/02/2025	14/02/2025
Preparación de los datos de entrada para la red neuronal	10	17/02/2025	28/02/2025

Continuación de la Tabla 3.

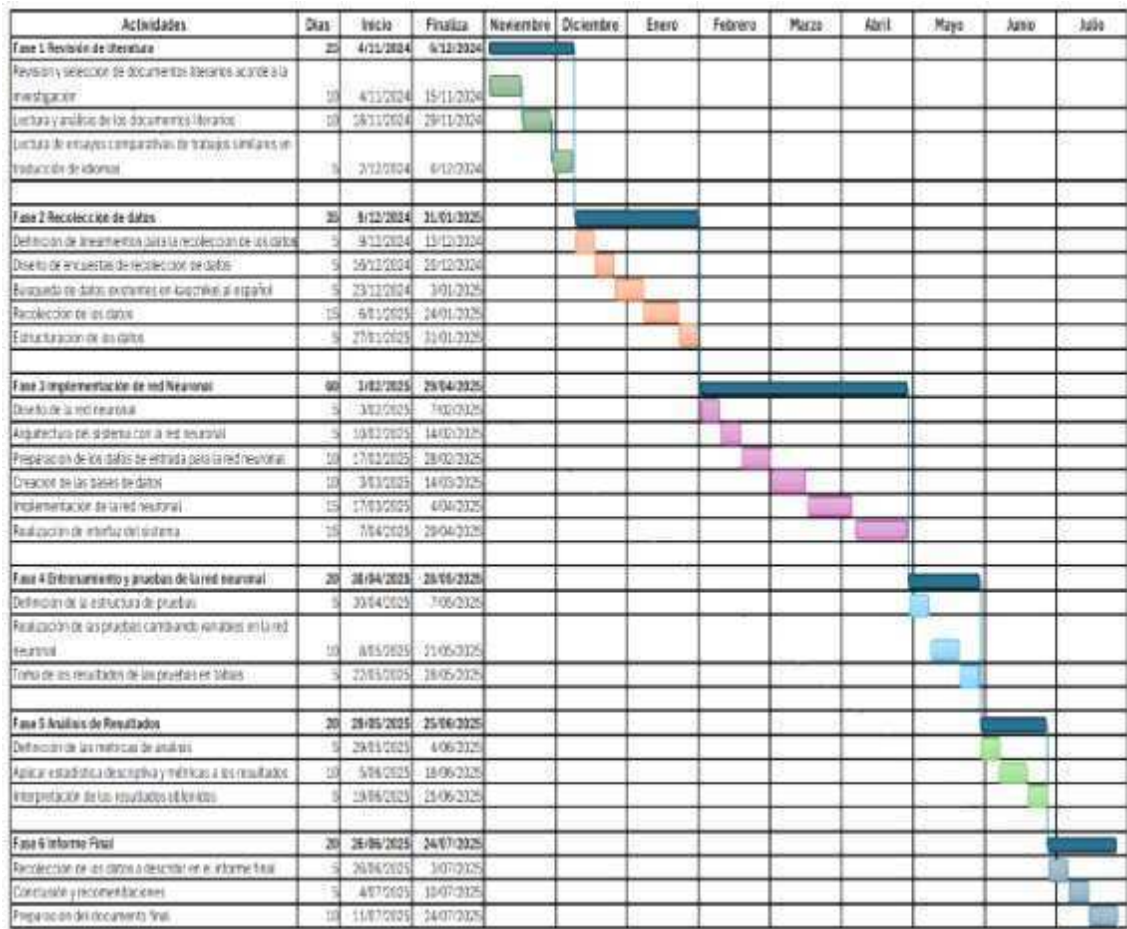
Creación de las bases de datos	10	3/03/2025	14/03/2025
Implementación de la red neuronal	15	17/03/2025	4/04/2025
Realización de interfaz del sistema	15	7/04/2025	29/04/2025
<b>Fase 4: Entrenamiento y pruebas de la red neuronal</b>	<b>20</b>	<b>30/04/2025</b>	<b>28/05/2025</b>
Definición de la estructura de pruebas	5	30/04/2025	7/05/2025
Realización de las pruebas cambiando variables en la red neuronal	10	8/05/2025	21/05/2025
Toma de los resultados de las pruebas en tablas	5	22/05/2025	28/05/2025
<b>Fase 5: Análisis de resultados</b>	<b>20</b>	<b>29/05/2025</b>	<b>25/06/2025</b>
Definición de las métricas de análisis	5	29/05/2025	4/06/2025
Aplicar estadística descriptiva y métricas a los resultados	10	5/06/2025	18/06/2025
Interpretación de los resultados obtenidos	5	19/06/2025	25/06/2025
<b>Fase 6: Informe final</b>	<b>20</b>	<b>26/06/2025</b>	<b>24/07/2025</b>
Recolección de los datos a describir en el informe final	5	26/06/2025	3/07/2025
Conclusión y recomendaciones	5	4/07/2025	10/07/2025
Preparación del documento final	10	11/07/2025	24/07/2025

*Nota.* Propuesta del cronograma, con sus fechas de inicio y finalización, de las actividades a realizar en la investigación. Elaboración propia, realizado con Microsoft Excel.

En la Figura 20 se puede observar la propuesta del cronograma de una forma más grafica de cómo estará conformada la investigación en fases, con los días de trabajo necesarios para la realización de la actividad, así como las fechas de inicio y finalización que tendrá cada una de las actividades.

**Figura 20.**

*Cronograma de actividades de la investigación*



*Nota.* Cronograma de las actividades para la realización de la investigación. Elaboración propia, realizado con Microsoft Excel.





## **13. FACTIBILIDAD DEL ESTUDIO**

En cuanto a la factibilidad técnica y financiera con que cuenta la investigación del intérprete del idioma maya kaqchikel al español se analizará si existen los recursos necesarios para la realización del trabajo, incluyendo recursos humanos, financieros, tecnológicos, acceso a información, equipo, infraestructura, entre otros, y se especificarán algunas fuentes de financiamiento.

### **13.1. Factibilidad técnica**

La factibilidad técnica del estudio del intérprete del idioma maya kaqchikel al español que utiliza aprendizaje automático presenta los siguientes factores que lo respaldan:

- Disponibilidad de datos: los datos lingüísticos del kaqchikel pueden ser más limitados que los de idiomas más estudiados, sin embargo, existen iniciativas como programas de estudios en escuelas y en la Universidad de San Carlos de Guatemala que enseñan el idioma kaqchikel y recursos en línea que podrían proporcionar un corpus de texto o audios lo suficientemente grandes para entrenar modelos de aprendizaje automático. Además, la recolección de nuevos datos a través de entrevistas, grabaciones y textos existentes puede aumentar significativamente la cantidad de datos disponibles.
- Recursos humanos: se puede contar con traductores jurados que pueden ayudar con la traducción de palabras y oraciones para la recolección de

los datos, como también maestros que enseñan el idioma kaqchikel en establecimientos educativos y personas en el interior de la república que hablen el idioma.

- Herramientas y bibliotecas tecnológicas: existen numerosas herramientas y bibliotecas de código abierto para el Procesamiento del Lenguaje Natural (PLN) y el aprendizaje automático, como SpaCy, NLTK, TensorFlow y PyTorch. Estas herramientas facilitan la creación y entrenamiento de modelos de traducción automática neuronal.
- Avance de la tecnología: el campo del aprendizaje automático ha experimentado un rápido desarrollo en los últimos años, con modelos cada vez más sofisticados y capaces de manejar lenguas menos representadas.

### **13.2. Factibilidad financiera**

La factibilidad financiera propuesta para esta investigación del intérprete del idioma kaqchikel al español se describe en la tabla 4, donde se incluyen los costos de los datos, equipos computacionales y de grabación, sueldos del personal trabajador y costos de software y licencias.

**Tabla 4.***Factibilidad financiera del estudio con costos en quetzales*

<b>Categoría</b>	<b>Elementos</b>	<b>Costos estimados</b>	<b>Precios en quetzales</b>
Personal	Científico de datos	Salario mensual x número de meses = Q15,000 x 9	Q135,000
	2 ingenieros de sistemas	Salario mensual x número de meses = Q15,000 x 9	Q270,000
	Traductor jurídico del idioma kaqchikel	Salario mensual x número de meses Q10,000 x 3	Q30,000
	Tester	Salario mensual x número de meses = Q10,000 x 3	Q30,000
	Recolección de los datos	Salario mensual x número de meses = Q10,000 x 3	Q30,000
Hardware	Computadora para el desarrollo	Costos de compra o uso	Q10,000
	Equipo de grabación	Costo de compra	Q2,000
Software	Licencias de Software (ejemplo, Python, TensorFlow, spaCy)	Utilizaciones gratuitas	Q0.00
	Servicios en la nube AWS	Se utilizará de forma gratuita o una reserva de Servicio de uso x mes = Q5,000 x 8	Q40,000
Datos	Corpus bilingüe	Costos de	Q10,000

Continuación de la Tabla 4.

	kaqchikel-español	adquisición o creación	
Otros	Oficina	Alquiler x mes = Q1200 x 9	Q10,800.00
	Internet	Precio x mes = Q250 x 9	Q 2,250.00
	Luz	Precio x mes Q100 x 9	Q 900.00
Total			Q570,950

*Nota.* Propuesta financiera con los costos que tendrá la realización de la investigación. Elaboración propia, realizado con Microsoft Word.

### 13.3. Fuentes de financiamiento

Sin embargo, existen varias formas de reducir los costos utilizando los siguientes factores:

- Utilización de recursos gratuitos: existen muchas plataformas en la nube que ofrecen créditos gratuitos para experimentar con el aprendizaje automático.
- Colaboración con instituciones académicas: las universidades y centros de investigación pueden proporcionar acceso a recursos computacionales y personal especializado. También pueden asignar a los estudiantes que estén realizando su ejercicio profesional supervisado (EPS) o prácticas de campo de carreras afines a esta tecnología para colaborar con el proyecto como parte de su culminación de estudios.

- Financiamiento externo: se pueden solicitar fondos a organizaciones gubernamentales, fundaciones o empresas interesadas en la preservación de las lenguas indígenas.
- Financiamiento propio: tomar en cuenta los recursos y servicios necesarios que pueda brindar la investigadora para la realización del proyecto.



## REFERENCIAS

- Ali, M., Rahman, M., Chaki, J., Dey, N. & Santosh, K. C. (2021). Machine translation using deep learning for universal networking language based on their structure. *International Journal of Machine Learning and Cybernetics*, 12(8), 2365-2376. <https://link.springer.com/article/10.1007/s13042-021-01317-5>
- Fiorini, S., Tezcan, A., Vanallemeersch, T., Szoc, S., Migdisi, K., Meeus, L. & Macken, L. (2023). *Translations and open science-exploring how translation technologies can support multilingualism in scholarly communication*. INCOMA Ltd. <https://acl-bg.org/proceedings/2023/HiT-IT%202023/pdf/2023.hitit2023-1.4.pdf>
- He, D., Xia, Y., Qin, T., Wang, L., Yu, N., Liu, T. & Ma, W. (2016). Dual learning for machine translation. *Advances in neural information processing systems*, 29(4), 5-10. <https://proceedings.neurips.cc/paper/2016/hash/5b69b9cb83065d403869739ae7f0995e-Abstract.html>
- Liu, S., Sun, Y. & Wang, L. (2021). Recent advances in dialogue machine translation. *Information*, 12(11), 484-486. <https://www.mdpi.com/2078-2489/12/11/484>
- Liu, S., Yang, N., Li, M. & Zhou, M. (2014, June). A recursive recurrent neural network for statistical machine translation. In: *Proceedings of the 52nd*



*Annual Meeting of the Association for Computational Linguistics, Vol. 1* (pp. 1491-1500). Long Papers. <https://aclanthology.org/P14-1140.pdf>

Luong, M., Kayser, M. & Manning, C. (2015, July). Deep neural language models for machine translation. In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*, (pp. 305-309). ACL. <https://aclanthology.org/K15-1031.pdf>

Ohidujjaman, F. Faysal, S. & Huda, M. (2021). Automatic machine translation for bangla and english resolving ambiguities. *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*, DHAKA, Bangladesh, 2021, pp. 27-32, [doi: 10.1109/ICREST51555.2021.9331085](https://doi.org/10.1109/ICREST51555.2021.9331085).  
<https://ieeexplore.ieee.org/document/9331085>

Özkaya, E. (2023). The Transformative Role of Artificial Intelligence and Machine Learning in Interpreting and Language Services. *RumeliDE Dil Ve Edebiyat Araştırmaları Dergisi*, 27(36), 1591-1598.  
<https://doi.org/10.29000/rumelide.1372500>

Ramírez, G., Pérez, J., Sánchez, F., Rossi, C., Kenny, D., Superbo, R. & Torres, O. (2021). MultiTraiNMT: training materials to approach neural machine translation from scratch. In *TRITON 2021 (Translation and Interpreting Technology Online)*. ACLBG. <https://aclbg.org/proceedings/2021/TRITON%202021/pdf/2021.triton-1.21.pdf>

Singh, S., Kumar, A., Darbari, H., Singh, L., Rastogi, A. & Jain, S. (2017). Machine translation using deep learning: an overview. In *2017*

*international conference on computer, communications and electronics (comptelix)*, (pp. 162-167). IEEE. [10.1109/COMPTELIX.2017.8003957](https://doi.org/10.1109/COMPTELIX.2017.8003957)

Tapo, A., Coulibaly, B., Diarra, S., Homan, C., Kreutzer, J., Luger, S. & Leventhal, M. (2020). *Neural machine translation for extremely low-resource African languages: a case study on Bambara*. Arxiv. <https://arxiv.org/pdf/2011.05284>

Wen, F. & Yue, Y. (2021). A study of deep learning combined with phonetic models in foreign language teaching. *Wireless Personal Communications*, 119(1), 275-288. <https://link.springer.com/article/10.1007/s11277-021-08207-7>

Yang, Y. & Yue, Y. (2020). English speech sound improvement system based on deep learning from signal processing to semantic recognition. *International Journal of Speech Technology*, 23(3), 505-515. <https://link.springer.com/article/10.1007/s10772-020-09733-8>

Zhang, J. & Zong, C. (2015). Deep neural networks in machine translation: an overview. *IEEE Intell. System*, 30(5), 16-25. <https://nlpr.ia.ac.cn/cip/ZongPublications/2015/IEEE-Zhang-8-5.pdf>