

# Procesamiento del lenguaje natural como eje central de la inteligencia artificial generativa

Rafael  
Muñoz Guillena



LECCIÓN INAUGURAL  
CURSO ACADÉMICO 2024-2025



UNIVERSIDAD  
DE LA RIOJA



Procesamiento del lenguaje  
natural como eje central  
de la inteligencia  
artificial generativa



**Prof. Dr. Rafael Muñoz Guillena**

Presidente de la Sociedad Española de Procesamiento del Lenguaje Natural

Secretario del Centro de Inteligencia Digital

Departamento de Lenguajes y Sistemas Informáticos

Universidad de Alicante

# Procesamiento del lenguaje natural como eje central de la inteligencia artificial generativa

Lección Inaugural del Curso Académico 2024-2025

Universidad de La Rioja  
Servicio de Publicaciones

2024

Nombres: Muñoz Guillena, Rafael, autor.

Título: Procesamiento del lenguaje natural como eje central de la inteligencia artificial generativa :  
Lección Inaugural del Curso Académico 2024-2025 / Rafael Muñoz Guillena.

Descripción: Primera edición. | Logroño : Universidad de La Rioja, Servicio de Publicaciones, 2024.

Temas: Inteligencia artificial. | Proceso en lenguaje natural (Informática).

Clasificación: CDU 004.8 | CDU 81'322.2 | Thema 1.0 UYQ | Thema 1.0 CFX

© 2024

Rafael Muñoz Guillena

Universidad de La Rioja. Servicio de Publicaciones

Depósito Legal: LR 1111-2024

Diseño de colección: Servicio de Relaciones Institucionales y Comunicación de la UR

Imprime: ABZ Impresores

Esta obra se difunde bajo una Licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Unported. Permisos que vayan más allá de lo cubierto por esta licencia pueden solicitarse a los titulares del copyright.



# Índice

<b>Prefacio</b>	<b>11</b>
<b>1. Introducción</b>	<b>13</b>
<b>2. Procesamiento de Lenguaje Natural</b>	<b>17</b>
<b>2.1. Métodos de procesamiento y su aproximación a la inteligencia artificial</b>	<b>22</b>
<b>2.2. Tareas de procesamiento de lenguaje natural</b>	<b>25</b>
<b>2.3. Aplicaciones específicas</b>	<b>28</b>
2.3.1. Traducción automática	28
2.3.2. Recuperación de información	32
2.3.3. Extracción de información	33
2.3.4. Búsqueda de respuestas	34
2.3.5. Generación de resúmenes	35
2.3.6. Clasificación de textos	40
2.3.7. Análisis de sentimientos	42
2.3.8. Reputación social	42
2.3.9. Identificación de haters o trolls	44
2.3.10. Cyberbullying	45
2.3.11. NER (Name Entity Recognition)	45
2.3.12. Categorización de Noticias	47
2.3.13. Fake news/Bulos	48
2.3.14. Similitud textual semántica	51
2.3.15. Identificación de paráfrasis	51
2.3.16. Inferencia del Lenguaje Natural	52
2.3.17. Implicación textual	53
2.3.18. Modelado de temas	54
<b>2.4. Recursos léxicos</b>	<b>57</b>
<b>3. Inteligencia Artificial</b>	<b>61</b>
<b>3.1. Tipos de inteligencia artificial</b>	<b>63</b>
<b>3.2. Inteligencia artificial generativa</b>	<b>66</b>
<b>3.3. Tipos de problemas</b>	<b>70</b>

3.3.1. Clasificación	70
3.3.2. Regresión	71
3.3.3. Reconocimiento de imágenes	71
3.3.4. Procesamiento de Lenguaje Natural (PLN)	72
3.3.5. Series temporales y predicción	73
3.3.6. Recomendación	74
3.3.7. Reconocimiento de voz y audio	74
3.3.8. Juegos y agentes autónomos	75
3.3.9. Modelos generativos	75
3.3.10. Detección de anomalías	76
3.3.11. Optimización y control	76
<b>3.4. Redes neuronales</b>	<b>77</b>
3.4.1. Perceptrón	78
3.4.2. Perceptrón multicapa	81
3.4.3. Redes Backforward	82
3.4.4. Redes neuronales recurrentes	82
3.4.5. Redes neuronales convolucionales	83
3.4.6. Redes LSTM	85
3.4.7. Generative Adversarial Networks (GAN)	86
<b>3.5. Modelos de lenguajes a Gran Escala (LLM)</b>	<b>88</b>
3.5.1. Arquitectura de un LLM	91
3.5.2. Entrenamiento de un modelo de lenguaje	97
<b>3.6. Hugging Face</b>	<b>106</b>
<b>3.7. Retos de la inteligencia artificial</b>	<b>109</b>
<b>4. Empleabilidad en el PLN e IA</b>	<b>136</b>
4.1. Previsiones del mercado de procesamiento del lenguaje natural	139
<b>5. Conclusión</b>	<b>148</b>
<b>6. Referencias</b>	<b>150</b>



# Índice de ilustraciones

Figura 1: La computadora ENIAC en 1946	13
Figura 2: Disciplinas dentro de la inteligencia artificial	17
Figura 3: Tareas de procesamiento del lenguaje	27
Figura 4: Representación del modelo de codificador-decodificador (Inglés-Alemán)	31
Figura 5: Arquitectura genérica de un sistema IR	32
Figura 6: Arquitectura genérica de un sistema IE	34
Figura 7: Arquitectura genérica de un sistema de QA	35
Figura 8: Arquitectura general para la generación de resúmenes	35
Figura 9: Ejemplo de resumen extractivo	36
Figura 10: Arquitectura genérica de un sistema TC. Una clasificación por texto	41
Figura 11: Arquitectura genérica de un sistema TC. Varias clasificaciones por texto	41
Figura 12: Herramienta de sentimientos Social Analytics	43
Figura 13: Trols y Haters	44
Figura 14: Cyberbullying	45
Figura 15: NER en historias clínicas	47
Figura 16: Acciones frente a las fake news	49
Figura 17: Extracción de entidades, eventos y relaciones	56
Figura 18: Neurona natural y artificial (Perceptrón)	61
Figura 19: Asistentes virtuales y fechas de creación	68
Figura 20: Evolución de las redes neuronales	77
Figura 21: Simulación de una red neuronal biológica con una artificial	78
Figura 22: Entrenamiento de un perceptrón para la función OR	79
Figura 23: Perceptrón multicapa	82
Figura 24: Algoritmo para el entrenamiento Backpropagation	84
Figura 25: Arquitectura red CNN	84
Figura 26: Arquitectura red LSTM	85

Figura 27: Arquitectura de una red GAN	87
Figura 28: Árbol evolución modelos de lenguaje	91
Figura 29: Arquitectura TRANSFORMER	92
Figura 30: Modelos de lenguaje hasta marzo 2023	96
Figura 31: Modelos de lenguaje 2024	97
Figura 32: Imagen del sistema COMPAS para predecir la reincidencia de presos	112
Figura 33: Previsiones del mercado PLN a nivel global	140
Figura 34: Previsión sobre la variación en los presupuestos dedicado a PLN	141
Figura 35: Evolución del mercado mundial del PLN	141
Figura 36: Evolución del mercado europeo del PLN	142
Figura 37: Evolución del mercado en España del PLN	143
Figura 38: Agentes del PLN (%) por comunidad autónoma	143
Figura 39: Facturación de las empresas del sector del PLN	145

# Prefacio

La inteligencia artificial generativa y el procesamiento del lenguaje natural (PLN) están revolucionando la manera en que interactuamos con la información y el conocimiento. Esta lección tiene como objetivo explorar estos campos emergentes, presentando una visión integral.

Este trabajo ha sido elaborado con motivo de la lección inaugural del curso 2024/2025 de la Universidad de La Rioja, un evento que marca el inicio de un nuevo ciclo académico lleno de oportunidades para el aprendizaje y la innovación. Para la elaboración de esta lección se ha utilizado material y recursos propios y otros creados en dos cursos de la Universidad de Alicante: el curso de Experto en Procesamiento de Lenguaje Natural y la asignatura de Minería de Texto del Máster en Ciencia de Datos. Estos cursos han proporcionado una base sólida en los principios y técnicas que sustentan el PLN y la IA generativa, permitiéndonos ofrecer un contenido detallado y práctico.

Quiero expresar mi más sincero agradecimiento a mis compañeros del Grupo de Investigación de Procesamiento de Lenguaje Natural y Sistemas Información (GPLSI) de la Universidad de Alicante por la cesión de parte de este material. Su apoyo y participación han sido cruciales para la realización de esta lección. Su colaboración, conocimientos compartidos y entusiasmo han enriquecido significativamente este trabajo,

haciendo posible una obra que, espero, será de valor para estudiantes, investigadores y profesionales interesados en la inteligencia artificial y el procesamiento del lenguaje.

Gracias por acompañarnos en esta exploración del PLN y la IA generativa. Espero que encuentren en estas páginas una fuente de inspiración y conocimiento.

**RAFAEL MUÑOZ GUILLENA**

Julio, 2024

# 1. Introducción

Desde la aparición de los primeros ordenadores, como el ENIAC<sup>1</sup> en 1945, poder interactuar con ellos no era una tarea baladí, ni a nivel de usuario ni mucho menos a nivel de programación. La comunicación con los ordenadores en un principio se hacía mediante la configuración de una serie de cables e interruptores, un proceso conocido como programación manual o de tableros de enchufes. Estos primeros ordenadores, como el Colossus<sup>2</sup> y el UNIVAC<sup>3</sup>, eran máquinas enormes que requerían configuraciones físicas para realizar cálculos básicos.



*Figura 1: La computadora ENIAC en 1946  
(Fuente: Historical [Corbis via Getty Images])*

- .....
1. <https://www.britannica.com/technology/ENIAC>
  2. <https://www.britannica.com/technology/Colossus-computer>
  3. <https://www.britannica.com/technology/UNIVAC>

A medida que la tecnología de los ordenadores avanzaba, surgió la necesidad de una forma más eficiente y menos propensa a errores para programar estos dispositivos. En primer lugar se desarrolló el lenguaje en *código máquina* que se basaba en una secuencia de bits (0 y 1) que hacía las veces de la existencia de un cable enchufado o no. Por ejemplo, 10110000 01100001. La tecnología siguió avanzando y entra en juego el lenguaje *ensamblador*. El lenguaje ensamblador es un lenguaje de programación de bajo nivel que utiliza códigos nemotécnicos para representar las instrucciones del código máquina, permitiendo a los programadores escribir de una manera más legible y comprensible para los humanos. Cada instrucción en lenguaje ensamblador se traduce directamente en una instrucción específica de la CPU, lo que hace que el programa resultante sea extremadamente eficiente en términos de velocidad y uso de memoria.

Por ejemplo, en lugar de escribir la secuencia de bits como 10110000 01100001, un programador puede escribir una instrucción en ensamblador como `MOV AL, 61h`, lo cual es mucho más fácil de entender y manipular.

A partir de 1956, la interacción con los ordenadores comenzó a cambiar con la introducción del primer lenguaje de programación de alto nivel, Fortran (Backus et al. 1956), desarrollado por IBM. Fortran (FORmula TRANslation) permitió a los programadores escribir instrucciones en un formato más comprensible que el ensamblador y mucho más que la manipulación de cables e interruptores.

Por ejemplo, para el cálculo del área de un círculo de radio “*r*” se deben dar las siguientes instrucciones en un programa escrito en lenguaje FORTRAN

```
program circulo
real :: r, area
real, parameter :: pi = 3.14159
! Este programa lee un número real r y muestra
! el área del círculo con radio r.
```

```
write (*,*) 'Escribe el radio r:'  
read (*,*) r  
area = pi*r*r  
write (*,*) 'Area = ', area  
stop  
end
```

lo cual es mucho más fácil de entender y escribir que su equivalente en lenguaje *ensamblador*, pero mucho menos intuitivo si pudiéramos indicarle:

"calcula el área de un círculo de radio r"

Es decir, para decirle a un ordenador lo que quisiéramos que hiciera, las personas debían utilizar y aprender un lenguaje que no era natural para nosotros.

Es en este momento en el que se empezó a soñar, por qué no usar nuestra lengua para interactuar con las máquinas. Podemos decir que el Procesamiento del Lenguaje Natural (PLN) nace con un doble objetivo. Por un lado, reducir la brecha de comunicación entre las máquinas y los humanos y, por otro lado, poder realizar un conjunto de tareas para el uso, tratamiento y extracción de conocimiento de forma automática de la información digitalizada (texto o voz).

La ciencia ficción como género narrativo que especula sobre avances científicos y su impacto en la sociedad parece que siempre propone algo inalcanzable en ese momento, pero esta especulación a menudo, con el paso del tiempo, se convierte en realidad. Cuando el cineasta Stanley Kubrick creó en 1968 la película "2001: Una odisea del espacio", imaginó al super ordenador HAL 9000 capaz de comprender tanto el habla como el lenguaje humano, así como de producir habla, leer los labios de los tripulantes o incluso ser capaz de identificar sentimientos cuando le dice a Bowman "Mire, Dave, veo que esto le molesta mucho". En este y en otros momentos de la película, HAL demuestra ser capaz de analizar e identi-

ficar las emociones de los humanos. De hecho, en sus últimas escenas, se diría que HAL es capaz incluso de sentir dichas emociones, cuando afirma “Tengo miedo, Dave”, ante la perspectiva de la desconexión.

Imaginen un mundo donde las máquinas no solo puedan procesar información, sino que también puedan entender nuestros pensamientos y sentimientos con una precisión sorprendente. Un mundo donde el lenguaje natural ya no es una barrera para la comunicación con las máquinas, y donde las conversaciones con nuestros dispositivos son tan naturales como las conversaciones con nuestros seres queridos.

La revolución del procesamiento del lenguaje natural ha abierto las puertas a este futuro donde los asistentes virtuales pueden ayudarnos a manejar nuestras vidas, los sistemas de recomendación pueden anticipar nuestras necesidades y los chatbots pueden proporcionar atención al cliente 24/7.



## 2. Procesamiento de Lenguaje Natural

Con Procesamiento del Lenguaje Natural (PLN) se hace referencia a todos aquellos aspectos de la inteligencia artificial relacionados con la capacidad de comunicación hombre-máquina mediante una lengua natural. El procesamiento del lenguaje natural es una rama de la inteligencia artificial que se centra en la capacidad de las máquinas para comprender, interpretar y generar lenguaje humano mediante el desarrollo de técnicas de diferentes áreas con una fuerte componente interdisciplinar (ciencias de la computación, lógica, matemáticas, psicología cognitiva, lingüística, etc.).

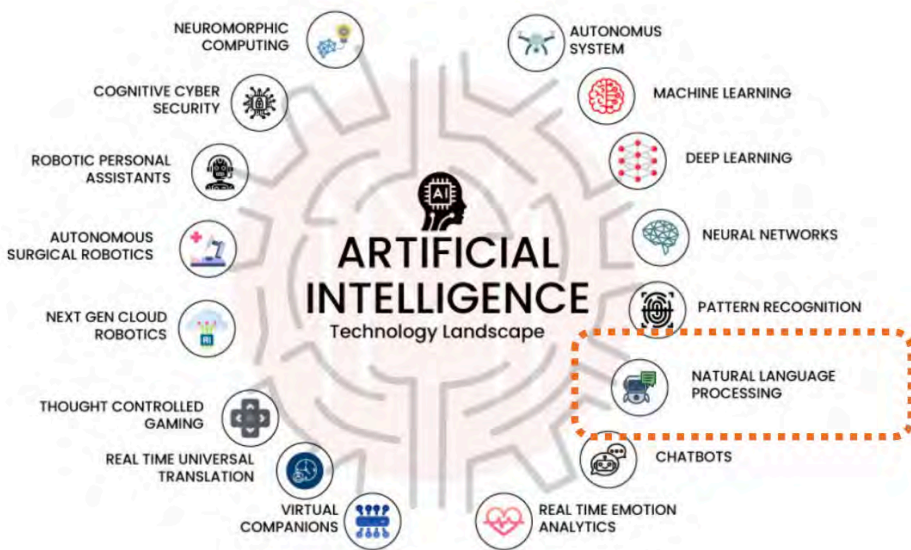


Figura 2: Disciplinas dentro de la inteligencia artificial

(Fuente: Curso de experto de Procesamiento del Lenguaje Natural de la Universidad de Alicante)

En general, un sistema de PLN es una emulación computacional de la capacidad humana para generar e interpretar textos en un idioma concreto (español, inglés, euskera, gallego, catalán, etc.). El sistema trata de generar o interpretar textos tal y como lo haría un ser humano. Esto no quiere decir que los métodos para crear/interpretar textos deban ser los mismos métodos cognitivos utilizados por los humanos.

Las técnicas de PLN pueden imitar o estar inspiradas en lo que conocemos de los modelos lingüísticos y cognitivos humanos, o simplemente utilizar modelos computacionales propios. En cualquier caso, el resultado es un texto y/o su interpretación “como si lo hubiera hecho un humano”; es decir, que un humano no pueda decir si el texto o su interpretación la ha realizado otro humano o una máquina.

Así, un sistema de PLN puede actuar en tres escenarios:

1. Sistemas en los que el texto/voz es la entrada (input) del sistema. En este caso, el proceso computacional es un proceso de interpretación automática. Ejemplo de este escenario son sistemas de detección de opiniones y emociones, sistemas de extracción de información, sistemas text-to-image o text-to-video (generación automática de imágenes o video, respectivamente, a partir de un texto), etc.
2. Sistemas en los que el texto es la salida (output) del sistema. En este caso, el proceso computacional es la generación o creación del texto. Ejemplos de este escenario son los sistemas de descripción de imágenes, la generación de texto a partir de plantillas de información o la creación automática de poesía, entre otros.
3. Sistemas en los que tanto la entrada como la salida serán textos. En estos casos deben darse ambos procesos: la interpretación y la generación (o a la inversa, según la finalidad del sistema). Ejemplos de sistemas texto- texto son los sistemas de traducción automática, resumen automático, sistemas de diálogo (chat hombre-máquina), etc.

Desde la perspectiva lingüística se le suele denominar también al PLN como “**Lingüística computacional**” (LC), que es término más tradicional. En ocasiones se utiliza LC para los aspectos más lingüísticos del PLN (gramáticas computacionales, análisis de rasgos, anotación de corpus, etc.). Otros términos que se utilizan son a menudo “**procesamiento del lenguaje humano**”, “**tecnologías del lenguaje humano**”, “**ingeniería lingüística**” o “**comprensión del lenguaje natural**”, entre otros.

El procesamiento de lenguaje natural aborda el lenguaje desde todos sus niveles:

### **1. Nivel Fonético y Fonológico:**

Se ocupa de los sonidos del habla y su representación. Incluye la conversión de texto a voz (TTS) y de voz a texto (ASR), así como la segmentación de palabras en fonemas (Martin, 2009).

### **2. Nivel Morfológico:**

Se centra en la estructura de las palabras y sus componentes, como raíces, prefijos y sufijos. Las tareas en este nivel incluyen la tokenización, lematización y la derivación, así como el análisis y generación de formas gramaticales, (Haspelmath & Sims, 2013).

### **3. Nivel Léxico:**

Trata con el significado y uso de palabras individuales. Incluye la desambiguación de palabras (determinar el sentido correcto de una palabra en un contexto dado) y el análisis de sinónimos y antónimos (Felbaum, 1998).

### **4. Nivel Sintáctico:**

Se enfoca en la estructura gramatical de las oraciones. Implica el análisis sintáctico (*parsing*) para determinar cómo se combinan las palabras para formar frases y oraciones correctas, (Manning & Schütze, 1999).

## **5. Nivel Semántico:**

Se ocupa del significado de las palabras y las frases. Incluye la comprensión de frases y oraciones, la inferencia semántica y el reconocimiento de entidades nombradas (NER) (Martin, 2009).

## **6. Nivel Pragmático:**

Trata de cómo se usa el lenguaje en contextos específicos. Implica la interpretación del significado en función del contexto situacional, la intención del hablante y las convenciones sociales (Levinson, 1983).

## **7. Nivel Discursivo:**

Se centra en la estructura y coherencia de textos más largos, como párrafos y documentos completos. Incluye la anáfora (resolución de referencias cruzadas) y la cohesión del discurso (Brown & Yule, 1983)

## **8. Nivel de Representación del Conocimiento:**

Involucra la comprensión y utilización del conocimiento del mundo para interpretar y generar lenguaje. Incluye el uso de ontologías y bases de conocimiento para enriquecer la comprensión semántica (Sowa, 1999)

Como vemos, los sistemas de PLN giran en torno al texto/voz, que es la unidad de comunicación lingüística humana. Antes de profundizar en las técnicas de PLN para procesar textos, se va a exponer qué es un texto y cómo es su procesamiento desde un punto de vista lingüístico (teórico-cognitivo). En algunos casos se hará alusión a sus implicaciones computacionales

Según Chomsky (1986), filósofo y lingüista, “el verdadero desafío en el procesamiento del lenguaje natural no es simplemente acumular y analizar datos lingüísticos, sino entender la estructura y el significado subyacente del lenguaje humano”. En definitiva, para que las máquinas sean verdaderamente inteligentes, deben ser capaces de entender el len-

guaje y el contexto humano. El procesamiento del lenguaje natural es un componente crucial en este desafío.

El lenguaje natural es la forma más eficaz de comunicación humana, y es a través del lenguaje que podemos expresar nuestros pensamientos, sentimientos y deseos. Al permitir que las máquinas comprendan y generen lenguaje natural, estamos abriendo la puerta a una nueva era de interacción entre humanos y máquinas.

Además, el PLN permite a las máquinas aprender de manera autónoma, sin necesidad de programación explícita. Mediante el análisis de patrones lingüísticos y semánticos, las máquinas pueden identificar relaciones y conceptos abstractos, y generar nuevo conocimiento a partir de ellos.

## 2.1. Métodos de procesamiento y su aproximación a la inteligencia artificial

Según Navarro (2024), existen dos paradigmas de representación que corresponden a los dos métodos de procesamiento lingüístico principales: el racional (simbólico) y el neuronal (conexionista). A estos dos modelos se les une un tercero, que ha sido el modelo de procesamiento preponderante en los últimos veinte años: el modelo de procesamiento simbólico empírico.

### a) El modelo simbólico racional.

Se basa en la creación explícita de reglas de procesamiento lingüístico. Es el modelo más cercano a la lingüística teórica, influido sobre todo por la teoría generativo-transformacional de N. Chomsky (Chomsky & Bouallal, 1968). Esta teoría formaliza los procesos de interpretación lingüística mediante reglas. Esta misma idea es la que se adaptó al PLN. Este tipo de modelos racionalistas se desarrollaron sobre todo en los años 80 del siglo XX.

Estas reglas son reglas de manipulación simbólica, es decir, las reglas transforman un símbolo en otro. Un token en un texto es un símbolo. Una regla puede transformar ese token en su categoría gramatical (que también es un símbolo). A su vez, otra regla, a partir de la etiqueta categorial, puede transformar el token de entrada en un número que identifique su significado en una ontología, por ejemplo. Este identificador es de nuevo un símbolo que se ha transformado en otro: una palabra en su significado. Aplicado a un texto completo, estas transformaciones de símbolos sobre *Madrid*, sería el proceso de interpretación.

Madrid → Nombre\_propio → Q2807

Un conjunto de reglas forma una gramática, y una gramática completa (en teoría) podría procesar cualquier texto. Hoy día se sabe que es imposible representar todo el conocimiento lingüístico con una sola gramática, pero durante muchos años así se creía. Se pensaba que con más reglas se podría llegar a interpretar cualquier texto. Hoy día las gramáticas de reglas explícitas se utilizan tanto para formalizar el conocimiento lingüístico de un idioma como para dar cuenta de casos concretos y específicos. La principal ventaja de estos métodos es que son claros y transparentes. El sistema realiza el análisis tal y como haya sido diseñado por el humano, que es quien crea las reglas. Si bien desde un punto de vista lingüístico esta capacidad explicativa es una ventaja, desde un punto de vista computacional no es la aproximación más eficiente.

### **b) El modelo neuronal (conexionista).**

Es el modelo de representación propio de las redes neuronales artificiales (RNA) y hoy, gracias a los grandes modelos de lenguaje (LLM, *Large Language Models*), es el paradigma de representación predominante. Este paradigma solo tiene sentido y se puede comprender dentro de una RNA. Una RNA, aparte de una entrada y una salida, está formada por unidades de procesamiento conectadas entre sí llamadas neuronas. La falta de capacidad computacional hizo que apenas pudiera usarse en tareas complejas de PLN hasta principios del 2010. En este modelo nos basaremos para esta lección y por tanto lo abordaremos más adelante.

### **c) El modelo empírico.**

Una solución a estos problemas computacionales vino con los métodos empíricos. Además, es mediante este tipo de modelo cuando empieza a abordarse el PLN desde otras perspectivas no solo la lingüística, como son las matemáticas a través de los métodos estadísticos.

Estos métodos también utilizan símbolos y reglas, pero esas reglas:

1. son deducidas directamente de los textos mediante técnicas de aprendizaje automático.
2. incorporan pesos estadísticos de tal manera que se pueden aplicar o no según el contexto.

Hay dos tipos de métodos en aprendizaje automático: los métodos supervisados y los no supervisados, según partan o no de información previa.

En PLN, los métodos supervisados parten de corpus que han sido ya previamente analizados por expertos. Estos han marcado el corpus con un lenguaje formal (XML, JSON...) y han incluido la etiqueta (el símbolo) a cada fenómeno lingüístico que se quiere modelar. En el caso, por ejemplo, de crear un sistema de PLN que analice las categorías gramaticales de un texto, se parte de un conjunto de textos cuyas palabras tienen la etiqueta categorial correspondiente, que ha sido marcada por expertos. A partir de ahí, los algoritmos de aprendizaje aprenden de cada palabra cuál es su categoría según el contexto donde aparece, y (si está bien hecho) harán el proceso de abstracción de tal manera que ante una palabra desconocida, pueda decidir por su contexto cuál es su categoría gramatical de manera correcta.

Los procesos no supervisados realizan procesos de inferencia sobre los corpus, de tal manera que extraen la información sin necesidad de que haya sido previamente marcada por expertos. Normalmente hacen agrupaciones.

Ahora es el momento de reflexionar sobre cómo podemos aprovechar el poder del procesamiento del lenguaje natural para mejorar nuestras vidas, y no solo para hacérsela más fácil, democratizando la comunicación hombre-máquina.



## 2.2. Tareas de procesamiento de lenguaje natural

En el procesamiento de lenguaje natural podemos identificar tres tareas principales:

### 1) Procesamiento del lenguaje natural

El Procesamiento del Lenguaje Natural (PLN) es un campo de la inteligencia artificial y la lingüística computacional que se enfoca en la interacción entre las computadoras y el lenguaje humano. El objetivo principal del PLN es permitir que las computadoras comprendan, interpreten y respondan al lenguaje humano de manera efectiva. Abarca todas las tareas relacionadas con la interacción de computadoras con el lenguaje humano, incluyendo tanto el análisis como la generación de texto.

Las tareas comunes en el PLN incluyen, entre otras:

- Análisis de texto: Tokenización, lematización, etiquetado de partes del discurso (POS tagging).
- Reconocimiento de entidades nombradas (NER): Identificación de nombres de personas, lugares, organizaciones, etc.
- Traducción automática: Convertir texto de un idioma a otro.
- Análisis de sentimiento: Determinar las emociones expresadas en un texto.
- Clasificación de texto: Asignar categorías a textos.

### 2) Entendimiento del lenguaje natural

El Entendimiento del Lenguaje Natural (NLU) es una subdisciplina del PLN que se centra específicamente en la capacidad de las máquinas para

comprender el significado y la intención detrás del lenguaje humano. NLU va más allá del procesamiento superficial del texto para interpretar el contexto, el propósito y las inferencias que los humanos hacen al comunicarse. Se enfoca en la comprensión profunda del significado y la intención detrás del lenguaje humano.

Las tareas típicas de NLU incluyen, entre otras:

- **Análisis semántico:** Comprensión del significado de las palabras y frases en contexto.
- **Desambiguación de palabras:** Determinar el sentido correcto de una palabra que tiene múltiples significados.
- **Inferencia lógica:** Extraer conclusiones a partir de información implícita en el texto.
- **Respuesta a preguntas:** Proporcionar respuestas precisas a preguntas formuladas en lenguaje natural.

### **3) Generación de lenguaje natural**

La Generación de Lenguaje Natural (GLN) es otra subdisciplina del PLN que se ocupa de la creación de texto en lenguaje humano a partir de datos estructurados o no estructurados. El objetivo de la GLN es producir texto que sea coherente, fluido y adecuado para el contexto y la audiencia. Se especializa en la creación de texto en lenguaje humano que sea coherente y significativo a partir de datos.

Las aplicaciones de GLN incluyen:

- **Generación de informes:** Crear resúmenes y análisis de datos financieros, médicos, etc.
- **Sistemas de diálogo:** Generar respuestas en sistemas de chatbots y asistentes virtuales.

- Producción de contenido: Redactar artículos, descripciones de productos y otros textos de manera automática.

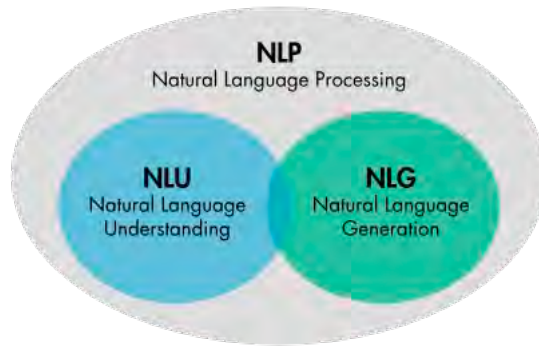


Figura 3: Tareas de procesamiento del lenguaje<sup>4</sup>

.....  
4. <https://es.mathworks.com/discovery/natural-language-processing.html>

## 2.3. Aplicaciones específicas

El lenguaje natural es muy complejo y deben abordarse lo que podríamos definir como tareas intermedias y aplicaciones específicas. Denominamos tareas intermedias a aquellas que son básicas desde el punto de vista lingüístico para aportar información, conocimiento a los sistemas para abordar una tarea final. Entre estas tareas intermedias resolveremos, entre otras, el análisis morfológico, sintáctico, la desambiguación del sentido de las palabras, la resolución de las correferencias y el análisis de semántico.

Dentro de las aplicaciones generales del PLN existen una gran diversidad de variantes. A continuación, se mencionan algunas de ellas.

### 2.3.1. Traducción automática

La traducción automática (MT) es la aplicación por excelencia del PLN. El primer desarrollo se conoce como «*El experimento de Georgetown*», en 1954, consistió en la traducción automática de más de sesenta oraciones rusas al inglés en un trabajo conjunto entre IBM y la Universidad de Georgetown. Implica el uso de técnicas matemáticas y algorítmicas para traducir documentos de un idioma a otro. Realizar una traducción eficaz es intrínsecamente complejo incluso para los humanos, y requiere competencia en áreas como morfología, sintaxis y semántica, así como una comprensión y un discernimiento expertos de las sensibilidades culturales, tanto para los idiomas (y sociedades asociadas) en consideración (Jurafsky & Martin, 2000).

Existen tres grandes grupos de propuestas:

- Basados en Reglas
- Basados en Datos
- Redes neuronales

### 2.3.1.1. Traducción automática basada en reglas

Consiste en realizar transformaciones a partir del original, reemplazando las palabras por su equivalente más apropiado. Al conjunto de este tipo de transformaciones del texto original se le llama predicción de textos.

Tipos de traductores basados en reglas:

- Directos (traducción automática basada en diccionarios) los cuales asignan la entrada a la salida con reglas básicas.
- De transferencia (traducción automática basada en transferencia) emplean análisis morfológico y sintáctico.
- Interlingüísticos (Interlingua) utilizan un significado abstracto.

Como mínimo, para obtener una traducción de un idioma A hacia un idioma B, se necesita:

- Un diccionario que asignará a cada palabra en A una palabra de B adecuada
- Reglas que representan la estructura regular de las oraciones en A.
- Reglas que representan la estructura regular de las oraciones en B.

### 2.3.1.2. Traducción automática basada en datos

Distinguimos dos tipos:

#### **a) Traducción automática basada en corpus**

La traducción automática a partir de un corpus lingüístico se basa en el análisis de muestras reales con sus respectivas traducciones. Entre los mecanismos que utilizan corpus se incluyen los métodos estadísticos y los basados en ejemplos.

- **Estadística:** El objetivo de la traducción automática estadística es generar traducciones a partir de métodos estadísticos basados en corpus de textos bilingües, como por ejemplo las actas del parlamento europeo, que se encuentran traducidas en todos los idiomas oficiales de la UE. A medida que se generan y se analizan corpus de textos multilingües, se mejoran iterativamente los resultados al traducir textos de ámbitos similares.
- **Basada en ejemplos:** La traducción automática basada en ejemplos se caracteriza por el uso de un corpus bilingüe como principal fuente de conocimiento en tiempo real. Es esencialmente una traducción por analogía (relación de semejanza) y puede ser interpretada como una implementación del razonamiento a partir de casos base empleado en el aprendizaje automático, que consiste en la resolución de un problema basándose en la solución de problemas similares.

### **b) Traducción automática basada en el contexto**

La traducción automática basada en el contexto utiliza técnicas basadas en hallar la mejor traducción para una palabra fijándose en el resto de palabras que la rodean.

Básicamente este método se basa en tratar el texto en unidades de entre 4 y 8 palabras, de manera que se traduce cada una de ellas por su traducción al idioma de destino, y se eliminan las traducciones que han generado una “frase” sin sentido. Luego, se mueve la ventana una posición (palabra), retraduciendo la mayoría de ellas de nuevo y volviendo a filtrar dejando solo las frases coherentes. Se repite dicho paso para todo el texto. Y luego se pasa a concatenar los resultados de dichas ventanas de manera que se logre una única traducción del texto.

El filtrado que se realiza donde se decide si es una frase con sentido utiliza un corpus del lenguaje destino, donde se cuentan el número de

apariciones de la frase buscada. Se trata, por tanto, de un método basado en ideas bastante simples que ofrecen muy buenos resultados en comparación con otros métodos. Como ventajas, aporta también la facilidad de añadir nuevas lenguas, ya que solo se necesita:

- un buen diccionario, que puede ser cualquier versión comercial adaptada mediante reglas gramaticales para tener los verbos conjugados y los nombres/adjetivos con sus variaciones en número y género, y
- un corpus en el lenguaje destino, que se puede sacar por ejemplo de Internet, sin que sea necesario traducir ninguna parte, como en los métodos estadísticos.

### 2.3.1.3. Redes neuronales. Codificador-Decoder

Entre las técnicas más actuales utilizadas en este ámbito está el uso de Redes Neuronales profundas y como parte de esto el uso de codificadores (encoders) y decodificadores (decoders). Una red neuronal (el codificador) procesa una oración de la fuente original para otra red (el decodificador). La segunda red neuronal predice las palabras en el lenguaje del destino.

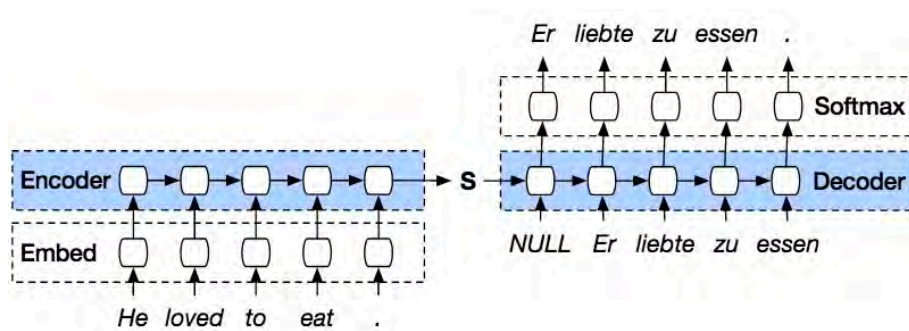


Figura 4: Representación del modelo de codificador-decodificador (Inglés-Alemán)

Fuente: <https://github.com/shangeth/Seq2Seq-Machine-Translation>

## 2.3.2. Recuperación de información

Después de la traducción automática, fue de las primeras aplicaciones necesarias con el boom de Internet en los 80 así como del crecimiento de la información digital. El propósito de los sistemas de recuperación de información (IR) es ayudar a las personas a encontrar la información correcta (la más útil) en el formato correcto (más conveniente) en el momento adecuado (cuando la necesitan) dentro de una colección enorme de documentos.

Entre muchos problemas de IR, un problema principal que debe abordarse es el de clasificar los documentos con respecto a una cadena de consulta, en términos de relevancia para tareas de recuperación *ad hoc* (hecha a medida), similar a lo que sucede en un motor de búsqueda.

Las búsquedas se pueden basar en el texto completo o en otra indexación basada en contenido. La *recuperación de información es la ciencia de buscar información en un documento*, buscar los propios documentos y también buscar los metadatos que describen los datos y las bases de datos de textos, imágenes o sonidos.

Un sistema de IR es un sistema de software que proporciona acceso a libros, revistas y otros documentos; almacena y gestiona esos documentos. El ejemplo de sistema de IR más usado es Google.

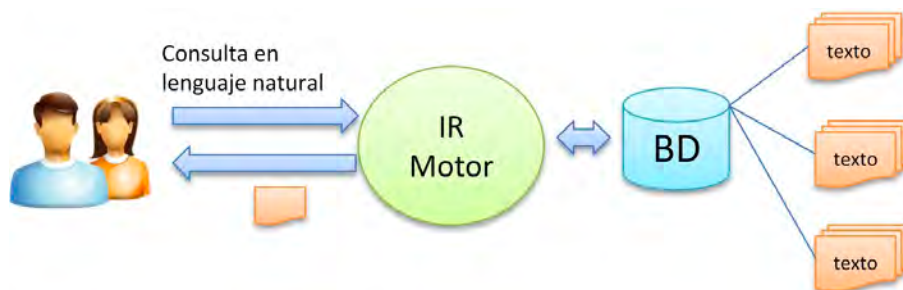


Figura 5: Arquitectura genérica de un sistema IR.



### 2.3.3. Extracción de información

La extracción de información identifica información explícita o implícita del texto. Los resultados de los sistemas varían, pero a menudo los datos extraídos y las relaciones entre ellos se guardan en bases de datos relacionales. No obstante, una etapa avanzada es poder almacenar esta información en estructuras de datos semánticas. La información comúnmente extraída incluye entidades nombradas y relaciones, eventos y sus implicaciones, información temporal y tuplas de hechos.

Este tipo de sistemas fueron muy usuales en la década de los 90 hasta el 2000 existiendo un conferencia denominada MUC (Message Understanding Conference) que era financiada por DARPA (Defense Advanced Research Projects Agency) cuyo objetivo era evaluar los sistemas de extracción de información incorporando en cada edición tareas diferentes como las que a continuación se enumeran y otras como por ejemplo el relleno de plantillas.

Las actividades más comunes implicadas en la extracción de información son:

- Reconocimiento de entidad nombrada (Named Entity Recognition)
- Extracción de eventos (Event Extraction)
- Extracción de relaciones (Relationship Extraction)

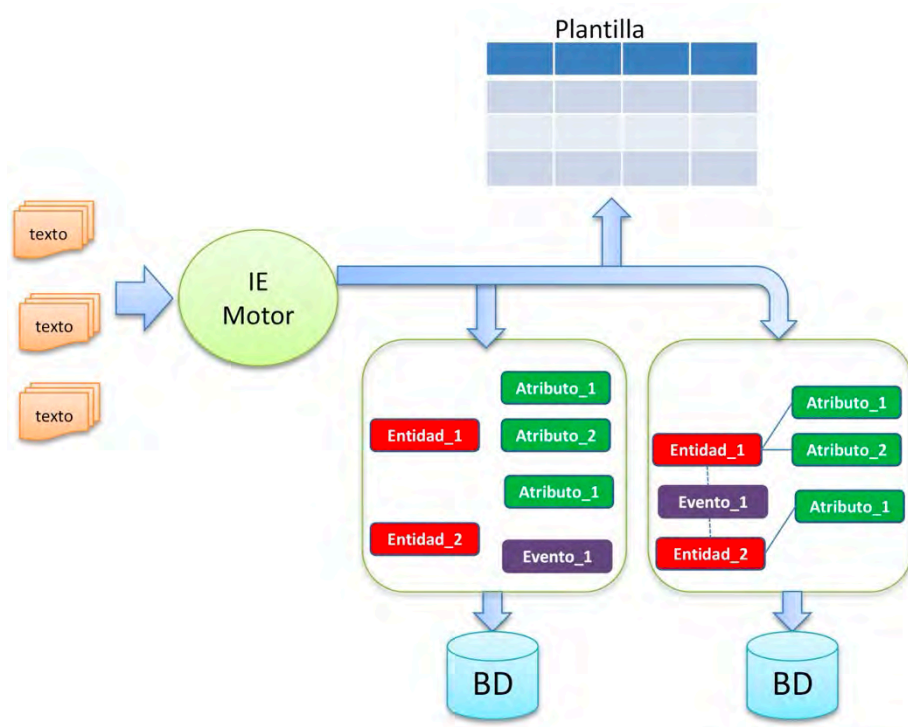


Figura 6: Arquitectura genérica de un sistema IE.

### 2.3.4. Búsqueda de respuestas

Es un tipo de recuperación de la información capaz de recuperar respuestas a preguntas planteadas en lengua natural. Se encarga de buscar una información en una cantidad más o menos grande de documentos, ya que esto debe extraer de dichos documentos un fragmento de texto que responda a una pregunta dada en lenguaje natural.

Estos sistemas están muy ligados a los buscadores web. Debe considerar listas, definiciones, y preguntas del tipo cómo, cuándo, dónde, por qué, etc.

Similar al resumen y la extracción de información, la respuesta a preguntas (QA) recopila palabras, frases u oraciones relevantes de un documento. Los sistemas de QA devuelven esta información de forma coherente en respuesta a una solicitud. Los métodos actuales se parecen a los de resumen.

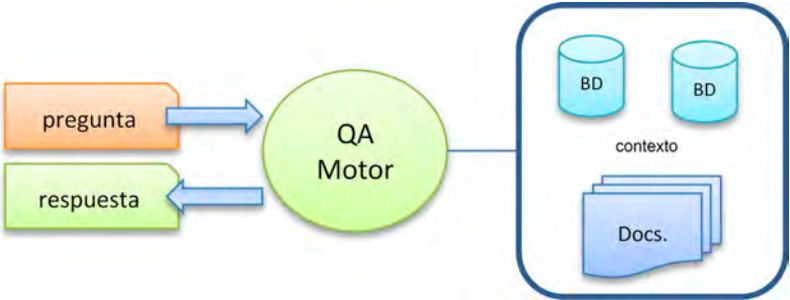


Figura 7: Arquitectura genérica de un sistema de QA.

### 2.3.5. Generación de resúmenes

La generación de resúmenes es la tarea de crear una versión breve de un documento o un conjunto de documentos que retenga las ideas y la información más importantes. Esta técnica es útil en diversas aplicaciones, como la visualización de grandes cantidades de información, la mejora de la eficiencia en la búsqueda de información, y la asistencia en la toma de decisiones. Encontramos diferentes tipos:

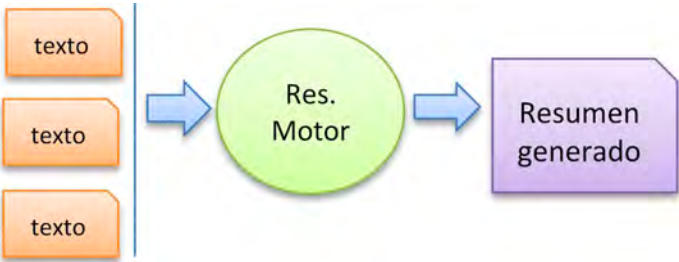
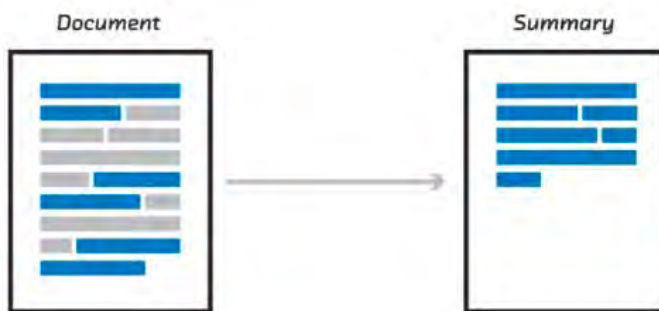


Figura 8: Arquitectura general para la generación de resúmenes

1. Resúmenes Extractivos: Consisten en seleccionar y combinar oraciones o fragmentos específicos del texto original para crear el resumen. Los métodos más usados para la generación son ranking de oraciones, frecuencia de términos en el documento y seleccionar las oraciones que contienen esos términos y algoritmos de grafos (por ejemplo, TextRank). Son más fáciles de implementar y generalmente garantizan que el contenido del resumen sea coherente con el texto original. Pueden no ser tan fluidos y coherentes como los resúmenes abstractivos, ya que están limitados a las oraciones originales.

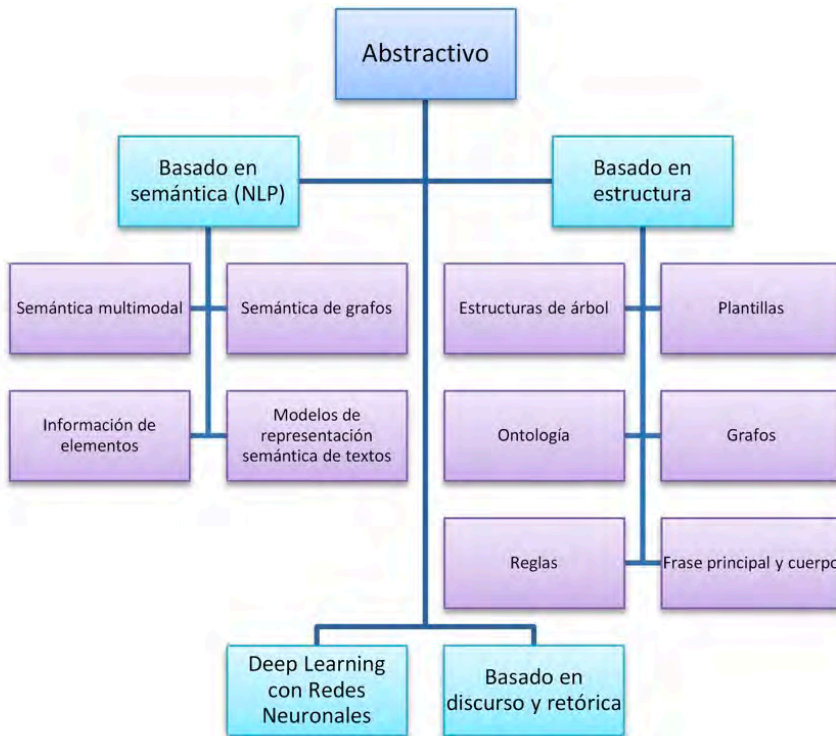


*Figura 9: Ejemplo de resumen extractivo.*

*Fuente: <https://medium.com/@ondenyei.eric/extractive-text-summarization-techniques-with-sumy-3d3b127a0a32>*

2. Resúmenes Abstractivos: Implican generar nuevas oraciones que capturan la esencia del texto original, utilizando un lenguaje propio en lugar de copiar directamente del texto original. Los métodos más usados para la generación son Modelos de Secuencia a Secuencia (Seq2Seq) y mecanismos de atención. Utilizan redes neuronales recurrentes (RNNs) o transformadores (Transformers) para generar resúmenes basados en el contexto del documento. Produce resúmenes más coherentes y naturales, ya que generan texto nuevo. Son más complejos de entrenar y pueden cometer errores semánticos, especialmente si el modelo no está bien ajustado.

Existen varios tipos de resúmenes abstractivos tal y como muestra la siguiente figura<sup>5</sup>



### 2.3.5.1. Lectura fácil

La lectura fácil podríamos considerarla un tipo especial de simplificación de textos con el objetivo de hacerlos más accesibles y comprensibles para una audiencia más amplia. Estos mecanismos son especialmente útiles para personas con dificultades de lectura, como aquellos con dislexia, discapacidades intelectuales, o que están aprendiendo un nuevo idioma. La idea es transformar textos complejos en versiones más sencillas sin perder la esencia del contenido original. Los objetivos de la lectura fácil son:

.....  
 5. <https://doi.org/10.1016/j.eswa.2018.12.011>

1. **Accesibilidad:** Facilitar la comprensión del texto para personas con diferentes niveles de habilidad lingüística y cognitiva.
2. **Inclusión:** Asegurar que la información esté disponible para todos, independientemente de sus habilidades de lectura.
3. **Educación:** Ayudar a las personas en proceso de aprendizaje a entender textos que de otra manera serían demasiado complejos.

Las técnicas más usadas son:

### **1. Simplificación Léxica:**

- **Sustitución de Palabras Difíciles:** Reemplazar palabras complejas o raras por sinónimos más comunes y fáciles de entender.
- **Uso de Listas de Vocabulario Controlado:** Limitar el vocabulario a un conjunto de palabras previamente definido que sea fácil de comprender.

### **2. Simplificación Sintáctica:**

- **Reestructuración de Oraciones:** Dividir oraciones largas y complejas en oraciones más cortas y sencillas.
- **Eliminación de Estructuras Complejas:** Simplificar o eliminar cláusulas subordinadas, frases parentéticas, y construcciones pasivas.

### **3. Generación de Textos Explicativos:**

- **Añadir Ejemplos y Explicaciones:** Incluir ejemplos, definiciones y explicaciones para clarificar conceptos difíciles.
- **Uso de Listas y Puntos:** Presentar información en formato de lista o puntos para mejorar la legibilidad.

#### **4. Resúmenes:**

- **Generación de Resúmenes Simples:** Crear resúmenes que capturen los puntos clave del texto en un lenguaje más sencillo.
- **Preguntas y Respuestas:** Proveer secciones de preguntas frecuentes (FAQ) que aborden los puntos principales del texto.

#### **2.3.5.2. Lectura facilitada**

La lectura facilitada como la lectura fácil también busca hacer los textos más accesibles, pero se centra más en proporcionar herramientas y estrategias adicionales para ayudar a las personas a comprender el texto. Esto puede incluir:

1. **Tecnología de Asistencia:** Uso de software o dispositivos que ayudan en la lectura y comprensión, como lectores de pantalla, aplicaciones de conversión de texto a voz, y herramientas de resaltado de texto.
2. **Apoyo Humano:** Intervenciones directas como tutores, maestros o cuidadores que ayudan a explicar y discutir el contenido del texto.
3. **Material Complementario:** Suministro de recursos adicionales como glosarios, resúmenes, preguntas de comprensión y materiales didácticos.
4. Aunque la lectura fácil y la lectura facilitada comparten el objetivo de mejorar la accesibilidad de los textos, difieren en su enfoque y métodos. La lectura fácil simplifica el texto directamente, mientras que la lectura facilitada proporciona apoyo adicional para ayudar a las personas a comprender textos que de otra manera serían difíciles para ellos. Ambos métodos son complementarios y pueden ser usados juntos para maximizar la accesibilidad y comprensión del contenido escrito.

Aunque la lectura fácil y la lectura facilitada comparten el objetivo de mejorar la accesibilidad de los textos, difieren en su enfoque y métodos. La lectura fácil simplifica el texto directamente, mientras que la lectura facilitada proporciona apoyo adicional para ayudar a las personas a comprender textos que de otra manera serían difíciles para ellos.

Ambos métodos son complementarios y pueden ser usados juntos para maximizar la accesibilidad y comprensión del contenido escrito.

Es común en contextos educativos y terapéuticos, donde el objetivo es ayudar a individuos con dificultades específicas de lectura o aprendizaje a acceder a contenido complejo con apoyo adicional.

Un ejemplo que puede mostrar las diferencias es el siguiente. Supongamos que tenemos un documento complejo sobre información acerca de las obligaciones tributarias:

- **Lectura Fácil:** Se reescribiría el documento usando palabras simples y oraciones cortas, incluiría imágenes para ilustrar conceptos, y sería revisado por personas con dificultades cognitivas para asegurar su comprensibilidad.
- **Lectura Facilitada:** El documento original podría permanecer igual, pero se ofrecería junto con un glosario de términos difíciles, un lector de pantalla que lee el texto en voz alta, y la opción de tener una sesión con un tutor que explique el contenido.

## 2.3.6. Clasificación de textos

Esta área trata la clasificación de texto o la asignación de documentos de texto libre a clases predefinidas. La clasificación de documentos tiene numerosas aplicaciones.

Aunque es una tarea que se viene trabajando desde hace mucho tiempo con algoritmos clásicos de aprendizaje automático, Kim (2014) fue el



primero en utilizar vectores de palabras previamente entrenados en unas Redes Neuronales Convolucionales (CNN) para la clasificación a nivel de oración. El trabajo de Kim fue motivador, y demostró que las CNN simples, con una capa convolucional seguida de una capa densa con desactivación de neuronas y salida de softmax, podían lograr excelentes resultados. Los modelos de CNN propuestos pudieron mejorar el estado del arte en 4 de las 7 tareas diferentes presentadas como clasificación de oraciones, incluido el análisis de sentimientos y la clasificación de preguntas. Conneau *et. al.* (2017) mostró más tarde que las redes que emplean un gran número de capas convolucionales funcionan bien para la clasificación de documentos.

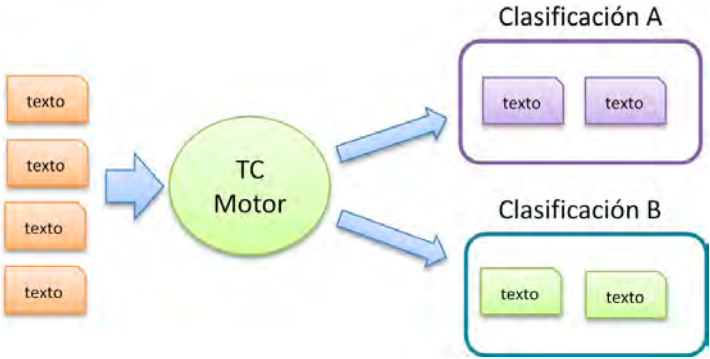


Figura 10: Arquitectura genérica de un sistema TC. Una clasificación por texto

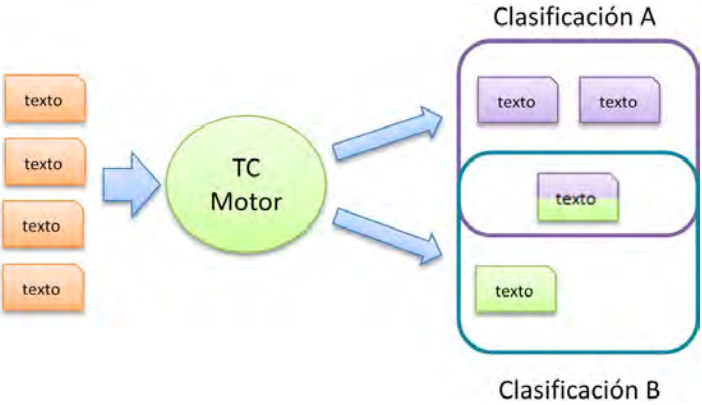


Figura 11: Arquitectura genérica de un sistema TC. Varias clasificaciones por texto.

## 2.3.7. Análisis de sentimientos

El análisis de sentimientos (SA) (o minería de opiniones/emociones) se refiere al uso del procesamiento del lenguaje natural para identificar, extraer, cuantificar y estudiar sistemáticamente estados afectivos e información subjetiva. Se suele aplicar contenido proveniente de redes sociales y en línea, y contenidos de cualquier otra índole que ofrezcan un criterio formulado. Sus aplicaciones de dominio desde el marketing hasta el servicio al cliente y la medicina clínica.

Entre los mayores promotores del estudio del análisis de sentimientos para idiomas iberoamericanos podemos encontrar el TASS<sup>6</sup> (Taller de Análisis Semántico de la Sociedad Española de Procesamiento del Lenguaje Natural).

## 2.3.8. Reputación social

La reputación de una entidad social (una persona, un grupo social, una organización, marca o un lugar) es una opinión sobre esa entidad, típicamente como resultado de una evaluación social sobre un conjunto de criterios, como el comportamiento o el desempeño. Cuando a esto le añadimos el componente social digital nos trasladamos entonces a crear mecanismos que cuantifiquen dicha opinión. Existen múltiples estrategias para cuantificar la opinión social, una de ellas es considerar elementos como,

### **Subjetivos:**

- criterios positivos/negativos/neutrales
- intensidad de los criterios
- manifestación de emociones
- otros

.....  
6. <http://tass.sepln.org/>

## Objetivos:

- likes
- visualizaciones
- contenidos compartidos
- audiencia
- otros

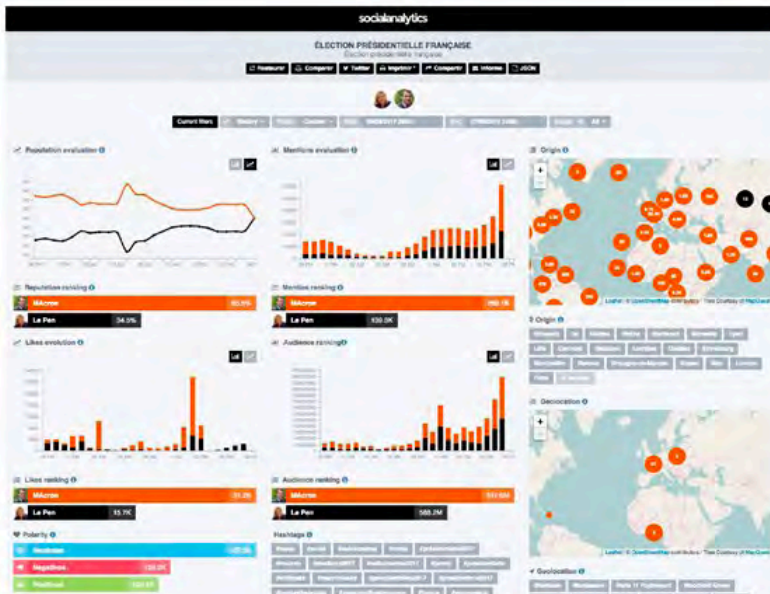


Figura 12: Herramienta de sentimientos Social Analytics

Con la combinación de estos criterios podemos formular una valoración de la reputación que nos permite establecer criterios de comparación y ordinalidad.

La Figura 12: Herramienta de sentimientos Social Analytics muestra un ejemplo de la herramienta Social Analytics<sup>7</sup>. Esta herramienta desarrollada en el Grupo de Procesamiento de Lenguaje Natural de la Universidad de Alicante monitoriza los comentarios o post de diversas fuentes de in-

7. <https://socialanalytics.gplsi.es>

formación como RRSS, blogs, webs de periódicos, y establecer un un mapa de sentimiento teniendo en cuenta el origen de la persona que genera el comentario, el tipo de sentimiento presente en el comentario, las palabras claves, los seguidores y haters más activos en los comentarios.

La herramienta establece un *dashboard* con toda la información o permite monitorizar en todo momento la reputación social de una persona, institución o producto.

### 2.3.9. Identificación de haters o trolls

El contenido (lenguaje) de odio en línea, o discurso de odio, se caracteriza por algunos aspectos clave (como la viralidad o el presunto anonimato) que lo distinguen de la comunicación fuera de línea y lo hacen potencialmente más peligroso y dañino. Por tanto, su identificación se convierte en una misión crucial en muchos campos.

Mediante el uso de técnicas de SA podemos modelar lenguajes con connotación negativa (dañina) para su seguimiento en el tiempo.

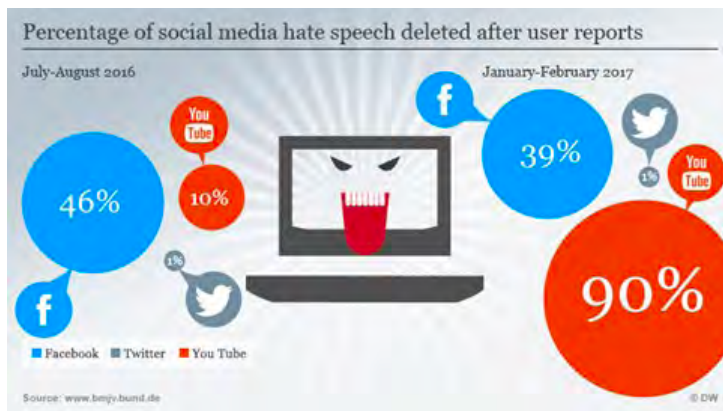


Figura 13: Trolls y Haters.

Fuente: <https://www.dw.com/en/german-anti-hate-speech-group-counters-facebook-trolls/a-38358671>

## 2.3.10. Cyberbullying

Hay evidencias que apuntan a que el cyberbullying es un fenómeno que cada vez más está presente entre los actuales problemas de nuestros jóvenes. El bullying siempre ha existido, pero con la aparición de las redes sociales y medios de comunicación instantáneos su efecto se hace global y con mayor impacto (Van Hee et al. 2018).

El cyberbullying consiste en enviar mensajes instantáneos o en un chat para herir a una persona, publicar fotos o vídeos vergonzosos en las redes sociales y crear rumores en línea.

¿Como el PLN puede intervenir en este problema? Pues creando tecnologías de detección precoz de mensajes de acoso, agresivos, vergonzosos, diseñar algoritmos para identificar los distintos roles (acosado, acosador, etc.) y derivar dicha información a las autoridades competentes.

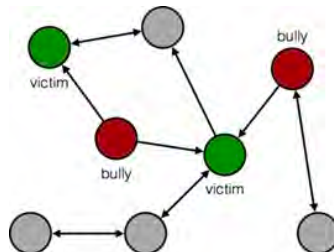


Figura 14: Cyberbullying.

Fuente: <https://dac.cs.vt.edu/research-project/semi-supervised-learning-cyberbullying-harassment-patterns-social-media/>

## 2.3.11. NER (Name Entity Recognition)




Named Entity Recognition, (NER) es la tarea o aplicación de NLP que ayuda a entender el qué, quién y dónde de una serie de documentos<sup>8</sup>.

.....  
8. <https://www.iic.uam.es/inteligencia/la-importancia-tener-ner/>








Nació como sustituto de la extracción de información relevante de los textos, se pensó:

- “si al menos tengo qué se ha hecho, quién lo ha hecho y dónde, tendré información relevante de un documento”, y así es.






Objetivo del NER. Es identificar personas, organizaciones y localizaciones. 3 etiquetas básicas:

- Personas 
- Organizaciones 
- Localizaciones 

También podemos encontrar NER de 7 etiquetas:

- Personas 
- Organizaciones 
- Localizaciones 
- Tiempo 
- Moneda 
- Calles 
- Colectivos 

Luego podemos tener NER para dominios específicos como por ejemplo para dominio clínico:

- Fármacos 
- Sustancias 
- Enfermedades 
- Diagnósticos 
- Procedimientos 
- etc.

Ejemplo: Tomado de la competición “CodiEsp: Clinical Case Coding in Spanish”<sup>9</sup> en la que se aplica NER para codificar terminologías en documentos clínicos.



Figura 15: NER en historias clínicas

## 2.3.12. Categorización de Noticias

El contenido de prensa suele utilizar diferentes estilos de redacción y presentación del contenido para captar una mayor atención de la audiencia. Entre los distintos criterios que pudiéramos necesitar caracterizar.

Las noticias de prensa podemos clasificarlas en:

- Propaganda
- Informativa
- De opinión

.....  
9. <https://temu.bsc.es/codiesp/>

Por su temática:

- Deporte
- Política
- Sociedad
- Economía
- Educación
- Otros

Por su ámbito en:

- Local
- Regional
- Nacional
- Internacional

### 2.3.13. Fake news/Bulos

Con respecto a valorar la calidad o fiabilidad de las noticias surge un nuevo, pero ya antiguo en su origen, concepto que busca corroborar la veracidad de los contenidos informativos o de opinión, fake news (FN).

Dentro de las fake news existen varios problemas que se pueden tratar con PLN:

- Sátira o parodia
- Falsa Conexión (no relación entre el titular y el contenido)
- Contenido Engañoso
- Contenido Falso
- Contenido Impostor
- Contenido Manipulado
- Contenido Fabricado

Acciones recomendadas para identificar noticias falsas:





Figura 16: Acciones frente a las fake news

Fuente: <https://www.ccbiblio.es/infografia-ifla-detectar-noticias-falsas/>

Estas acciones se pueden automatizar en función de la creación de tecnologías de PLN para el combate de las FN.

Técnicas para identificar noticias falsas:

- **Fact Checking (Verificación de hechos).** Consiste en la evaluación de la veracidad y precisión de la información publicada.
- **Stance detection (Detección o determinación de posturas).** Es un subtipo del análisis de opinión y sentimiento que se enfoca en determinar la actitud o postura que un autor tiene respecto a una determinada declaración, tema, o entidad en un texto. A diferencia del análisis de sentimientos, que se concentra en determinar si un texto tiene una valoración positiva, negativa o neutral, la detección de postura busca identificar la posición específica que se adopta, ya sea a favor, en contra, o neutral respecto a una cuestión particular.

- **Sentiment Polarity (Análisis de Sentimientos).** Como hemos visto en el punto anterior.
- **Deception detection (Detección de engaños).** es el proceso de identificar cuándo una persona está mintiendo o intentando engañar. Este proceso utiliza diversas técnicas y herramientas para analizar y evaluar señales verbales y no verbales que pueden indicar que alguien no está siendo sincero. La detección de engaños es una disciplina que abarca campos como la psicología, la lingüística, y la inteligencia artificial, especialmente en el análisis de texto y el reconocimiento de patrones.
- **Clickbait detection (Atracción de la atención y animar a dar click).** Se refiere a la identificación de títulos o titulares diseñados para atraer la atención y provocar clics, a menudo a costa de la precisión y la relevancia del contenido. Los clickbaits suelen utilizar frases exageradas, sensacionalistas o engañosas para incrementar el tráfico hacia ciertos sitios web, pero a menudo no cumplen con las expectativas creadas por el titular.
- **Credibility (Credibilidad).** Se refiere al proceso de evaluar la veracidad y la confiabilidad de la información y las fuentes en contextos digitales. Este proceso es crucial para combatir la desinformación y asegurar que el contenido que se consume y comparte en línea sea fiable y preciso.
- **Writing style (Estilo de escritura).** se refiere al análisis de las características estilísticas del texto para identificar patrones que puedan indicar desinformación o contenido manipulado, como por ejemplo patrones sensacionalistas. Este enfoque puede ser útil para diferenciar entre noticias verdaderas y falsas al analizar cómo se presentan y estructuran las noticias.

### 2.3.14. Similitud textual semántica

La similitud textual semántica trata de determinar qué tan similares son dos piezas de texto. Se refiere a la tarea de evaluar el grado de equivalencia en el significado entre dos fragmentos de texto. Esta evaluación va más allá de la coincidencia superficial de palabras y se centra en capturar la relación semántica entre los textos, considerando contextos y sinónimos, entre otros factores. Las tareas relacionadas son la identificación parafraseada o duplicada.

Por ejemplo, si usamos la herramienta SentEval<sup>10</sup> con las siguientes frases:

**Premisa:**

“Fuga de amoníaco líquido mata a 15 personas en Shanghai”.

**Hipótesis:**

“Fuga de amoníaco líquido mata al menos a 15 personas en Shanghai”.

Nos da un valor de salida: 4.6 (la mayoría de herramientas asignan una puntuación del 1 al 5).

### 2.3.15. Identificación de paráfrasis

La detección de paráfrasis es un problema de clasificación de PLN donde, dado un par de oraciones, el sistema determina la similitud semántica entre las dos oraciones. Si las dos oraciones tienen el mismo significado, entonces se etiqueta como paráfrasis; de lo contrario, se etiqueta como no parafraseado.

Esta técnica se puede aplicar, por ejemplo, al problema de detección de plagio, ya que no solo mide el *match* léxico entre ambos textos, sino también el *match* semántico.

.....  
10. [http://nlpprogress.com/english/semantic\\_textual\\_similarity.html](http://nlpprogress.com/english/semantic_textual_similarity.html)

Ejemplo tomado de Conneau and Kiela (2018):

**Premisa:**

“El procedimiento generalmente se realiza en el segundo o tercer trimestre”.

**Hipótesis:**

“La técnica se utiliza durante el segundo y, ocasionalmente, el tercer trimestre del embarazo.”

**Salida:**

Parafrase.

## 2.3.16. Inferencia del Lenguaje Natural

La inferencia del lenguaje natural es la tarea de determinar si una “hipótesis” es verdadera (implicación), falsa (contradicción) o indeterminada (neutral) dada una “premisa”. Los siguientes ejemplos muestran cada uno de los posibles tipos<sup>11</sup>:

### a) Contradicción

**Premisa:**

“Un hombre inspecciona todo el recinto antes de abrir las puertas.”

**Hipótesis:**

“El hombre duerme”.

**Resultado:**

Contradicción.

.....  
11. [http://nlpprogress.com/english/natural\\_language\\_inference.html](http://nlpprogress.com/english/natural_language_inference.html)

## **b) Neutral**

**Premisa:**

«Un hombre mayor y uno más joven sonrían.»

**Hipótesis:**

«Dos hombres sonrían y se ríen de los gatos que juegan en el suelo.»

**Resultado:**

Neutral.

## **c) Implicación**

**Premisa:**

“Un partido de fútbol con varios hombres jugando.”

**Hipótesis:**

“Algunos hombres practican algún deporte.”

**Resultado:**

Implicación.

## **2.3.17. Implicación textual**

Describe el problema en el que se requiere reconocer la relación direccional entre fragmentos de texto. La relación se mantiene siempre que la verdad de un fragmento de texto se sigue de otro texto. Los textos vinculantes y vinculados se denominan texto (t) e hipótesis (h), respectivamente.

“*t* implica *h*” ( $t \rightarrow h$ ) si, típicamente, una lectura humana *t* inferiría que *h* es probablemente verdadera.

La relación es direccional porque incluso si “*t* implica *h*”, la inversa “*h* implica *t*” es mucho menos segura.

**Premisa(t):**

“Una niña pequeña está montada en un carrusel”.

**Hipótesis(h):**

“El carrusel se mueve”.

**Resultado:**

Implicación.

## 2.3.18. Modelado de temas

El análisis de temas (también llamado detección de temas, modelado de temas o extracción de temas) es una técnica de aprendizaje automático que organiza y comprende grandes colecciones de datos de texto, mediante la asignación de etiquetas o categorías según el tema o temas de cada texto individual.

¿Qué es un tema en el modelado de temas? Es un tipo de modelo estadístico para descubrir los temas abstractos que ocurren en una colección de documentos. Los temas producidos por las técnicas de modelado de temas son grupos de palabras similares.

Dado que un documento trata sobre un tema en particular, uno esperaría que aparecieran palabras específicas en el documento con más o menos frecuencia.

Veamos un ejemplo de TopicModeling<sup>12</sup>, TopicModeling<sup>2</sup><sup>13</sup>:

- “perro” y “hueso” aparecerán con más frecuencia en documentos sobre perros,
- “gato” y “miau” aparecerán en documentos sobre gatos,
- y “el” y “es” aparecerán aproximadamente por igual en ambos.

Dos métodos de modelado de temas mas usados son:

- Análisis semántico latente (LSA).
- y Asignación de Dirichlet latente (LDA1).

### 2.3.18.1. Extracción de eventos

Es el proceso de recopilar conocimiento sobre incidentes periódicos que se encuentran en los textos, identificando automáticamente información sobre lo que sucedió y cuándo sucedió.

Un elemento vital es identificar el tema central para conseguir su seguimiento.

Por ejemplo<sup>14</sup>:

1. 2018/10 – El gobierno del presidente Donald Trump **prohibió** a los países **importar petróleo iraní** con exenciones a siete países.
2. 2019/04 – El secretario de Estado de Estados Unidos, Mike Pompeo, **anunció** que su país **no abriría más excepciones** después de la fecha límite.
3. 2019/05 – Estados Unidos **terminó con exenciones** que permi-

.....  
12. <https://monkeylearn.com/blog/introduction-to-topic-modeling/>

13. <https://www.kaggle.com/rcushen/topic-modelling-with-lsa-and-lda>

14. <https://towardsdatascience.com/natural-language-processing-event-extraction-f20d634661d3>

tían a los países **\*\*importar petróleo de Irán\*\*** sin sufrir las sanciones estadounidenses.

La extracción de eventos brinda la capacidad de contextualizar la información, nos permite conectar eventos distribuidos en el tiempo y asimilar sus efectos, y cómo se desarrolla un conjunto de episodios a lo largo del tiempo.

### 2.3.18.2. Extracción de eventos en textos biomédicos

Por otro lado, podemos ver otro tipo de extracciones de eventos para dominios concretos como por el ejemplo en textos biomédicos. Ver más detalles en KaggleEvent y eHealthKD2020.



Figura 17: Extracción de entidades, eventos y relaciones  
Fuente: <https://knowledge-learning.github.io/ehealthkd-2020/>



## 2.4. Recursos léxicos

Los recursos léxicos proporcionan al procesamiento del lenguaje natural la base de conocimientos sobre el significado, la estructura y el uso de las palabras en un idioma. Estos recursos son esenciales para una amplia gama de tareas en PLN, como la comprensión de texto, la traducción automática, el análisis de sentimientos y la generación de texto, entre otras, como hemos visto en la sección anterior.

Sin estos recursos, sería extremadamente difícil para las máquinas entender y generar lenguaje humano de manera efectiva. Utilizando estos recursos, los sistemas de PLN pueden mejorar significativamente en tareas como por ejemplo la desambiguación semántica o el análisis sintáctico.

Los recursos léxicos más habituales son:

- **Desambiguación Semántica:** Los recursos léxicos ayudan a determinar el significado correcto de una palabra en un contexto dado. Por ejemplo, la palabra “banco” puede referirse a una institución financiera o a un asiento en un parque, dependiendo del contexto.

Los recursos más utilizados son:

- Wordnet y Eurowordnet. Wordnet (Miller, 1990) es una base de datos léxica monolingüe que agrupa palabras en conjuntos de sinónimos (synsets), proporciona definiciones cortas y muestra las relaciones semánticas entre estas. EuroWordNet (Vossen, 1998) es similar a WordNet, pero para múltiples idiomas europeos proporcionando un módulo interlingua en el que relaciona la palabra con una acepción concreta de un idioma origen con un idioma destino.
- FrameNet (Baker, 1998): Ofrece una base de datos de marcos

semánticos, describiendo las relaciones entre los participantes de una acción y cómo se expresan en lenguaje natural.

- VerbNet (Schuler, 2005): Clasifica los verbos ingleses en un amplio conjunto de clases de verbos con información sobre su comportamiento sintáctico y semántico.
- **Análisis Morfológico:** Permiten descomponer palabras en sus componentes morfológicos (raíces, prefijos, sufijos) para entender su estructura y significado. Esto es crucial para el procesamiento de idiomas con morfología compleja.

Los recursos más utilizados son:

- MorphoLex (Sánchez-Gutiérrez *et al.*, 2018): Proporciona información sobre la estructura morfológica de las palabras, útil para tareas de análisis morfológico y generación de lenguaje.
- CELEX (Baayen *et al.* 1996): Ofrece información detallada sobre la frecuencia de las palabras, su pronunciación, morfología y estructura sintáctica en varios idiomas.
- **Análisis Sintáctico:** Facilitan la identificación de las relaciones gramaticales entre las palabras en una oración, lo que es fundamental para construir árboles sintácticos y para la generación de lenguaje.

Los recursos más utilizados son:

- VerbNet.
- Penn Treebank Corpus (Marcus *et al.* 1993) contiene una colección de textos en inglés anotados con información sintáctica y morfológica. La anotación incluye la estructura sintáctica de las oraciones, con etiquetas que indican las categorías gramaticales (como sustantivos, verbos, adjetivos, etc.) y las relaciones entre palabras (como sujeto, objeto, modificador, etc.)

- **Otros recursos:**

- Ontologías y Bases de conocimiento:
  - DBpedia (Lehmann et al. 2015): Extrae información estructurada de Wikipedia, proporcionando una base de conocimiento amplia y rica en datos.
  - YAGO (Suchanek, 2007): Una ontología que combina la información de Wikipedia y WordNet, útil para la desambiguación semántica y la inferencia de conocimiento.
- Corpus anotados. Además del corpus del Penn TreeBank mencionado anteriormente, algunos de los más usados son:
  - 1) CoNLL-2003 (Sang & Veenstra, 1999): Contiene anotaciones para el reconocimiento de entidades nombradas (NER) en inglés, alemán, neerlandés y español.
  - 2) SQuAD (Stanford Question Answering Dataset) (Rajpurkar *et al.* 2016). Un corpus de preguntas y respuestas sobre artículos de Wikipedia en inglés.
  - 3) Brown Corpus (Francis & Kucera, 1964).. Contiene textos en inglés con anotaciones de partes del discurso y género.
  - 4) TIMIT (Garofolo, *et al.* 1993). Corpus de habla en inglés con anotaciones fonéticas y transcripciones, útil para reconocimiento de habla.
  - 5) Google Ngrams (Michel *et al.* (2011). Contiene datos de frecuencias de n-gramas en inglés extraídos de libros digitalizados.
  - 6) Corpus CESS-CAST (Martí, *et al.* 2007). Un corpus anotado con partes del discurso y frases en español, útil para análisis sintáctico y etiquetado gramatical.

- 7) Corpus de Referencia del Español Actual (CREA) (RAE, 2005). Un corpus del español moderno, proporcionado por la Real Academia Española, que incluye textos de varios géneros.
- 8) AnCora (Taule, *et al.* 2008). Corpus anotado para el análisis de texto en español, que incluye anotaciones de partes del discurso y sintaxis
- 9) Spanish Billion Words Corpus (SBWC) (Rodríguez & Gómez, 2014). Un corpus de texto grande en español con una variedad de géneros y fuentes.
- 10) TASS (SemEval Task on Sentiment Analysis) (Taboada, *et al.* 2016). Incluye anotaciones para el análisis de sentimientos en textos en español.

# 3. Inteligencia Artificial

El término “inteligencia artificial” (IA) parece que ha irrumpido en nuestras vidas de repente y se ha convertido en la panacea que todo lo soluciona. No crean que la inteligencia artificial ha surgido en el presente, más bien es algo del siglo pasado, los primeros trabajos sobre inteligencia artificial datan de mediados de siglo XX. Frank Rosenblatt en 1956 creó el Perceptrón (Rosenblatt, 1958), la unidad de cálculo desde donde nacerían y evolucionarían las actuales redes neuronales artificiales. El perceptrón no es más que la simulación matemática del funcionamiento de una neurona humana.

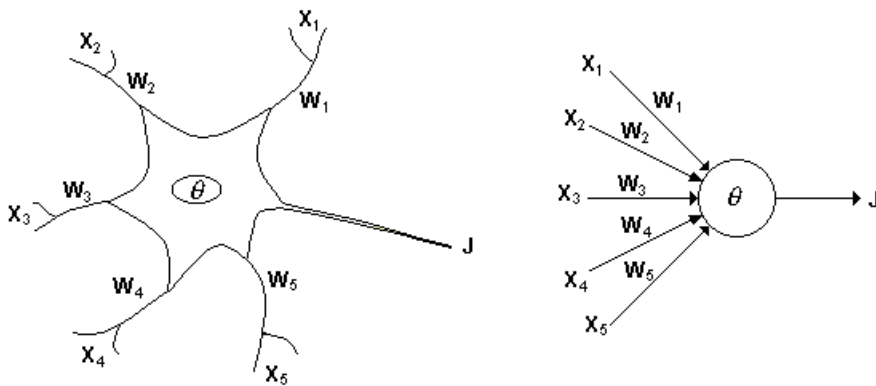


Figura 18: Neurona natural y artificial (Perceptrón)

La RAE, define el término “inteligencia artificial” como “programa informático que ejecuta operaciones comparables con las que realiza la mente humana, como el aprendizaje o el razonamiento lógico”.

Según John Macrthy, considerado el padre de la inteligencia artificial, “La inteligencia artificial es la habilidad de hacer que las máquinas hagan cosas que normalmente requieren inteligencia humana”.

El procesamiento del lenguaje natural es la piedra angular de la inteligencia artificial generativa porque permite a las máquinas, como hemos dicho, entender y generar lenguaje de manera autónoma. La capacidad de procesar lenguaje natural es fundamental para que las máquinas puedan aprender de grandes cantidades de datos, comprender el contexto y las relaciones semánticas y generar respuestas coherentes y relevantes. Pero antes de profundizar en la inteligencia artificial generativa debemos dejar claro que existen otros tipos de inteligencia artificial que se han usado y se siguen usando para la resolución de otros tipos de problemas diferentes a la generación de nuevo contenido.

No debemos acotar el término “inteligencia artificial” a las redes neuronales; existen otros métodos, como los métodos estadísticos, que forman parte de lo que conocemos como aprendizaje automático y que son adecuados para resolver los diferentes tipos de problemas. Si bien es cierto que lo que podríamos llamar la inteligencia artificial moderna resuelve la mayoría de los problemas con redes neuronales.

## 3.1. Tipos de inteligencia artificial

Muchas veces se usa el término “inteligencia artificial” en genérico, pero se deben distinguir 5 tipos en función de su aplicabilidad, como son:

### 1. IA descriptiva.

La inteligencia artificial descriptiva se centra en el análisis y resumen de datos históricos para proporcionar una comprensión clara de lo que ha ocurrido en el pasado. Utiliza técnicas de minería de datos, análisis estadístico y visualización de datos para identificar patrones, tendencias y relaciones en los conjuntos de datos. Los métodos más habituales en la inteligencia artificial descriptiva incluyen el análisis de series temporales para identificar tendencias a lo largo del tiempo, el *clustering* para agrupar datos similares y las visualizaciones de datos como gráficos de barras, histogramas y diagramas de dispersión para representar la información de manera comprensible.

Ejemplos de aplicaciones incluyen informes financieros que muestran el rendimiento pasado de una empresa; análisis de ventas que destacan los productos más vendidos en un período determinado, y visualizaciones de datos que ilustran la distribución demográfica de una población en un área específica.

### 2. IA predictiva.

La inteligencia artificial predictiva se enfoca en el uso de datos históricos y algoritmos de aprendizaje automático para hacer predicciones sobre eventos futuros. Utiliza métodos como regresión lineal y logística, árboles de decisión, bosques aleatorios (*random forests*), redes neuronales y modelos de series temporales para anticipar tendencias y comportamientos.

Ejemplos de aplicaciones de la IA predictiva incluyen la predicción de ventas futuras en el comercio minorista; el pronóstico de la demanda de energía en servicios públicos; la predicción de fallos en equipos industriales para mantenimiento predictivo, y la evaluación del riesgo crediticio en el sector financiero. Estos modelos permiten a las organizaciones tomar decisiones informadas basadas en las previsiones derivadas de datos históricos.

### **3. IA prescriptiva.**

La inteligencia artificial prescriptiva va un paso más allá de la predictiva al no solo anticipar lo que podría suceder, sino también recomendar acciones específicas para influir en esos resultados futuros. Utiliza métodos como la optimización matemática, simulaciones, algoritmos de aprendizaje por refuerzo y análisis de decisiones para determinar la mejor estrategia o curso de acción en un contexto dado.

Ejemplos de aplicaciones de la IA prescriptiva incluyen la planificación y optimización de rutas logísticas en el transporte y la cadena de suministro; la gestión de inventarios en tiempo real; la personalización de ofertas y promociones en el comercio electrónico, y la planificación de tratamientos personalizados en el ámbito de la salud. Estas soluciones ayudan a las organizaciones a tomar decisiones más efectivas y eficientes basadas en análisis avanzados y recomendaciones automatizadas.

### **4. IA adaptativa.**

La inteligencia artificial adaptativa se caracteriza por su capacidad para modificar su comportamiento y mejorar su funcionalidad en respuesta a cambios en el entorno o en los datos de entrada. Utiliza métodos como el aprendizaje por refuerzo, algoritmos evolutivos y sistemas de aprendizaje en línea, que permiten a los modelos ajustar continuamente sus parámetros y estrategias basándose en la retroalimentación recibida.



Ejemplos de aplicaciones de la IA adaptativa incluyen sistemas de recomendación que ajustan dinámicamente las sugerencias de contenido en plataformas de streaming según las preferencias cambiantes de los usuarios; vehículos autónomos que adaptan su conducción en tiempo real a las condiciones del tráfico y del clima, y aplicaciones de ciberseguridad que detectan ataques y vulneraciones adaptando las medidas de protección.

## 5. IA generativa.

Crea nuevos contenidos o datos en función de los datos de aprendizaje usados. En este caso el ejemplo es bien conocido, y hay que reconocer que es el que ha socializado a la inteligencia artificial: no es otro que ChatGPT. Si bien esto es un producto, hablaremos más de la tecnología que lo sustenta que es la de los modelos de lenguaje (*Large Language Models*, en inglés; LLM, su abreviatura).

Igual que decíamos que la IA es la panacea que todo lo soluciona, aparecen voces que advierten de los riesgos de que la inteligencia artificial gobierne a las personas. A corto plazo, se puede afirmar que eso no va a pasar, pero a la velocidad con la que la IA avanza, no se sabe hasta dónde llegará su capacidad de razonamiento y de tomar sus propias decisiones. La IA debe estar al servicio del ser humano y debe servir como apoyo para mejorar la vida de las personas, por tanto, se debe velar por un desarrollo responsable y ético.

No se puede rechazar directamente la IA por los riesgos que conlleva, son innumerables los beneficios en diversos ámbitos, como el de la salud, donde es capaz de descifrar la estructura molecular del ADN o ARN, entre otras miles de aplicaciones. Como curiosidad quiero decirles que este desarrollo proviene a partir de un algoritmo que aprendió a jugar al ajedrez dándole solo las instrucciones del juego y aprendió a base de jugar millones de partidas contra sí mismo.

## 3.2. Inteligencia artificial generativa

En esta lección nos vamos a centrar en la IA generativa y en cómo influye el procesamiento del lenguaje natural. La IA generativa es capaz de crear un nuevo contenido, ya sea texto, imagen o voz. Es capaz de generar una conversación natural, o un texto, o resumir un documento, etc. Los resultados que nos ofrecen los LLMs han maravillado a la sociedad por su calidad y por las enormes posibilidades de aplicación que nos ofrecen: son una especie de navaja suiza o herramienta multiuso.

La IA generativa se ha convertido en ese “cuñado” que todo lo sabe; le preguntas de cualquier tema y lo sabe, te da, incluso, más explicaciones de las que necesitas. E incluso, como “el cuñado”, a veces se lo inventa: es lo que se conoce como alucinación. Más allá del conocimiento que atesora, la clave de su éxito está en la generación de contenido de forma natural, que sea imperceptible si el autor es una persona o una máquina. Alan Turing en 1950, mediante la publicación del artículo titulado «Inteligencia», proponía lo que se conoce como “la prueba de Turing”. En ella, un humano interactúa por escrito con dos participantes: una persona real y una máquina. Sin saber quién es quién, el interrogador intenta determinar cuál de ellos es el humano basándose en sus respuestas. Si la máquina logra engañarle haciéndole creer que es un humano, se considera que ha superado la prueba de inteligencia.

Los modelos de lenguaje han hecho realidad esa ciencia ficción de hace unos años y han iniciado lo que se conoce como la revolución 5.0. La interacción hombre-máquina se está simplificando hasta niveles de mantener una conversación natural con los dispositivos. Las aplicaciones o tareas a partir de estos modelos de lenguaje que la ciudadanía puede usar son numerosas, desde una traducción, hasta la elaboración de un escrito o a la planificación de un viaje, entre otras muchas.

Esta revolución se debe, principalmente, a 3 factores:

1. Acceso a los datos. La sociedad digital en la que vivimos ha supuesto la existencia de enormes cantidades de datos digitalizados. Los modelos de lenguaje actuales necesitan trillones de tokens para aprender; si hablamos de textos, serían palabras.
2. Capacidad de cómputo. La supercomputación ha permitido poder desarrollar algoritmos cada vez más complejos que puedan ofrecer resultados en un tiempo razonable. Esto ha dado pie al tercer factor.
3. Redes neuronales artificiales. Las redes neuronales han evolucionado mucho desde 1956. Actualmente, ya no son un pequeño conjunto de neuronas conectadas conocidas como redes neuronales superficiales, ahora son redes profundas con infinitud de capas de neuronas interconectadas y billones de parámetros. Sin la supercomputación sería imposible entrenar y utilizar estas redes.

La inteligencia artificial generativa crea nuevo contenido multimodal: video, imagen, voz o texto. En ello, el Procesamiento del Lenguaje Natural (PLN) juega un papel crucial dentro de la inteligencia artificial generativa. Según Chomsky, filósofo y lingüista, “El verdadero desafío en el procesamiento del lenguaje natural no es simplemente acumular y analizar datos lingüísticos, sino entender la estructura y el significado subyacente del lenguaje humano”. En definitiva, para que las máquinas sean verdaderamente inteligentes, deben ser capaces de entender el lenguaje y el contexto humano. El procesamiento del lenguaje natural es un componente crucial en este desafío.

Hace apenas dos años, en noviembre de 2022, la sociedad empezó a conocer de primera mano las potencialidades de la IA generativa con la presentación de ChatGPT. ¿Quién no conoce hoy qué es ChatGPT? ¿Quién no lo ha probado? La calidad y la naturalidad de los resultados que proporciona ha sorprendido a la sociedad, es capaz de generar texto como si fuera una perso-

na o asistente, te traduce, te sintetiza, te realiza tareas como la planificación de un viaje o te da un menú para cumplir una dieta de unas calorías determinadas. Desde el punto de vista tecnológico, es la combinación de un modelo de lenguaje, GPT3, GPT4 o GPT4.o, y un sistema conversacional o chatbot.

Alguien podría decir que antes existían los asistentes virtuales como Siri o Alexa, ¡¡¡cierto!!! quien no ha dicho:

*“Alexa, ¿qué tiempo va a hacer hoy en Logroño?”*

y nos ha contestado una voz sintetizada:

*“En Logroño hoy tendremos un día soleado con temperaturas máximas de 25 grados”.*

Estos sistemas buscaban la información y, si la encontraban, nos la proporcionaban. Sin embargo, los modelos de lenguaje actuales van mucho más allá, no buscan la información, la construyen, la generan. Si se pregunta a ChatGPT o cualquier otro modelo como Copilot de Microsoft o Gemini de Google:

*“¿puedes darme un discurso sobre inteligencia artificial de 20 minutos para una lección inaugural en la Universidad de La Rioja?”*

Nos lo escribiría, prometo que ni soy un holograma ni esta lección es cosecha de los modelos de lenguaje. No quiero continuar sin que quede claro que ya hoy, Siri o Alexa, incorporan algunos de estos modelos de lenguajes.

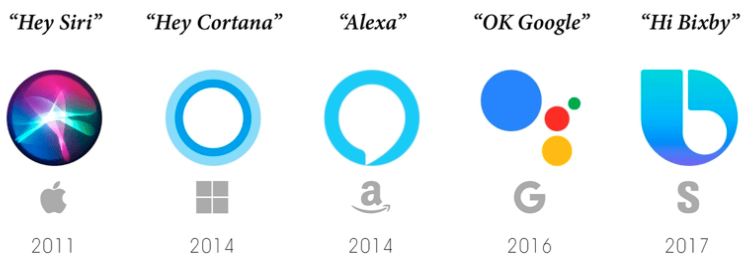


Figura 19: Asistentes virtuales y fechas de creación  
(Fuente: Medium)

La interacción hombre-máquina requiere de una comunicación natural y clara, requiere que la máquina entienda la forma de expresarse por parte de las personas y que las personas entiendan la respuesta de la máquina. Una vez que la máquina entienda la lengua (inglés, español, gallego, etc.) en la que se expresa una persona, ya se tiene parte del camino hecho. En la época actual caminamos hacia la personalización: no todos somos iguales. No solo se necesita que se interprete de forma adecuada un lenguaje, sino también que se adapte a las personas con las que se comunica. ¿Quién no querría que la voz o la terminología que usan estos sistemas sean los típicos de nuestra región de origen? ¿Por qué no tener sistemas de inteligencia artificial con las variantes dialectales de cada región? Imaginen que la máquina nos conteste a la pregunta anterior:

*“En Logroño hoy tendremos un día soleado con temperaturas máximas de 25 grados, te recomiendo que uses una pantaloneta”*

En lugar de decir unos “*pantalones cortos*” o “*bermudas*” términos que serían de uso más común en otras zonas de España, como por ejemplo, en Alicante de donde yo soy o, si es en formato hablado, con un acento típico de Logroño.

Este es uno de los retos actuales, la personalización de la interacción con las máquinas. En este sentido, permítanme que destaque el proyecto ILENIA en el que estamos trabajando en 4 de los centros más importantes en tecnologías del lenguaje de España: CENID, en Alicante; CITIUS, en Santiago de Compostela; HITz, en San Sebastián y el Barcelona Super Computer. Este proyecto está desarrollando modelos de lenguaje en las lenguas cooficiales de España, teniendo en cuenta las variedades dialectales de cada zona. No habla igual un valenciano del norte que uno del sur, no habla igual un español que una persona de Sudamérica o incluso una persona de La Rioja que una de Andalucía.

## 3.3. Tipos de problemas

Esta lección no pretende hacer una revisión exhaustiva de todos los tipos de métodos que pueden resolver los diferentes tipos de problemas. En esta clasificación se pretende, exclusivamente, citar algunos de los métodos más usados para resolver cada tipo de problemas. La elección del método depende de varios factores, además del tipo de problema, como por ejemplo, el tipo de datos de los que se dispone o la calidad de estos.

Los tipos de problemas que se pueden resolver son:

### 3.3.1. Clasificación

La clasificación es el problema de identificar entre dos o más categorías a cuál de ellas pertenece una nueva observación, una vez entrenado el sistema con un conjunto de datos de entrenamiento cuyas categorías se conocen. Se diferencian dos tipos:

**a) Binaria.** Solo existen dos categorías. Los métodos más adecuados para resolver este tipo de problemas son:

- Métodos estadísticos: regresión logística, máquinas de vectores de soporte (SVM), K-Nearest Neighbors (KNN), árboles de decisión, Random Forest y Gradient Boosting.
- Redes neuronales apropiadas: Perceptrón Multicapa (MLP), Redes Neuronales Convolucionales (CNN) para imágenes, Redes Neuronales Recurrentes (RNN) para datos secuenciales.

**b) Multiclase.** Existen más de dos categorías a las que pueden pertenecer las muestras. Los métodos más adecuados para resolver este tipo de problemas son:

- Métodos estadísticos: regresión logística multinomial, SVM de múltiples clases, KNN, árboles de decisión, Random Forest y Gradient Boosting.
- Redes neuronales apropiadas: MLP, CNN y RNN.

### 3.3.2. Regresión

La regresión es un tipo de problema de predicción que trata de modelar la relación entre una variable dependiente y una o más variables independientes.

- Métodos estadísticos: regresión lineal, regresión polinómica, KNN para regresión, árboles de decisión para regresión, Random Forest y Gradient Boosting.
- Redes neuronales apropiadas: MLP, Redes Neuronales Convolucionales (CNN) para series temporales o datos espaciales, Redes Neuronales Recurrentes (RNN) para series temporales.

### 3.3.3. Reconocimiento de imágenes

El reconocimiento de imágenes es el proceso de identificar y etiquetar objetos, personas, lugares o acciones en imágenes o secuencias de imágenes.

#### ***a) Clasificación de imágenes***

- Métodos estadísticos: SVM con características extraídas manualmente, Random Forest, Gradient Boosting y KNN con características extraídas manualmente.
- Redes neuronales apropiadas: CNN.

### ***b) Detección de objetos***

- Métodos estadísticos: métodos basados en características (SIFT, HOG), SVM o Random Forest, Gradient Boosting y técnicas de segmentación tradicionales (k-means).
- Redes neuronales apropiadas: CNN, redes neuronales convolucionales profundas (Deep CNN), Redes de Propuesta de Región (RPN) en modelos como Faster R-CNN.

### ***c) Segmentación de imágenes***

- Métodos estadísticos: segmentación basada en umbrales, segmentación de crecimiento de regiones, Clustering (k-means, Gaussian Mixture Models) y métodos de contorno activo.
- Redes neuronales apropiadas: Fully Convolutional Networks (FCN) y U-Net.

## **3.3.4. Procesamiento de Lenguaje Natural (PLN)**

El PLN es un campo de la inteligencia artificial que se centra en la interacción entre computadoras y lenguaje humano, permitiendo que las máquinas comprendan, interpreten y respondan al lenguaje de manera efectiva.

### ***a) Análisis de sentimientos***

- Métodos estadísticos: regresión logística, SVM, Naive Bayes, Random Forest y Gradient Boosting.
- Redes neuronales apropiadas: RNN, Long Short-Term Memory (LSTM) y transformers.



### ***b) Traducción automática***

- Métodos estadísticos: modelos estadísticos de traducción (Phrase-based Machine Translation).
- Redes neuronales apropiadas: Sequence-to-Sequence (Seq2Seq) con LSTM y transformers.

### ***c) Generación de texto***

- Métodos estadísticos: modelos estadísticos de lenguaje n-gram (generación de texto basada en la probabilidad de secuencias de palabras).
- Redes neuronales apropiadas: RNN, LSTM y Transformers (como GPT).

## **3.3.5. Series temporales y predicción**

El análisis de series temporales y predicción se centra en el uso de datos históricos para prever valores futuros de una variable dependiente a lo largo del tiempo.

### ***a) Predicción de valores futuros***

- Métodos estadísticos: modelos ARIMA (Autoregressive Integrated Moving Average), modelos de suavizamiento exponencial, regresión lineal con términos de series temporales y modelos GARCH (Generalized Autoregressive Conditional Heteroskedasticity).
- Redes neuronales apropiadas: RNN, LSTM y Gated Recurrent Unit (GRU).

### ***b) Análisis de series temporales***

- Métodos estadísticos: modelos ARIMA, modelos de suavizamiento exponencial, modelos de descomposición de series temporales y modelos GARCH.
- Redes neuronales apropiadas: RNN, LSTM, GRU.

### 3.3.6. Recomendación

Los sistemas de recomendación son una subclase de sistemas de filtrado de información que buscan predecir la “valoración” o “preferencia” que un usuario daría a un elemento.

- Métodos estadísticos: filtrado colaborativo, filtrado basado en contenido, modelos de factorización de matrices y modelos de regresión logística multinivel.
- Redes neuronales apropiadas: Autoencoders, CNN para procesamiento de contenido visual y RNN para secuencias de interacción.

### 3.3.7. Reconocimiento de voz y audio

El reconocimiento de voz y audio se refiere a la capacidad de los sistemas de convertir el habla humana y otros sonidos en texto y de identificar y clasificar diferentes sonidos.

#### **a) Reconocimiento de voz**

- Métodos estadísticos: modelos ocultos de Markov (HMM), Gaussian Mixture Models (GMM) y modelos estadísticos basados en n-gramas.
- Redes neuronales apropiadas: RNN, LSTM y Transformers (como Wav2Vec).

#### **b) Clasificación de audio**

- Métodos estadísticos: SVM, KNN, Random Forest, Gradient Boosting y Naive Bayes.
- Redes neuronales apropiadas: CNN, RNN y LSTM.

### 3.3.8. Juegos y agentes autónomos

Los juegos y agentes autónomos se enfocan en el desarrollo de algoritmos que permitan a los agentes aprender a tomar decisiones óptimas a través de la interacción con su entorno.

#### **a) Aprendizaje por refuerzo**

- Métodos estadísticos: algoritmos clásicos de aprendizaje por refuerzo (Q-learning, SARSA), programación dinámica y algoritmos basados en teoría de juegos.
- Redes neuronales apropiadas: redes de aprendizaje profundo por refuerzo (Deep Q-Networks, DQN), redes de política proximal (PPO) y algoritmos Actor-Critic.

### 3.3.9. Modelos generativos

Los modelos generativos se utilizan para crear nuevos datos similares a un conjunto de datos existente, como imágenes, música y texto.

#### **a) Generación de imágenes**

- Métodos estadísticos: modelos basados en autoencoders variacionales (VAE), modelos estadísticos basados en reglas y en métodos tradicionales de gráficos por computadora.
- Redes neuronales apropiadas: Redes Generativas Antagónicas (GAN), Autoencoders Variacionales (VAE).

#### **b) Generación de música y arte**

- Métodos estadísticos: modelos basados en autoencoders, algoritmos evolutivos y técnicas de recombinación y transformación de patrones musicales.
- Redes neuronales apropiadas: GAN, RNN y LSTM.

### 3.3.10. Detección de anomalías

La detección de anomalías consiste en identificar patrones inusuales o no esperados en los datos, que pueden indicar problemas o eventos interesantes.

- Métodos estadísticos: análisis de componentes principales (PCA), Clustering (k-means, DBSCAN), modelos estadísticos de detección de cambios y modelos de series temporales.
- Redes neuronales apropiadas: Autoencoders, redes neuronales convolucionales (CNN), redes neuronales recurrentes (RNN).

### 3.3.11. Optimización y Control

La optimización y control se refiere a encontrar la mejor solución para un problema, bajo un conjunto de restricciones, y tomar decisiones para alcanzar los objetivos deseados.

- Métodos estadísticos: programación lineal, programación no lineal, algoritmos genéticos, métodos de búsqueda heurística y algoritmos de Monte Carlo.
- Redes neuronales apropiadas: redes de aprendizaje por refuerzo (Reinforcement Learning Networks), redes de programación neuro-evolutiva.

En resumen, aunque las redes neuronales son herramientas poderosas y versátiles para una amplia variedad de problemas, existen muchos otros métodos estadísticos y de aprendizaje automático que pueden ser igualmente efectivos, dependiendo del problema específico y las características de los datos.

## 3.4. Redes neuronales

Las redes neuronales son la base de los modelos de lenguaje. Desde sus inicios, allá por 1958, se vieron las posibilidades que ofrecían en la aplicación a la resolución de diferentes tipos de problemas, como se muestra en la siguiente sección. Desde el perceptrón simple hasta ahora, las redes neuronales han evolucionado mucho (Figura 20).

Las redes neuronales pretenden simular el comportamiento de las neuronas del cerebro humano (Figura 21). Las neuronas se conectan a través de las dendritas, estimulándose unas a otras en función de la reacción de la neurona a la que están conectadas. Las redes neuronales tienen la misma arquitectura o estructura: un conjunto de capas, cada una con un número de neuronas (no tienen por qué ser el mismo) y en las que cada neurona de una capa se conecta con todas las de las capas siguientes. A cada una de estas conexiones se le asigna un peso, que será el que, una vez entrenada la red, le dé más o menos valor a la salida que produce esa neurona particular dentro de toda la red.

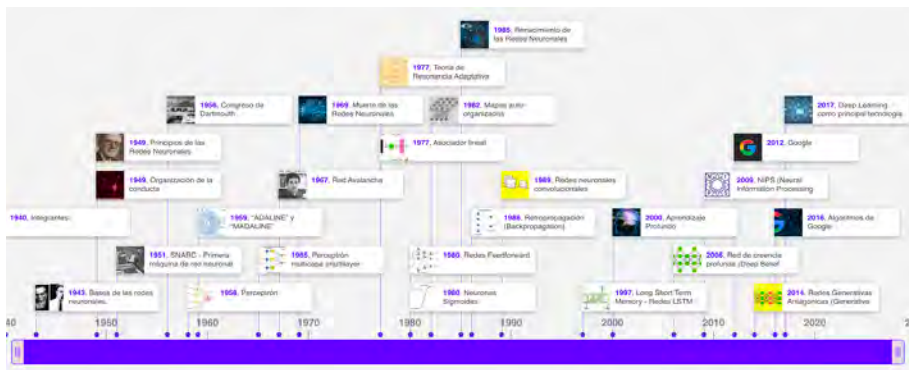


Figura 20: Evolución de las redes neuronales<sup>15</sup>

.....  
15. <https://www.timetoast.com/timelines/evolucion-de-las-redes-neuronales-artificiales-c20eb783-ed94-4675-ad4f-dde7e4b8470c>

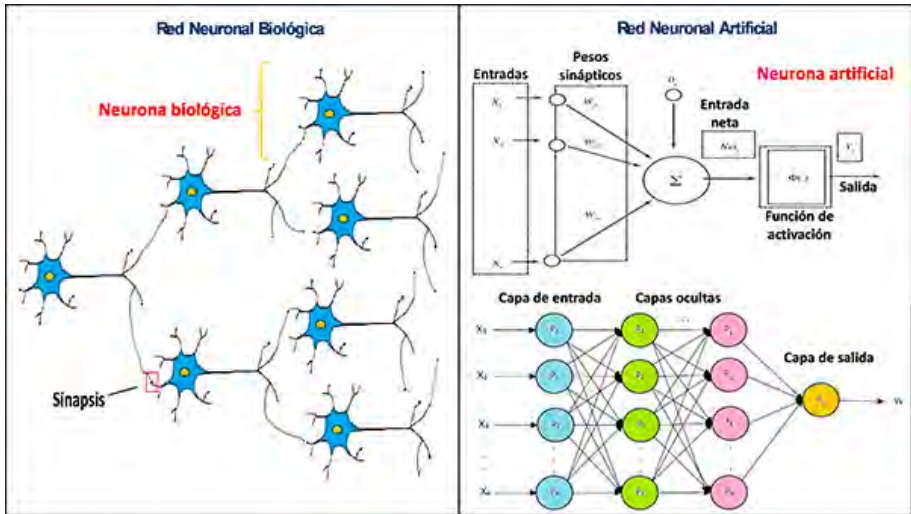
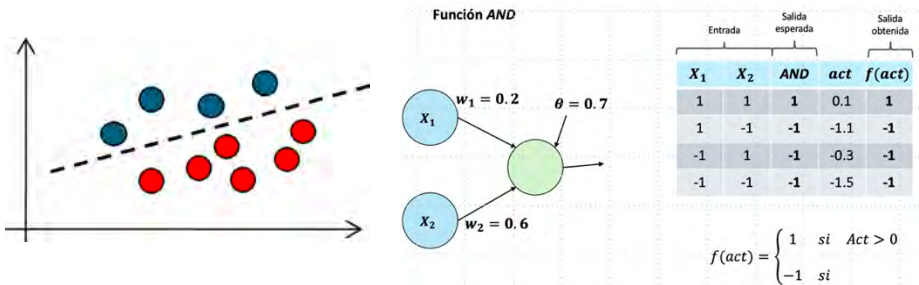


Figura 21: Simulación de una red neuronal biológica con una artificial  
(Fuente: Franco et al.)

### 3.4.1. Perceptrón

El **perceptrón simple** es capaz de resolver exclusivamente problemas linealmente separables, como por ejemplo los correspondientes a puertas lógicas OR, AND, etc.



El perceptrón simple aprende en la fase de entrenamientos. Para conseguir que aprenda, se le muestra un conjunto de ejemplos de entrada, calculando la salida para cada muestra. Esta salida se compara con el valor esperado y se ajustan los pesos para minimizar el error producido entre la entrada producida y la esperada. A continuación se muestra el proceso de aprendizaje de un perceptrón que simula una puerta lógica OR.

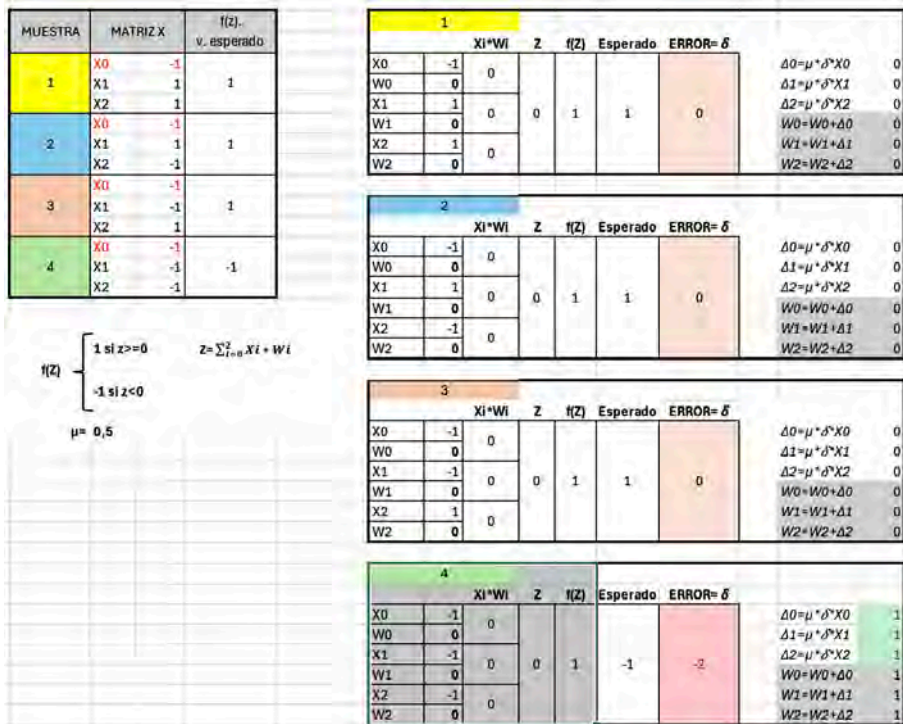


Figura 22: Entrenamiento de un perceptrón para la función OR

Como se observa en la Figura 22, para cada una de las 4 muestras se añade una característica adicional, que denominamos X0, con un valor fijo de -1 y se le asigna un peso inicial (W0) de 0; a esta entrada se le denomina entrada de control. Recordemos que la puerta lógica OR devuelve 1 si alguna de las entradas es 1 y devuelve 0 (-1) si ambas entradas

son 0 (-1). En la figura se observa que para la muestra 1 los valores de los pesos se han inicializado ambos ( $W_1$  y  $W_2$ ) con un valor de 1. El entrenamiento consiste en un conjunto de iteraciones, hasta que para todos los ejemplos que se les muestra el error producido es 0 o está por debajo del umbral establecido. En cada iteración de entrenamiento se le muestra el primer ejemplo (1,1) y el valor esperado (1). Como coincide con el valor obtenido por la función de activación no hay que corregirse los pesos. A continuación, se le muestra el siguiente ejemplo (1,-1) tampoco se produce error, en el tercer ejemplo (-1,1) tampoco se produce error, y en el último ejemplo (-1,-1), vemos que se produce un error (el valor calculado por la función de activación produce 1 y el esperado era -1). Por tanto, se produce un error de 2. Entonces tenemos que ajustar los pesos de la red ( $W_0$ ,  $W_1$  y  $W_2$ ). ¿Cómo se ajustan? ¿Se cambian de forma aleatoria? La respuesta es que se usa una fórmula de ajuste de pesos que tiene en cuenta el valor de la entrada ( $X_i$ ), el error producido ( $\delta$ ) y un factor de aprendizaje ( $\mu$ ). Para ajustar cada peso se le suma el valor de la multiplicación de  $\mu * \delta * X_i$ . Por tanto, el valor de cada peso se ajusta en función a la siguiente fórmula:

$$W'_i = W_i + \mu * \delta * X_i$$

Después de tres iteraciones, es decir, de mostrarle tres veces todos los ejemplos, el perceptrón converge. Para los cuatros ejemplos que se le muestra a la red, la salida esperada y la obtenida coinciden por tanto, el error ( $\delta$ ) es cero.



MUESTRA	MATRIZ X	f(z), v. esperado
1	X0 -1 X1 1 X2 1	1
2	X0 -1 X1 1 X2 -1	1
3	X0 -1 X1 1 X2 1	1
4	X0 -1 X1 -1 X2 -1	-1

$f(z) = \begin{cases} 1 & \text{si } z \geq 0 \\ -1 & \text{si } z < 0 \end{cases}$	
$\mu = 0,5$	$Z = \sum_{i=0}^2 X_i \cdot W_i$

1		XI*Wi	Z	f(Z)	Esperado	ERROR= $\delta$	$\Delta W = \mu \cdot \delta \cdot X_i$	
X0	-1	-1	1	1	1	0	$\Delta 0 = \mu \cdot \delta \cdot X_0$	0
W0	-1	1					$\Delta 1 = \mu \cdot \delta \cdot X_1$	0
X1	1	1					$\Delta 2 = \mu \cdot \delta \cdot X_2$	0
W1	1	1	1	1	1	0	$W_0 = W_0 + \Delta 0$	-1
X2	1	1					$W_1 = W_1 + \Delta 1$	1
W2	1	1					$W_2 = W_2 + \Delta 2$	1

2		XI*Wi	Z	f(Z)	Esperado	ERROR= $\delta$	$\Delta W = \mu \cdot \delta \cdot X_i$	
X0	-1	-1	1	1	1	0	$\Delta 0 = \mu \cdot \delta \cdot X_0$	0
W0	-1	1					$\Delta 1 = \mu \cdot \delta \cdot X_1$	0
X1	1	1					$\Delta 2 = \mu \cdot \delta \cdot X_2$	0
W1	1	1	-1	-1	-1	0	$W_0 = W_0 + \Delta 0$	-1
X2	-1	-1					$W_1 = W_1 + \Delta 1$	1
W2	1	1					$W_2 = W_2 + \Delta 2$	1

3		XI*Wi	Z	f(Z)	Esperado	ERROR= $\delta$	$\Delta W = \mu \cdot \delta \cdot X_i$	
X0	-1	-1	1	1	1	0	$\Delta 0 = \mu \cdot \delta \cdot X_0$	0
W0	-1	1					$\Delta 1 = \mu \cdot \delta \cdot X_1$	0
X1	-1	-1					$\Delta 2 = \mu \cdot \delta \cdot X_2$	0
W1	1	1	1	1	1	0	$W_0 = W_0 + \Delta 0$	-1
X2	1	1					$W_1 = W_1 + \Delta 1$	1
W2	1	1					$W_2 = W_2 + \Delta 2$	1

4		XI*Wi	Z	f(Z)	Esperado	ERROR= $\delta$	$\Delta W = \mu \cdot \delta \cdot X_i$	
X0	-1	-1	1	-1	-1	0	$\Delta 0 = \mu \cdot \delta \cdot X_0$	0
W0	-1	1					$\Delta 1 = \mu \cdot \delta \cdot X_1$	0
X1	1	-1					$\Delta 2 = \mu \cdot \delta \cdot X_2$	0
W1	1	1	-1	-1	-1	0	$W_0 = W_0 + \Delta 0$	-1
X2	-1	-1					$W_1 = W_1 + \Delta 1$	1
W2	1	-1					$W_2 = W_2 + \Delta 2$	1

### 3.4.2. Perceptrón multicapa

Ante la problemática que presenta el perceptrón simple, que solo es capaz de resolver problemas linealmente separables, el **perceptrón multicapa** es la primera red neuronal artificial compuesta por varias capas. Esta estructura (Figura 23: Perceptrón multicapa) le permite abordar problemas predictivos para una o más variables dependientes. Las neuronas de una capa se conectan con todas las neuronas de la siguiente capa. Normalmente tienen una capa de entrada y una de salida, y unas pocas capas ocultas.

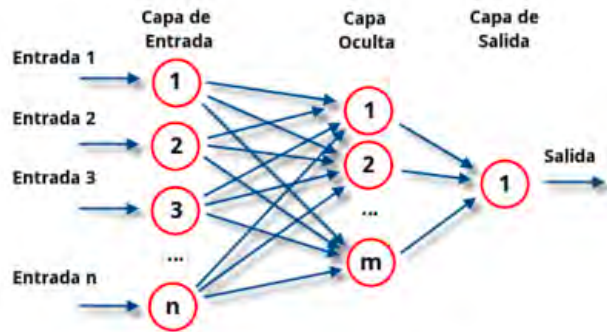


Figura 23: Perceptrón multicapa

### 3.4.3. Redes Backforward

El algoritmo de **backpropagation** hizo posible entrenar redes neuronales de múltiples capas de manera supervisada. A la diferencia entre el valor de salida y el esperado se le llama “error”. El ajuste de ese error se hace haciendo pequeños ajustes a los pesos de las capas anteriores, para ir minimizando el error en cada iteración. Por iteración se entiende el proceso de calcular la salida de todas las muestras de las que se dispone. Obviamente, el ajuste de los pesos no se hace de forma aleatoria, sino que se aplica una función de pérdida.

Como vemos (Figura 24) este aprendizaje es supervisado: en función del valor esperado y el producido, se van ajustando los pesos en “n” iteraciones, hasta que la pérdida es 0 o converge a un umbral permitido.

### 3.4.4. Redes neuronales recurrentes

Este algoritmo de propagación hacia atrás permite que las redes neuronales ya no solo propaguen la información exclusivamente hacia adelan-

te (feedback), sino que también puedan retroalimentarse, dando origen a un nuevo tipo de redes que se denominan **redes neuronales recurrentes** (RNN) (Williams & Zipser, 1989). Este tipo de redes son apropiadas para el manejo de datos secuenciales o series temporales. La principal característica de las RNN es la capacidad de retener datos procesados previamente. A medida que se procesa cada pieza de la secuencia, las conexiones internas de una RNN le permiten mantener una memoria interna o estado oculto. Este tipo de redes, aunque tienen memoria, la tienen a muy corto espacio: de hecho, solo recuerdan la aparición anterior. Son adecuadas para secuencias cortas y tareas donde las dependencias a corto plazo son suficientes.

La Figura 24 muestra el algoritmo de entrenamiento de una red *back-progation* usando la variante on-line.

### 3.4.5. Redes neuronales convolucionales

Otro de los avances que se produjeron fue la aparición de las **redes neuronales convolucionales** (CNN). Este tipo de redes es ideal para el reconocimiento de imágenes. La primera red se utilizó para reconocer caracteres manuscritos (LeCunn, 1998). La arquitectura tiene dos etapas: la primera es la de extracción de características y la segunda es la de clasificación. Una vez extraídas las características, hace un *pooling* para quedarse solo con las características relevantes. Este proceso puede repetirse varias veces (convolucionan). Luego se hace una nueva convolución y otro *pooling* que alimenta una red *feedforward* multicapa. La salida final de la red es un grupo de nodos que clasifican el resultado, por ejemplo, un nodo para cada número del 0 al 9 (es decir, 10 nodos, que se “activan” de a uno).

Backpropagation Variante On-line  
(estocástica)

**Comienzo:**

1 - Inicializar  $w_j^i$  aleatoriamente.  $\Delta w_j^i = 0$

2 - Hasta que la condición de parada se satisfaga:

2.1 - Para cada vector  $\langle x^d, t^d \rangle \in \mathcal{E}$  calcular:

-  $o^d$

- Para las neuronas de la capa de salida:  $d^o = (t^d - o^d) f'(Act^d)$

- Para las neuronas de las capas ocultas:  $d^h = \sum_{j=1}^C d^j w_h^j f'(Act^h)$

- Cambiar los pesos de acuerdo con:  $w_i^j = w_i^j + \eta \Delta w_i^j$   
donde  $\Delta w_i^j = \mathbf{h}^i o^j$

2.2 - Volver al paso 2.

**Fin.**

$C$ : Cantidad de neuronas en la capa siguiente a la neurona  $h$ .

$w_h^j$ : Peso con que la neurona  $j$  pondera el estímulo recibido desde la neurona  $h$ .

Notar que  $o^{ij}$  es la salida de la  $i$ -ésima neurona, siendo a su vez un impulso transmitido a las neuronas de la siguiente capa.

Figura 24: Algoritmo para el entrenamiento Backpropagation

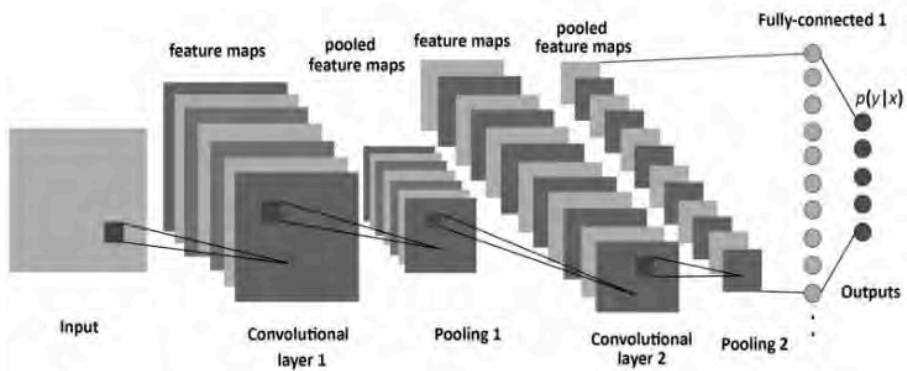


Figura 25: Arquitectura red CNN. (Fuente: IALATAM2019)

### 3.4.6. Redes LSTM

En 1996 aparecen las LSTM (Long short-term memory) (Hochreiter & Schmidhuber, 1996). Este tipo de red neuronal solventa el problema de recordar solo la aparición anterior y puede retener información que haya aparecido en varias palabras anteriores. Están diseñadas para manejar dependencias a largo plazo, con una estructura más compleja que incluye celdas de memoria y puertas de control. Son más efectivas y estables para tareas con secuencias largas. Son muy usadas en tareas de traducción automática y generación de texto. Este fue el principio de los modelos de lenguaje que veremos posteriormente. A partir de este momento empezó la era de lo que conocemos como aprendizaje profundo (*deep learning*).

Este diagrama (Figura 26) ilustra el flujo de datos a través de una capa de LSTM con entrada  $\mathbf{x}$ , salida  $\mathbf{y}$ , y con  $T$  unidades de tiempo. En el diagrama,  $\mathbf{h}_t$  determina la salida (también conocida como el *estado oculto*) y  $\mathbf{c}_t$  denota el *estado de celda* en la unidad de tiempo  $t$ .

Si la capa genera la secuencia completa, entonces genera  $\mathbf{y}_1, \dots, \mathbf{y}_T$ , que es equivalente a  $\mathbf{h}_1, \dots, \mathbf{h}_T$ . Si la capa genera únicamente la última unidad de tiempo, entonces genera  $\mathbf{y}_T$ , que es equivalente a  $\mathbf{h}_T$ . El número de canales en la salida coincide con el número de unidades ocultas de la capa de LSTM.

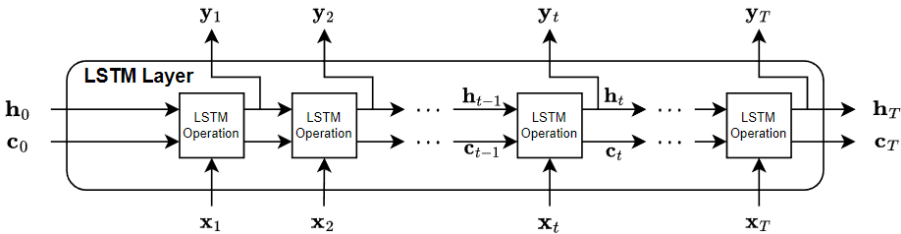


Figura 26: Arquitectura red LSTM

Fuente: <https://es.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html>

En la década del 2000: comienza la era del aprendizaje profundo y se popularizan la función de activación ReLU y la técnica de regularización Dropout, facilitando el entrenamiento de redes neuronales más profundas.

### 3.4.7. Generative Adversarial Networks (GAN)

**En 2014** se desarrollan las **Generative Adversarial Networks (GAN)** por Goodfellow *et al.* (2014). Las GAN introducen un enfoque innovador para la generación de datos sintéticos mediante un enfoque competitivo entre dos redes neuronales. La idea detrás de GAN es la de tener dos modelos de redes neuronales que compiten entre sí. Uno, llamado “generador”, toma inicialmente cualquier dato como entrada y genera muestras. El otro modelo, llamado “discriminador”, recibe a la vez muestras del generador y del conjunto de entrenamiento (real) y deberá ser capaz de diferenciar entre las dos fuentes. Estas dos redes juegan una partida continua donde el generador aprende a producir muestras más realistas y el discriminador aprende a distinguir entre datos reales y muestras artificiales. Estas redes son entrenadas simultáneamente para, finalmente, lograr que los datos generados no puedan diferenciarse de los datos reales.

La arquitectura consta de dos partes. Un generador  $G$  y un discriminador  $D$ .  $G$  toma el ruido como entrada y se lo da a  $D$ ,  $D$  estima la probabilidad de que la muestra proceda de los datos de entrenamiento en lugar de proceder de  $G$ . Durante el entrenamiento,  $G$  tiene que maximizar la probabilidad de que  $D$  cometa un error, mientras que  $D$  tiene que minimizar su error. La Figura 27 (Arquitectura de una red GAN) muestra los componentes básicos de la arquitectura. La mejor salida de  $D$  es 0,5, ya que esto significa que no puede decir si la entrada son datos de entrenamiento o datos generados.

Este tipo de redes son muy apropiadas para la generación de imágenes y videos.

El siguiente gran avance se produjo en 2017, con la aparición de los Transformer.

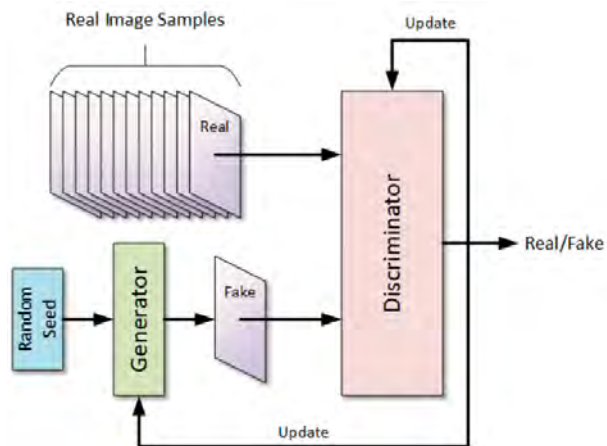


Figura 27: Arquitectura de una red GAN

Fuente: [https://semiengineering.com/knowledge\\_centers/artificial-intelligence/neural-networks/generative-adversarial-network-gan/](https://semiengineering.com/knowledge_centers/artificial-intelligence/neural-networks/generative-adversarial-network-gan/)

## 3.5. Modelos de lenguajes a gran escala (LLM)

En la sociedad, nos hemos acostumbrado a nombres comerciales como ChatGPT o Copilot, pero en realidad lo que hay detrás de toda esta revolución son los modelos de lenguajes.

Los modelos de lenguaje a gran escala (LLM) se han hecho muy populares debido a varios factores, como la mejora de la comprensión del lenguaje o la accesibilidad a estas tecnologías por parte de la sociedad. Esto ha llevado a la inteligencia artificial generativa a la vanguardia del interés público y a impulsar la adopción de la IA como herramienta fundamental para mejorar la productividad en diversas funciones empresariales y aplicaciones prácticas.

Para quienes no están inmersos en el ámbito empresarial, puede parecer que los LLM han aparecido de repente, junto con los avances recientes en IA generativa. Los LLM se han estado utilizando durante años para mejorar sus capacidades de comprensión y procesamiento del lenguaje natural (NLU y NLP). Esto ha ido de la mano de los avances en aprendizaje automático, modelos de machine learning, algoritmos, redes neuronales y arquitecturas de transformadores, que constituyen la base de estos sistemas de IA.

Los LLM son modelos fundamentales entrenados con enormes cantidades de datos, que proporcionan capacidades esenciales para múltiples aplicaciones y tareas. Esto contrasta con la creación de modelos específicos para cada aplicación, lo cual resulta prohibitivo en términos de costo e infraestructura, limita las sinergias y puede llevar a un rendimiento inferior.



Más allá de las palabras de Chomsky, ¿son necesarias las técnicas de PLN en los modelos de lenguaje? ¿Qué aporta entonces el procesamiento de lenguaje natural a estos modelos de lenguaje?

El lenguaje no es una secuencia de palabras que cumple unas reglas y por lo tanto es interpretable de forma unívoca. El lenguaje es complejo y con una serie de fenómenos lingüísticos, desde el nivel morfológico al semántico o, incluso, discursivo, como son la ambigüedad estructural o semántica, la elipsis, la anáfora y otros muchos, que deben ser resueltos para una correcta comunicación hombre-máquina. Un ejemplo típico para comprender la dificultad de que una máquina pueda resolver algunos de estos problemas está en la frase: “El policía vio al ladrón con el telescopio”. ¿Quién tenía el telescopio? ¿El ladrón? ¿El policía? Si ya es difícil para una persona, para una máquina lo es mucho más. Hay que dotar a la máquina de conocimiento y de contexto.

La realidad es, como decía Chomsky, que el Procesamiento del Lenguaje Natural (PLN) proporciona las bases y técnicas esenciales que permiten a los modelos de lenguaje comprender, procesar y generar lenguaje humano de manera efectiva. Desde la tokenización y el análisis sintáctico hasta la desambiguación de palabras y la generación de lenguaje natural. El PLN enriquece a los LLMs con la capacidad de manejar y aprovechar la complejidad del lenguaje, haciendo posible su aplicación en una amplia gama de tareas y dominios, facilitando la interacción con humanos e infiriendo o descubriendo nuevo conocimiento.

Los modelos de lenguaje actuales no solo generan texto, sino que también infieren información. Si se realiza a cualquier modelo de lenguaje actual la siguiente pregunta:

*“Sabiedo que Alicia tiene tres hermanos y una hermana,  
¿cuántas hijas tienen los padres de Alicia?”*

se obtiene la respuesta:

*“Los padres de Alicia tienen 2 hijas, Alicia y su hermana”.*

Como ven, el modelo de lenguaje ha sido capaz de inferir información. Si el modelo del lenguaje ha sido instruido solo con un conjunto de tokens (texto, voz e imágenes), ¿cómo ha sido posible inferir la respuesta? ¿Cómo sabe que existe alguna relación de Alicia con el término “hermana”? La realidad es que los modelos de lenguaje no solo aprenden a partir de textos, sino que también usan otros recursos, como taxonomías, ontologías o cualquier otro recurso léxico, incluso código fuente de programas informáticos que proporciona cierta jerarquía en el razonamiento.

Los LLM representan un avance importante en PLN e inteligencia artificial y son accesibles al público a través de interfaces como Chat GPT-3 y GPT-4, de OpenAI, respaldados por Microsoft. Otros ejemplos incluyen los modelos Llama de Meta y modelos de Google como BERT/ROBERTa y PaLM.

En resumen, los LLM están diseñados para comprender y generar texto similar a un humano, así como otros tipos de contenido, basándose en los grandes volúmenes de datos con los que se entrenan. Pueden inferir contexto, generar respuestas coherentes y relevantes, traducir a varios idiomas, resumir textos, responder preguntas y ayudar en tareas creativas como la escritura o generación de código.

Gracias a sus miles de millones de parámetros, los LLM pueden captar patrones complejos en el lenguaje y realizar una amplia variedad de tareas lingüísticas. Están revolucionando aplicaciones en campos como chatbots, asistentes virtuales, generación de contenido, asistencia en investigación y traducción de idiomas.

A medida que los LLM continúan evolucionando y mejorando, están preparados para transformar nuestra interacción con la tecnología y el acceso a la información, convirtiéndose en una parte integral del panorama digital actual.

Actualmente nos encontramos con infinitud de LLM.



eficiencia. La arquitectura propuesta es la que se conoce como Transformer que podemos ver en la Figura 29: Arquitectura TRANSFORMER. Los componentes principales de la arquitectura Transformer son:

- **Incrustaciones de entrada (*Input Embedding*)**. Este componente utiliza técnicas de incrustación (*word embedding*) para transformar el texto, que se trocea en las unidades de información básica (unos vectores de números). Por ejemplo, si la entrada es una oración, los trozos son palabras. Los vectores contienen información semántica y sintáctica, representada como números, y sus atributos se aprenden durante el proceso de entrenamiento.
- **Codificación posicional (*Positional Encodig*)**. La codificación posicional es un componente crucial en la arquitectura del transformador dado que lo que pretende es procesar de forma paralela, pero no debe olvidar el orden o secuencia de la entrada original.

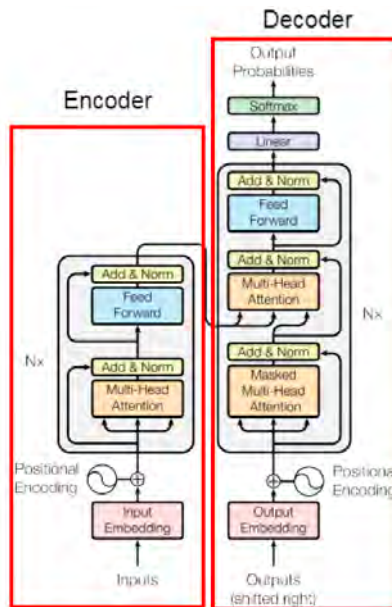


Figura 29: Arquitectura TRANSFORMER

- **Bloque transformador.** Un modelo de transformador típico tiene varios bloques de transformadores apilados juntos. Cada bloque transformador tiene dos componentes principales:
  - un mecanismo de autoatención con múltiples cabezales. La autoatención permite agrupar los elementos relevantes para el contexto: el significado de una palabra depende de las palabras que la acompañan en la frase.
  - una red neuronal de retroalimentación por posición.

La capa de alimentación directa tiene componentes adicionales que ayudan al modelo del transformador a entrenarse y funcionar de manera más eficiente. Cada bloque incluye lo siguiente:

- Conexiones que permiten el flujo de información de una parte de la red a otra, omitiendo ciertas operaciones intermedias.
  - La normalización de capas mantiene los números dentro de un rango para que el modelo aprenda sin problemas.
  - La transformación lineal consigue adaptar el modelo a la tarea en la que se está entrenando.
- **Bloques lineales y softmax.** Es en este momento donde la arquitectura predice la palabra que debe proporcionar. El resultado de esta capa es un conjunto de puntuaciones para cada token posible. La función softmax es la que toma las puntuaciones y las normaliza en una distribución de probabilidad.

### 3.5.1.1. Tamaño del modelo de lenguaje

Uno de los aspectos que más llama la atención cuando aparece un nuevo modelo es la cantidad de parámetros que tiene. ¿Cómo se decide el número de parámetros que tendrá un modelo?

El número de parámetros en un LLM se determina por una combinación de objetivos del modelo, elección de arquitectura, disponibilidad de recursos y resultados experimentales. Este proceso es crucial para balancear la capacidad del modelo para aprender y generalizar con la eficiencia computacional necesaria para entrenarlo y desplegarlo.

Algunos de los aspectos más importantes que influyen en esta decisión son:

### **1. Capacidad y complejidad del modelo**

- **Objetivo del modelo:** la complejidad de las tareas que el modelo debe realizar. Tareas más complejas y variadas pueden requerir un mayor número de parámetros.
- **Dimensionalidad de las representaciones:** la dimensionalidad de las representaciones internas (vectores de características) también afecta el número de parámetros. Representaciones de mayor dimensión pueden capturar más detalles y sutilezas del lenguaje.

### **2. Arquitectura del modelo**

- **Tipo de modelo:** diferentes arquitecturas de Transformer tienden a necesitar más parámetros en función de la capacidad de atención.
- **Número de capas:** el número de capas en la red afecta directamente el número de parámetros. Más capas significan más parámetros.
- **Tamaño de las capas:** el tamaño de las capas (es decir, el número de neuronas por capa) también impacta el número total de parámetros.

### **3. Capacidad de entrenamiento y recursos**

- **Disponibilidad de datos:** la cantidad y calidad de los datos disponibles para entrenar el modelo pueden influir en el número de parámetros.
- **Recursos computacionales:** la disponibilidad de recursos compu-

tacionales, como GPUs, y la infraestructura para entrenar el modelo también son factores limitantes. Modelos con más parámetros requieren más capacidad de procesamiento y memoria.

#### 4. Experimentación

- Pruebas y ajustes: durante el desarrollo del modelo se realizan pruebas y ajustes iterativos para encontrar el equilibrio adecuado entre el número de parámetros y el rendimiento del modelo.

Ejemplo: Arquitectura *Transformer*

Para un modelo basado en *transformers*, el número de parámetros se calcula, principalmente, considerando los siguientes componentes:

1. *Encoders* y *Decoders*: cada capa de *encoder* y *decoder* tiene múltiples subcomponentes, como la capa de atención y la red *feedforward*.
2. Mecanismo de atención: los parámetros en el mecanismo de atención incluyen matrices de pesos para las consultas (Q), claves (K) y valores (V), que se multiplican por el número de cabezas de atención y las dimensiones de estas matrices.
3. Redes *Feedforward*: cada capa incluye una red *feedforward* con sus propios conjuntos de pesos y sesgos.
4. *Embeddings*: los *embeddings* de entrada y salida, que mapean palabras a vectores de alta dimensión, también contribuyen significativamente al número total de parámetros.

Fórmula simplificada para parámetros en una capa del Transformer

$$\text{Número de Parámetros} \approx 4 \cdot (d_{\text{model}} \cdot d_{\text{ff}} + d_{\text{model}}^2)$$

donde:

- $d_{\text{model}}$  es la dimensionalidad del modelo.
- $d_{\text{ff}}$  es la dimensionalidad de la red *feedforward*.

El número de parámetros (las interconexiones entre neuronas existentes) es el que suele dar apellido al modelo, cuando se dice que es un 8B (8 billones americanos de parámetros) o un 175B. Para que se hagan una idea, GPT3 tiene 1.750 millones de parámetros, por lo que es un modelo 175B. Se pueden encontrar modelos de hasta 530B o incluso se especula, ya que no lo han hecho público, que GPT4 puede llegar a 1.000B. Obviamente, el tamaño de la red condiciona las necesidades tecnológicas de la máquina sobre la que se va a ejecutar. Por lo tanto, es necesario tener en cuenta la disponibilidad de recursos tecnológicos, como son las GPU que disponer. En la Figura 30; Modelos de lenguaje hasta marzo 2023, podemos ver diferentes LLM y el número de parámetros que cada una de ellas tiene. Como podemos observar, en 2023 los tamaños máximos son de 530B y en 2024 se espera que superen los 2T.

En España, el supercomputador más potente es el Marenostrum 5, que se encuentra en el Barcelona Supercomputing Center. Para realizar un modelo fundacional (modelo a partir de cero) necesitaría la dedicación del 50 % del superordenador durante 3 meses de cálculo, 24 horas al día.

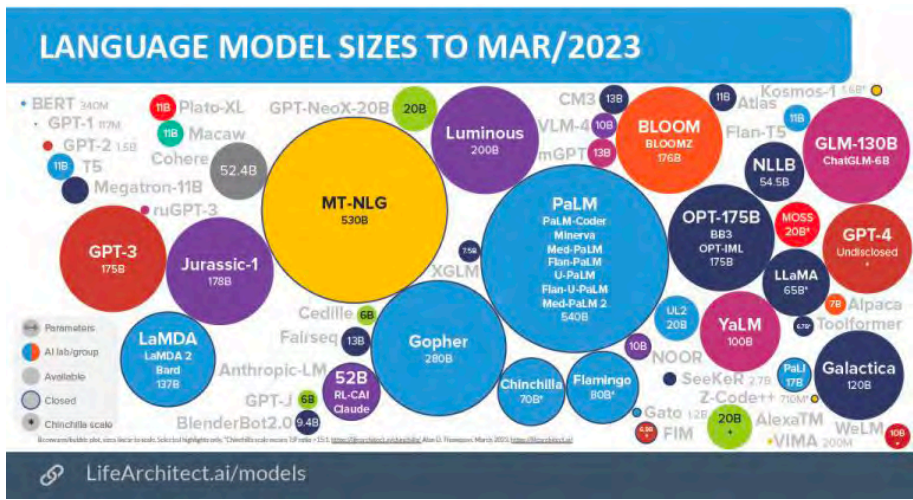


Figura 30: Modelos de lenguaje hasta marzo 2023

Fuente: <https://www.scalian-spain.es/la-revolucion-de-la-inteligencia-artificial-el-poder-transformador-de-los-modelos-de-lenguaje-a-gran-escala-llm/>



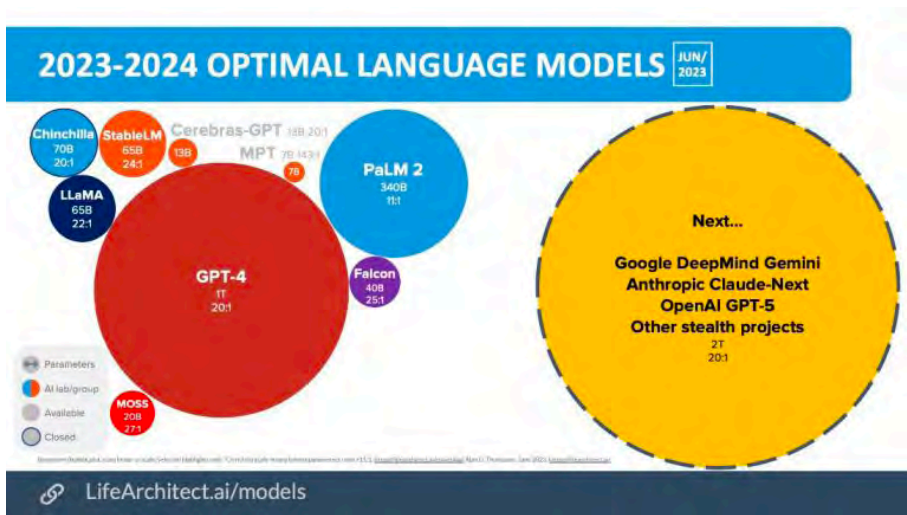


Figura 31: Modelos de lenguaje 2024

Fuente: <https://www.scalian-spain.es/la-revolucion-de-la-inteligencia-artificial-el-poder-transformador-de-los-modelos-de-lenguaje-a-gran-escala-llm/>

### 3.5.2. Entrenamiento de un modelo de lenguaje

El entrenamiento de un modelo de lenguaje depende de muchos factores, pero el principal es identificar la tarea que va a resolver: no se entrena igual un modelo de lenguaje para la traducción automática que para la generación de lenguaje o la extracción de información, entre otras.

De forma genérica, los pasos en el entrenamiento de un modelo de lenguaje son:

1. Recopilación de datos de texto: el aprendizaje de un modelo de lenguaje empieza con disponer de una colección de datos lo suficientemente grande para que el modelo aprenda. Se estima que, en función del número de parámetros que se deseen ajustar, se necesita tres veces la cantidad de texto para que el sistema aprenda. Estos datos pueden provenir de libros, sitios web, artículos o plataformas

de redes sociales o de cualquier fuente que proporcione información referente al problema a resolver. El objetivo es poder mostrar a la red la diversidad del lenguaje humano. En la mayoría de los casos es necesario complementar esta recopilación de datos con la generación de datos sintéticos. Esto se hace necesario en dos supuestos: el primero, cuando hay una escasez de datos reales, y el segundo, cuando es necesario balancear los datos para no generar sesgos.

2. Limpieza de los datos: los datos que se obtienen están llenos de caracteres o marcas que no aportan información al modelo, más bien producen ruido, y debe realizarse en un proceso llamado preprocesamiento que limpie los datos a aportar al modelo. Esto incluye tareas como eliminar caracteres no deseados, dividir el texto en partes más pequeñas, llamadas tokens, y ponerlo todo en un formato con el que el modelo pueda funcionar.
3. División de los datos: a continuación, los datos limpios se dividen en dos conjuntos. Un conjunto, de datos de entrenamiento que se usará para entrenar el modelo, y un conjunto de datos de validación, que se usarán más adelante para probar el rendimiento del modelo.
4. Configuración del modelo: luego se define la estructura del LLM, conocida como arquitectura. Esto implica seleccionar el tipo de red neuronal y decidir sobre varios parámetros, como la cantidad de capas y unidades ocultas dentro de la red. Esto es más importante de lo que parece, dado que de ello dependerá la cantidad de datos necesarios para que el modelo converja y, por otro lado, la necesidad de cómputo necesario.
5. Entrenamiento del modelo: el modelo LLM aprende con datos de entrenamiento procesando enormes corpus de texto, haciendo predicciones basadas en lo que ha aprendido hasta el momento. Realiza un aprendizaje no supervisado. Predice la próxima palabra en una secuencia, aprendiendo patrones de lenguaje, hechos e, in-

cluso, algunas habilidades de razonamiento. Una vez realizado el entrenamiento inicial se produce lo que se conoce como ajuste de instrucciones (instruction tuning): se van ajustando sus parámetros internos para reducir la diferencia entre sus predicciones y los datos reales, en función de la red y la tarea a desarrollar (por ejemplo, traducción o resumen) con datos etiquetados. Este proceso de ajuste de instrucciones personaliza el modelo para que funcione mejor en esas tareas.

6. Comprobación del modelo: el aprendizaje del modelo LLM se comprueba utilizando los datos de validación. Esto ayuda a ver qué tan bien está funcionando el modelo y a ajustar la configuración del modelo para un mejor rendimiento.
7. Usando el modelo: una vez el modelo ha aprendido está listo para usarse. En este punto pueden ocurrir varias cosas: disponer de un modelo “congelado” que no evoluciona y siempre actúa igual, o de un modelo “vivo”, que mejora o modifica su comportamiento en función de la interacción con el usuario. En este segundo caso, el sistema aprende a base de las apreciaciones o preguntas adicionales que el usuario hace a las respuestas del LLM. Es lo que se conoce como *aprendizaje por refuerzo*. Estos modelos se pueden integrar en aplicaciones o sistemas donde generarán texto en función de las nuevas entradas que se les proporcionen.
8. Mejorando el modelo: por último, siempre hay margen de mejora. El modelo LLM se puede perfeccionar aún más con el tiempo, utilizando datos actualizados o ajustando la configuración en función de los comentarios y el uso en el mundo real. La mejora de un LLM es un proceso iterativo y multifacético que involucra técnicas desde el preentrenamiento en grandes corpus de datos hasta la aplicación de sofisticados mecanismos de atención y técnicas de optimización. Cada una de estas técnicas contribuye a hacer que

los modelos sean más precisos, eficientes y capaces de manejar una amplia gama de tareas de procesamiento de lenguaje natural. Algunas de estas técnicas son:

a) Aumento de datos (Data Augmentation)

- **Parafraasis:** generar múltiples versiones de una oración o párrafo para incrementar la diversidad de los datos de entrenamiento.
- **Traducción y retrotraducción:** traducir el texto a otro idioma y luego volver a traducirlo al idioma original para generar variaciones semánticamente similares.

b) Preentrenamiento y ajuste fino (Pretraining and Fine-Tuning)

- **Ajuste fino en tareas específicas:** después del preentrenamiento, ajustar el modelo en conjuntos de datos específicos de tareas (por ejemplo, clasificación de texto; respuesta a preguntas) para mejorar el rendimiento en aplicaciones particulares.

c) Regularización y técnicas de optimización

- **Dropout:** desactivar aleatoriamente unidades de la red durante el entrenamiento para prevenir el sobreajuste.
- **Regularización L2:** añadir un término de penalización basado en los pesos de la red para controlar su magnitud y prevenir el sobreajuste.
- **Técnicas de optimización avanzadas:** utilizar algoritmos como Adam, LAMB o AdaBelief que adaptan las tasas de aprendizaje durante el entrenamiento para mejorar la convergencia.

#### d) Mecanismos de atención (Attention Mechanisms)

- Atención Auto-regresiva: implementar mecanismos de atención que permiten al modelo centrarse en diferentes partes del input para generar cada token.
- Multi-cabeza de atención: utilizar múltiples cabezas de atención para capturar diferentes aspectos de la información contextual en cada paso de procesamiento.

#### e) Escalabilidad y paralelización

- Paralelización de modelo y datos: dividir el modelo y los datos de entrenamiento entre múltiples GPUs o TPUs para entrenar más rápido y manejar modelos más grandes.
- Cuantización y Pruning: reducir la precisión de los pesos y eliminar conexiones innecesarias en la red para hacerla más eficiente sin sacrificar mucho rendimiento.

#### f) Incorporación de conocimiento externo

- Knowledge Distillation: transferir el conocimiento de un modelo grande (profesor) a uno más pequeño (estudiante) para hacer el modelo más eficiente.
- Inyección de conocimiento: utilizar bases de datos de conocimiento externo para enriquecer las representaciones internas del modelo.

#### g) Técnicas de interacción y aprendizaje activo

- Aprendizaje por refuerzo: utilizar técnicas de aprendizaje por refuerzo, como el Proximal Policy Optimization (PPO), para mejorar las respuestas generadas por el modelo en interacciones en tiempo real.

- **Feedback Humano:** incorporar retroalimentación de usuarios humanos para ajustar las salidas del modelo y mejorar su precisión y relevancia.

#### h) Ensamblaje de modelos (Model Ensembling)

- **Ensamblaje de modelos:** combinar múltiples instancias del modelo o diferentes tipos de modelos para mejorar la robustez y la precisión de las predicciones.

#### i) Evaluación y monitoreo continuo

- **Evaluación continua:** implementar un sistema de evaluación continua para monitorear el rendimiento del modelo y detectar posibles degradaciones.
- **Retraining periodico:** realizar entrenamientos periódicos del modelo con datos nuevos y actualizados para mantener su relevancia y precisión.

¿El aprendizaje o entrenamiento de un LLM es supervisado o no supervisado?

Si al modelo de lenguaje le damos solo corpus sin etiquetar parece que es un sistema no supervisado (text to text); pero la realidad, como hemos dicho al principio, es que el aprendizaje o entrenamiento depende de la tarea a resolver. Aunque no olvidemos que hay sistemas generalistas y se realiza el ajuste de instrucciones (instruction tuning). Entoces, en ese momento se le está enseñando con un conjunto de entrada y salida esperada.

Una vez que un LLM ha sido preentrenado, se pueden emplear varias técnicas para enseñarle a resolver tareas específicas. El ajuste fino es el método más común, pero también se utilizan técnicas como el aprendizaje por transferencia, meta-aprendizaje, prompt engineering y aprendizaje por refuerzo. Cada una de estas técnicas aprovecha el conocimiento

general adquirido por el LLM durante el preentrenamiento y lo adapta a aplicaciones específicas, mejorando su eficacia y utilidad en una amplia gama de tareas.

Estas técnicas se centran en ajustar y especializar el modelo para que sea eficaz en tareas concretas. Algunos de los métodos más comunes:

### **3.5.2.1. Ajuste fino (Fine-Tuning)**

El ajuste fino es el proceso más común para especializar un LLM en tareas específicas. Consiste en entrenar el modelo preentrenado en un conjunto de datos etiquetados para la tarea en cuestión. El proceso es el siguiente:

- a) Seleccionar el conjunto de datos: recopilar un conjunto de datos relevante para la tarea específica (por ejemplo, clasificación de texto, traducción automática o respuesta a preguntas).
- b) Configurar el modelo: cargar el LLM preentrenado y adaptar la arquitectura si es necesario (por ejemplo, añadir una capa de clasificación para tareas de clasificación).
- c) Entrenamiento: entrenar el modelo en el conjunto de datos específico durante varias épocas, ajustando hiperparámetros como la tasa de aprendizaje.
- d) Validación y evaluación: validar el rendimiento del modelo en un conjunto de datos de validación y ajustar según sea necesario.

### **3.5.2.2. Aprendizaje por transferencia (Transfer Learning)**

El aprendizaje por transferencia aprovecha el conocimiento adquirido por el LLM en una tarea anterior para mejorar el rendimiento en una nueva tarea. El proceso es:

- **Preentrenamiento general:** el modelo se preentrena en una tarea o conjunto de datos general.
- **Transferencia a la nueva tarea:** se transfiere el modelo preentrenado a la nueva tarea y se realiza un ajuste fino utilizando un conjunto de datos específico para la nueva tarea.

### 3.5.2.3. Meta-aprendizaje (Meta-Learning)

El meta-aprendizaje, también conocido como “aprendizaje para aprender”, enseña al modelo a adaptarse rápidamente a nuevas tareas con pocos ejemplos. El proceso es:

- **Entrenamiento meta:** entrenar al modelo en una variedad de tareas diferentes para que pueda aprender patrones generales de aprendizaje.
- **Adaptación rápida:** cuando se presenta una nueva tarea, el modelo puede adaptarse rápidamente con solo unos pocos ejemplos de la nueva tarea.

### 3.5.2.4. Prompt engineering

El prompt engineering es una técnica en la que se diseñan cuidadosamente las entradas al modelo para guiar su comportamiento y obtener respuestas deseadas. El proceso es:

- **Diseño del prompt:** crear prompts que proporcionen contexto y guía claros al modelo para la tarea específica.
- **Evaluación y ajuste:** probar diferentes formulaciones de prompts y ajustar según sea necesario para mejorar la calidad de las respuestas.



### 3.5.2.5. Aprendizaje por refuerzo (Reinforcement Learning)

En tareas interactivas o secuenciales, se puede utilizar el aprendizaje por refuerzo para entrenar al modelo a tomar decisiones que maximicen una recompensa a largo plazo. El proceso es:

- Definición del entorno y recompensas: definir el entorno en el que el modelo operará y las recompensas asociadas con las acciones correctas.
- Entrenamiento: utilizar algoritmos de aprendizaje por refuerzo para entrenar al modelo a tomar decisiones que maximicen la recompensa acumulada.

### 3.5.2.6. Incorporación de conocimiento adicional

En algunos casos, se puede incorporar conocimiento adicional al modelo para mejorar su rendimiento en tareas específicas. Las técnicas son:

- Inyección de conocimiento: incorporar información adicional, como bases de datos de conocimiento o gráficos de conocimiento, para enriquecer las representaciones del modelo.
- Fine-Tuning con datos externos: utilizar conjuntos de datos adicionales que contengan información relevante para la tarea específica durante el ajuste fino.

## 3.6. Hugging Face

Hugging Face es una empresa y comunidad líder en el campo del procesamiento del lenguaje natural (PLN) y el aprendizaje automático. Fundada en 2016, Hugging Face ha emergido como una plataforma clave en el desarrollo y la democratización de modelos de inteligencia artificial (IA) para tareas relacionadas con el lenguaje. Su enfoque principal es proporcionar herramientas y recursos accesibles que faciliten el trabajo con modelos de aprendizaje automático, especialmente en el ámbito del PLN.

Las principales herramientas de las que dispone son:

1. Biblioteca *Transformers*. Uno de los productos más destacados, proporciona una interfaz fácil de usar para trabajar con modelos de lenguaje basados en la arquitectura Transformer, como BERT, GPT, RoBERTa y T5. Esta biblioteca ofrece preentrenamiento y ajuste fino de modelos, permitiendo a los investigadores y desarrolladores implementar modelos de última generación con mínima configuración. Los modelos disponibles en Transformers son utilizados para tareas como clasificación de texto, generación de lenguaje, traducción automática y respuesta a preguntas.
2. Biblioteca Datasets Library. Ofrece acceso a una amplia variedad de conjuntos de datos de PLN. Esta biblioteca proporciona una interfaz eficiente para descargar, procesar y trabajar con datos en múltiples formatos, facilitando la preparación de datos para entrenamiento y evaluación de modelos. Datasets incluye recursos para tareas como análisis de sentimientos, traducción automática, y reconocimiento de entidades nombradas, entre otros.
3. Model Hub. Es una plataforma centralizada donde los investigadores y desarrolladores pueden compartir y acceder a modelos pre-

trenados. Este repositorio contiene una gran cantidad de modelos desarrollados y subidos por la comunidad, cubriendo una amplia gama de lenguajes y tareas. El Model Hub no solo facilita el acceso a modelos de alta calidad, sino que también permite a los usuarios contribuir con sus propios modelos, promoviendo la colaboración y el avance en el campo del PLN.

4. Spaces. Es una plataforma para construir y compartir aplicaciones de IA interactivas y demostraciones. Utilizando la infraestructura de Hugging Face, los desarrolladores pueden crear aplicaciones web que muestran cómo funcionan los modelos de PLN en tiempo real. Spaces permite a los usuarios explorar modelos y datos en un entorno interactivo, lo que facilita la experimentación y el desarrollo de nuevas ideas.

Hugging Face ha tenido un impacto significativo en la comunidad de PLN, proporcionando herramientas que simplifican el acceso a tecnología avanzada y fomentan la colaboración. La empresa ha contribuido a la popularización de los modelos Transformer y ha facilitado su adopción tanto en la investigación académica como en aplicaciones industriales.

Además, Hugging Face ha cultivado una comunidad activa de investigadores, desarrolladores y entusiastas del PLN a través de sus foros, conferencias y talleres. La filosofía de código abierto de la empresa ha permitido que los desarrolladores construyan sobre la base de modelos preexistentes y contribuyan a la mejora continua de la tecnología.

La visión de Hugging Face es seguir liderando el avance en el campo del PLN y la IA al hacer que las tecnologías más avanzadas sean accesibles para todos. La empresa continúa desarrollando nuevas herramientas y recursos que apoyan tanto a investigadores como a profesionales de la industria, con el objetivo de democratizar el acceso a la inteligencia artificial y promover la innovación.

En resumen, Hugging Face se ha consolidado como un pilar fundamental en el ecosistema del procesamiento del lenguaje natural, ofreciendo herramientas accesibles, recursos valiosos y una plataforma para la colaboración y el desarrollo continuo en el campo de la inteligencia artificial.

## 3.7. Retos de la inteligencia artificial

El acceso a los modelos de lenguaje está al alcance de la mano de cualquiera, son muchos los modelos que se pueden usar hoy en día. Por lo que se debe tener en cuenta una serie de cuestiones tanto de fiabilidad como sociales o éticas. Los modelos de lenguaje presentan una serie de retos y desafíos sociales significativos que deben abordarse para asegurar su uso responsable y beneficioso.

Actualmente hay abierta una serie de desafíos o retos a resolver como son: la mitigación de desigualdades, la privacidad de los datos, la explicabilidad de la IA, el uso ético de la IA. El Reglamento de Inteligencia Artificial de la Unión Europea apunta, entre las características que deben tener estas tecnologías, la de ser lo suficientemente explicables para generar confianza en sus resultados. La explicabilidad de la IA es crucial para generar confianza en la toma de decisiones y para identificar y abordar posibles sesgos y errores. Existen diversas técnicas para abordar la explicabilidad, como por ejemplo los mecanismos de atención, que permiten entender y explicar las decisiones y predicciones de los modelos, lo que es fundamental para la responsabilidad y la transparencia. Les adelanto que no va a ser tarea fácil igual que no lo es explicar el porqué de las actuaciones del cerebro humano.

Los riesgos relativos a la privacidad son la recopilación de datos biométricos y de salud, la falta de control y transparencia que puedan ocasionar problemas de ciberseguridad o de suplantación. Hoy en día, es sencillo para un sistema suplantar la voz de una persona simplemente teniendo una grabación de esa persona. Para evitar esto se deben implementar medidas de seguridad y protección de datos, garantizar la transparencia. Hay que desarrollar mecanismos para otorgar a las personas control sobre sus datos personales y desarrollar directrices éticas para el desarrollo y uso de la IA.

La IA presenta oportunidades emocionantes para la creación de contenido, pero también plantea desafíos complejos en cuanto a la propiedad intelectual y los derechos de autor. Es crucial establecer un marco legal claro y equilibrado que promueva la innovación responsable y proteja los derechos de todas las partes involucradas. Dejo para discusión ¿quién es el propietario de un contenido generado con IA que se ha basado en contenido de diversos autores? Seguramente, por un lado, habrá personas que argumenten que al igual que los seres humanos, se nutren de mucho conocimiento y crean obras originales, mientras que otros pensarán que su trabajo se basa en lo que han aprendido de otros, y por lo tanto, la propiedad pertenece a todos

Es fundamental garantizar que los sistemas de IA sean justos, éticos, transparentes y respetuosos con la privacidad de los usuarios, y que se implementen salvaguardas adecuadas para mitigar cualquier riesgo potencial de sesgo o discriminación. En esta línea y teniendo en cuenta el marco normativo en España y Europa, se están desarrollando actuaciones como la puesta en marcha de la Estrategia Nacional de Inteligencia Artificial de 2024, en la que se ha constituido un comité de gobernanza en el que participa, entre otros, los principales agentes de IA, como los representantes de Proyecto ILENIA anteriormente mencionados, las academias de las lenguas co-oficiales de España: por cierto, también participa la Fundación para la Transformación de La Rioja. Además, recientemente, se ha puesto en marcha la Agencia Española de Supervisión e Inteligencia Artificial (AESIA), cuyos objetivos son la promoción del uso responsable, sostenible y confiable de la inteligencia artificial; la concienciación, divulgación y promoción de la formación en materia de inteligencia artificial.

El enorme crecimiento de la IA ha introducido una serie de cambios en las organizaciones y empresas. Los miedos a estos cambios y con la enorme capacidad de automatización de tareas y trabajos ha supuesto el crecimiento de lo que permítanme que denomine “IAfobia”. Por ejemplo, muchos docentes se oponen a la utilización de la IA en el entorno

educativo. Desde mi punto de vista es como si en su momento se hubiera prohibido el uso de las calculadoras o de los ordenadores en el aula. Lo que se debe hacer es aprovechar esta utilidad en el proceso educativo integrándolo convenientemente.

Otro de los aspectos negativos que ha surgido en la sociedad respecto a la IA es la preocupación por la cantidad de puestos de trabajo que podría eliminar. Yo diría que, más bien, la IA transformará y creará nuevos empleos; obviamente, algunos desaparecerán, como ha ocurrido en todas las revoluciones industriales, donde las tareas manuales o repetitivas han sido reemplazadas por actividades más inteligentes.

A continuación, se detallan algunos de los principales desafíos:

### **1. Sesgo y Equidad.**

Los principales retos que resolver son:

- **Discriminación y Prejuicios:** Los LLM pueden perpetuar y amplificar sesgos presentes en los datos con los que fueron entrenados, llevando a resultados discriminatorios en aplicaciones como la contratación de personal, decisiones judiciales, y recomendaciones de productos.
- **Representación Injusta:** Grupos minoritarios o menos representados en los datos de entrenamiento pueden verse perjudicados por modelos que no capturan adecuadamente sus características y necesidades.

La mitigación de desigualdades está estrechamente relacionada con los sesgos existentes en la IA. Se refiere a la presencia de prejuicios sistemáticos en los resultados generados por estos sistemas como consecuencia de haber sido entrenados con datos “incompletos” o “no balanceados”. Estos prejuicios pueden manifestarse de diversas maneras, incluyendo por ejemplo la discriminación (racial, de género, por edad), exclusión, daños a la reputación e injusticia social. El sis-

tema COMPAS, sistema de justicia penal, utilizado en Estados Unidos para predecir la reincidencia criminal, discriminaba a las personas negras. El sistema tenía más probabilidades de predecir erróneamente que las personas negras reincidirían, lo que resultó en penas más severas para ellas.



*Figura 32: Imagen del sistema COMPAS para predecir la reincidencia de presos*

Las estrategias para abordar el problema de los sesgos se centran en utilizar conjuntos de datos diversos y representativos, promover la transparencia, balanceo de los ejemplos de todas las categorías.

A continuación, se describen varias estrategias y prácticas recomendadas para abordar este problema:

a) Diversidad en los Datos de Entrenamiento

- **Recolección de Datos Diversos:** Asegurar que los conjuntos de datos utilizados para entrenar los modelos sean representativos de la diversidad de la población a la que se aplicarán, incluyendo variaciones en género, raza, edad, idioma, y contexto socioeconómico.
- **Auditorías de Datos:** Realizar auditorías regulares de los datos para identificar y corregir sesgos. Esto incluye la eliminación



de datos duplicados y la inclusión de datos de grupos subrepresentados. La duplicidad o desbalanceo de los modelos puede ocurrir cuando a un modelo de lenguaje se le realiza técnicas de aprendizaje continuo o cualquier otra técnica para instruirlo en otra tarea en la que el corpus usado pueda ya estar presente en el modelo original.

#### b) Preprocesamiento de Datos

- Normalización de Datos: Implementar técnicas de normalización para reducir el impacto de características que pueden introducir sesgos.
- Anotación Imparcial: Asegurar que los procesos de anotación de datos sean imparciales y revisados por múltiples anotadores para minimizar la introducción de sesgos subjetivos.

#### c) Diseño y Entrenamiento del Modelo

- Algoritmos de Equidad: Utilizar algoritmos y técnicas de aprendizaje que tienen en cuenta la equidad, como el aprendizaje adversarial para mitigar el sesgo.
- Regularización del Sesgo: Implementar términos de regularización que penalicen el sesgo en las predicciones del modelo durante el entrenamiento.
- Evaluación Continua: Evaluar continuamente el modelo con métricas de equidad para identificar y corregir sesgos en las predicciones.

#### d) Transparencia y Explicabilidad

- Modelos Interpretables: Desarrollar y utilizar modelos que sean más interpretables y transparentes, facilitando la identificación de sesgos.

- **Explicabilidad de Decisiones:** Proporcionar explicaciones claras y comprensibles de las decisiones tomadas por el modelo para aumentar la transparencia y la confianza.

#### e) Gobernanza y Regulación

- **Políticas de Equidad:** Establecer políticas y marcos regulatorios que promuevan la equidad en el desarrollo y uso de IA.
- **Auditorías Externas:** Realizar auditorías externas independientes para evaluar el sesgo y la equidad en los sistemas de IA.

#### f) Participación Comunitaria

- **Inclusión de Diversas Perspectivas:** Involucrar a comunidades diversas y a partes interesadas en el proceso de desarrollo de IA para asegurar que se consideren múltiples perspectivas y necesidades.
- **Educación y Sensibilización:** Promover la educación y la sensibilización sobre los problemas de sesgo y equidad en la IA entre desarrolladores, investigadores y usuarios finales.

#### g) Mantenimiento y Actualización Continua

- **Actualización de Modelos:** Actualizar regularmente los modelos con nuevos datos y técnicas para asegurar que continúen siendo equitativos y relevantes.
- **Monitoreo de Desempeño:** Implementar sistemas de monitoreo continuo para detectar y corregir sesgos que puedan surgir a medida que cambian los datos y el entorno.

#### h) Evaluación y Mitigación de Sesgo

- **Herramientas de Evaluación:** Utilizar herramientas y frameworks específicos para la evaluación del sesgo en los modelos de IA, como Fairness Indicators de Google o AI Fairness 360 de IBM.

- **Técnicas de Mitigación:** Aplicar técnicas específicas para mitigar el sesgo, como la reponderación de datos, la modificación de etiquetas, y el ajuste de umbrales de decisión.

En este punto, se podría abrir un debate que dejo para otra ocasión: Si la IA intenta simular la inteligencia humana, ¿no estamos las personas también sesgadas? En realidad, ningún conjunto de datos reales está completamente equilibrado. Es como si se pretendiese que para aprender a pronosticar el tiempo de Londres se deba tener los mismos datos de días de lluvia que de sol. Como ejemplo del excesivo intento de mitigar los sesgos está el caso del modelo de lenguaje GEMINI de Google; cuando se presentó el sistema, si le decías que te generara una imagen de un general nazi te generaba un general de raza afroamericana. Obviamente, se pensó que el sistema funciona mal, lo que realmente pasaba es que las reglas para mitigar los sesgos no tenían en cuenta que en este caso “histórico” no se debía equilibrar los ejemplos. El sistema fue bloqueado durante un tiempo hasta arreglar este problema.

Tener fe ciega en la IA tiene un peligro si no se aborda el problema de caja negra de los algoritmos. La IA genera contenido, pero no se sabe cómo está configurada, cómo ha sido entrenada, nos proporciona una respuesta, pero ¿cómo ha llegado a esa conclusión? Se desconoce. ¿hay que fiarse de un sistema que decide algo que pueda poner en riesgo nuestras vidas sin al menos una explicación de los motivos que ha llevado al sistema a tomar esas decisiones?

## **2. Privacidad y Seguridad**

Abordar la privacidad y seguridad en los modelos de lenguaje es crucial debido a la cantidad de datos que procesan y la variedad de aplicaciones en las que se utilizan. Los principales retos que resolver son:

- **Fuga de Datos Sensibles:** LLM pueden memorizar y revelar información sensible presente en los datos de entrenamiento, comprometiendo la privacidad de individuos.

- **Manipulación y Abuso:** Los LLM pueden ser utilizados para generar desinformación, propaganda, y realizar ataques de ingeniería social, como phishing.

Algunas estrategias para enfrentar estos desafíos son:

a) Privacidad

- **Anonimización de Datos:** Antes de usar los datos para entrenar los LLM, eliminar o modificar información que pueda identificar a las personas, como nombres, direcciones y números de teléfono.
- **Técnicas de Desidentificación:** Utilizar técnicas avanzadas de desidentificación para proteger la privacidad de los datos.
- **Entrenamiento Descentralizado:** En lugar de centralizar los datos, usar el aprendizaje federado, que permite a los modelos entrenarse en dispositivos locales y solo compartir parámetros del modelo, no los datos crudos.
- **Seguridad de los Datos en el Dispositivo:** Asegurar que los datos se procesen y queden en los dispositivos de los usuarios para minimizar los riesgos de privacidad.
- **Perturbación Controlada:** Introducir ruido en los datos antes de su procesamiento para garantizar que los resultados agregados no puedan ser usados para identificar información específica sobre los individuos.
- **Privacidad Garantizada:** Implementar técnicas de privacidad diferencial para mantener la utilidad del modelo mientras se protegen los datos sensibles.

b) Seguridad

- **Cifrado en Tránsito y en Reposo:** Asegurar que todos los datos

utilizados por los LLM estén cifrados tanto cuando se transmiten como cuando están almacenados.

- **Cifrado de Parámetros del Modelo:** Proteger los parámetros del modelo mediante cifrado para evitar accesos no autorizados.
- **Autenticación y Autorización:** Implementar estrictos mecanismos de autenticación y autorización para controlar quién puede acceder y modificar los LLM.
- **Roles y Permisos:** Utilizar un sistema de roles y permisos para limitar el acceso a los datos y al modelo solo a personal autorizado.
- **Registro de Actividades:** Mantener un registro detallado de todas las actividades relacionadas con los LLM, incluyendo accesos y modificaciones.
- **Auditorías Regulares:** Realizar auditorías periódicas para asegurar que se cumplen las reglas establecidas.
- **Evaluación de Vulnerabilidades:** Llevar a cabo evaluaciones regulares de vulnerabilidades y pruebas de penetración para identificar y corregir debilidades en la seguridad del sistema.
- **Respuesta a Incidentes:** Establecer un plan de respuesta a incidentes para gestionar rápidamente cualquier brecha de seguridad.

### c) Regulación y Cumplimiento

- **Cumplimiento Normativo:** Asegurarse de que el uso de los LLM cumpla con las leyes y regulaciones de privacidad y seguridad de datos, como el GDPR o reglamentos de IA en cualquier parte del mundo. En este sentido España y Europa están a la vanguardia frente a la regulación en EEUU.
- **Políticas Internas:** Desarrollar y seguir políticas internas que

reflejen las mejores prácticas y requisitos legales.

- **Informar a los Usuarios:** Ser transparente con los usuarios sobre cómo se recopilan, usan y protegen sus datos.
- **Responsabilidad del Uso de Datos:** Asegurar que se utilicen los datos de manera responsable y ética.

### **3. Transparencia y Explicabilidad**

Abordar la transparencia y la explicabilidad en los modelos de lenguaje es fundamental para garantizar que sus decisiones y procesos sean comprensibles y confiables para los usuarios

Los principales retos que resolver son:

- **Caja Negra:** Los LLM son a menudo complejos y opacos, lo que dificulta entender cómo toman decisiones y por qué generan ciertas respuestas.
- **Responsabilidad:** La falta de transparencia hace difícil asignar responsabilidad en casos de mal uso o errores significativos.

Algunas estrategias para enfrentar estos desafíos:

#### a) Transparencia

- **Documentación Detallada:** Proporcionar documentación clara y detallada sobre la arquitectura del modelo, incluyendo la estructura de capas, los tipos de nodos y la forma en que se entrenan. Describir los conjuntos de datos utilizados para entrenar el modelo, incluyendo la procedencia, el tamaño y las características de los datos.
- **Tarjetas de Modelo:** Crear “model cards” que resuman la información clave sobre el modelo, como su propósito, los datos de entrena-

miento, las métricas de rendimiento y las limitaciones conocidas. Incluir ejemplos específicos de cómo se puede utilizar el modelo, junto con escenarios en los que podría fallar o no ser adecuado.

- **Procedimientos de Evaluación:** Implementar y publicar los resultados de pruebas de sesgo para identificar y mitigar posibles sesgos en el modelo. Permitir auditorías externas para evaluar la transparencia y la fiabilidad del modelo.

#### b) Explicabilidad

- **Métodos de Interpretación del Modelo:** Utilizar técnicas como SHAP (SHapley Additive exPlanations) y LIME (Local Interpretable Model-agnostic Explanations) para explicar las predicciones del modelo a nivel local y global.
- **Visualización de Atención:** En modelos basados en Transformers, usar mapas de atención para mostrar qué partes del texto influyen más en las decisiones del modelo.
- **Descomposición del Modelo:** Descomponer el modelo en componentes más simples que pueden ser explicados individualmente. Generar explicaciones en lenguaje natural que describan cómo el modelo llegó a una determinada conclusión.
- **Interfaces de Usuario Explicativas:** Crear interfaces de usuario que muestren no solo los resultados, sino también la confianza del modelo y las características más importantes que influyeron en la decisión. Permitir que los usuarios proporcionen feedback sobre las explicaciones y ajusten las interpretaciones basadas en este feedback.

#### c) Colaboración y Ética

- **Involucrar a Expertos Multidisciplinarios:** Involucrar a expertos en ética, leyes, ciencias sociales y tecnología para abordar los problemas de transparencia y explicabilidad de manera integral.

- **Revisión Ética:** Establecer comités de revisión ética para evaluar y supervisar el desarrollo y la implementación de los LLM.
- **Comunicación Clara:** Comunicar de manera clara y accesible, evitando tecnicismos innecesarios, para que los usuarios no técnicos también puedan comprender cómo funciona el modelo. Proporcionar recursos educativos para ayudar a los usuarios a entender los fundamentos de los LLM y cómo interactuar con ellos de manera efectiva.

La transparencia y la explicabilidad en los LLM son esenciales para ganar la confianza de los usuarios y asegurar que los modelos sean utilizados de manera ética y responsable. La combinación de documentación detallada, técnicas de interpretación del modelo, interfaces de usuario explicativas y la colaboración con expertos multidisciplinarios puede ayudar a abordar estos desafíos de manera efectiva.

#### **4. Impacto Laboral y Económico**

Abordar el impacto laboral y económico de los modelos de lenguaje es crucial para asegurar que la adopción de estas tecnologías beneficie a la sociedad en general y no solo a unas pocas entidades. Los desafíos a los que se enfrenta son:

- **Desplazamiento de Empleos:** La automatización impulsada por LLM puede desplazar empleos en sectores como el servicio al cliente, redacción de contenido, y traducción.
- **Desigualdad Económica:** La implementación de LLM puede beneficiar desproporcionadamente a grandes empresas tecnológicas, aumentando la brecha económica y concentrando el poder en pocas manos.

Algunas estrategias para abordar estos desafíos son:

a) Impacto Laboral



### 1. Reentrenamiento y Recualificación

- **Programas de Capacitación:** Implementar programas de capacitación y reentrenamiento para ayudar a los trabajadores a adquirir nuevas habilidades relevantes para la era de la IA.
- **Educación Continua:** Ofrecer oportunidades de educación continua y certificaciones en áreas como el procesamiento del lenguaje natural (NLP), la inteligencia artificial (IA) y la ciencia de datos.

### 2. Nuevas Oportunidades de Empleo

- **Nuevos Roles:** Crear nuevos roles que surjan a medida que se desarrollan y adoptan los LLM, como ingenieros de IA, analistas de datos y especialistas en ética de la IA.
- **Apoyo a la Innovación:** Fomentar la innovación y el emprendimiento en el campo de la IA para generar nuevas oportunidades de empleo.

### 3. Adaptación Organizacional

- **Cambio de Funciones:** Redefinir las funciones laborales para aprovechar las capacidades de los LLM, permitiendo a los empleados centrarse en tareas más estratégicas y creativas.
- **Asistencia en el Trabajo:** Utilizar LLM como herramientas de asistencia para aumentar la productividad y mejorar la toma de decisiones en lugar de reemplazar a los trabajadores.

## b) Impacto Económico

### 1. Equidad en la Distribución de Beneficios

- **Políticas de Inclusión:** Desarrollar políticas que aseguren que los beneficios económicos de los LLM se distribuyan de manera equitativa en toda la sociedad.

- Apoyo a Pequeñas y Medianas Empresas (PYMES): Proporcionar recursos y apoyo a las PYMES para que puedan adoptar y beneficiarse de los LLM.

## 2. Regulación y Supervisión

- Marco Regulatorio: Establecer marcos regulatorios claros que guíen el desarrollo y la implementación de los LLM de manera ética y responsable.
- Supervisión Continua: Implementar sistemas de supervisión continua para monitorear los impactos económicos y ajustar las políticas según sea necesario.

## 3. Inversión en Investigación y Desarrollo (I+D)

- Financiación Pública y Privada: Fomentar la inversión en I+D tanto del sector público como del privado para impulsar la innovación en el campo de los LLM.
- Colaboración Academia-Industria: Promover la colaboración entre instituciones académicas y la industria para desarrollar tecnologías de LLM que sean económicamente sostenibles y beneficiosas.

Abordar el impacto laboral y económico de los LLM requiere un enfoque multidisciplinar que incluya la recualificación de la fuerza laboral, la creación de nuevas oportunidades de empleo, la implementación de políticas inclusivas y equitativas, la regulación y supervisión adecuada, y la inversión en I+D. Estas estrategias ayudarán a asegurar que la adopción de los LLM sea beneficiosa para todos los sectores de la sociedad y contribuya al crecimiento económico sostenible.

## 5. Desinformación y Confianza

Abordar la desinformación y la confianza en los Modelos de lenguaje

es esencial para asegurar que estos sistemas sean utilizados de manera responsable y beneficiosa para la sociedad. Los desafíos a los que se enfrentan son:

- Generación de fake news: los LLM pueden ser utilizados para generar contenido falso o engañoso que es difícil de detectar y combatir.
- Erosión de la Confianza: La proliferación de desinformación generada por LLM puede erosionar la confianza pública en la información y en las instituciones.

Algunas de las estrategias para enfrentar estos desafíos son:

#### a) Abordando la Desinformación

##### 1. Mejora en la Calidad de los Datos de Entrenamiento

- Selección Cuidadosa de Datos: Asegurar que los datos utilizados para entrenar LLM provengan de fuentes fiables y verificadas.
- Filtrado de Datos: Implementar técnicas de filtrado de datos para eliminar información falsa, sesgada o de baja calidad antes de que se utilice en el entrenamiento.

##### 2. Implementación de Mecanismos de Verificación

- Fact-Checking Automatizado: Integrar herramientas de verificación automatizada de hechos que pueden verificar la veracidad de la información generada por los LLM.
- Referencia a Fuentes Confiables: Hacer que los LLM proporcionen referencias a fuentes confiables y verificadas cuando generen información.

##### 3. Capacitación del Usuario

- **Educación del Usuario:** Implementar programas de educación y formación para usuarios sobre cómo identificar y evitar la desinformación.
- **Transparencia de Limitaciones:** Informar a los usuarios sobre las limitaciones de los LLM, incluyendo la posibilidad de generar información incorrecta.

## b) Fomentando la Confianza

### 1. Transparencia y Explicabilidad

- **Modelos Transparentes:** Desarrollar modelos que puedan explicar sus decisiones y procesos internos de manera comprensible para los usuarios.
- **Auditorías Externas:** Realizar auditorías externas de los LLM para garantizar su transparencia y rendición de cuentas.

### 2. Gobernanza y Políticas

- **Normas Éticas:** Establecer y seguir normas éticas claras para el desarrollo y uso de LLM.
- **Políticas de Privacidad:** Implementar y hacer cumplir políticas de privacidad rigurosas para proteger los datos del usuario.

### 3. Evaluación y Monitoreo Continuos

- **Monitoreo en Tiempo Real:** Establecer sistemas de monitoreo en tiempo real para detectar y corregir rápidamente cualquier instancia de desinformación.
- **Evaluación Continua:** Realizar evaluaciones continuas del rendimiento y la precisión de los LLM para identificar áreas de mejora.

## c) Tecnologías y Herramientas Complementarias

### 1. Sistemas de Detección de Sesgos

- **Detección de Sesgos:** Utilizar herramientas que detecten y mitiguen sesgos en los datos y en la salida de los LLM.
- **Diversificación de Datos:** Asegurar que los datos de entrenamiento sean diversos y representen una amplia gama de perspectivas.

### 2. Integración con Sistemas de Fact-Checking Humanos

- **Colaboración con Fact-Checkers:** Colaborar con organizaciones y plataformas de fact-checking para revisar la información generada por los LLM.
- **Revisión Humana:** Incluir una capa de revisión humana para validar información crítica o sensible antes de su difusión.

Abordar la desinformación y fomentar la confianza en los LLM requiere un enfoque integral que combine la mejora de la calidad de los datos de entrenamiento, la implementación de mecanismos de verificación, la capacitación de los usuarios, la transparencia y explicabilidad de los modelos, y la gobernanza y políticas robustas. Además, la evaluación y monitoreo continuos, junto con la utilización de tecnologías complementarias, son cruciales para asegurar que los LLM operen de manera ética, precisa y confiable.

## 6. Control y Gobernanza

Abordar el control y la gobernanza en los modelos de lenguaje es fundamental para garantizar su uso responsable, ético y beneficioso. Los retos que se plantean son:

- **Regulación:** La falta de regulaciones claras y consistentes a nivel global crea incertidumbre y puede permitir el uso irresponsable de LLM.

- **Monopolio Tecnológico:** Grandes corporaciones tecnológicas pueden dominar el desarrollo y uso de LLM, limitando la competencia y la innovación abierta.

Algunas de las estrategias y prácticas clave para enfrentar estos desafíos:

#### a) Estrategias para Control y Gobernanza en LLM

##### 1. Desarrollo de Políticas y Normas

- **Normas Éticas y Directrices:** Establecer y seguir normas éticas claras y directrices para el desarrollo, despliegue y uso de LLM.
- **Regulación:** Colaborar con reguladores y legisladores para desarrollar marcos regulatorios que aseguren el uso ético y responsable de los LLM.

##### 2. Transparencia y Rendición de Cuentas

- **Transparencia en el Desarrollo:** Publicar detalles sobre los datos de entrenamiento, arquitecturas de modelos y procesos de desarrollo utilizados para crear los LLM.
- **Auditorías Externas:** Permitir auditorías independientes de los modelos y sus operaciones para asegurar la transparencia y la rendición de cuentas.

##### 3. Monitoreo y Evaluación Continuos

- **Monitoreo en Tiempo Real:** Implementar sistemas de monitoreo en tiempo real para detectar y mitigar comportamientos no deseados de los LLM.
- **Evaluación Regular:** Realizar evaluaciones periódicas del rendimiento, precisión y equidad de los LLM para identificar y corregir posibles problemas.

#### 4. Gestión de Sesgos y Equidad

- **Detección y Mitigación de Sesgos:** Utilizar herramientas y técnicas para detectar y mitigar sesgos en los datos de entrenamiento y en la salida de los LLM.
- **Diversidad en Datos:** Asegurar que los datos de entrenamiento sean diversos y representen una amplia gama de perspectivas y grupos demográficos.

#### 5. Control de Acceso y Uso

- **Control de Acceso:** Establecer controles estrictos sobre quién puede acceder y utilizar los LLM, incluyendo autenticación y autorización robustas.
- **Términos de Uso:** Definir y hacer cumplir términos de uso claros que describan las condiciones bajo las cuales los LLM pueden ser utilizados.

### b) Herramientas y Prácticas Complementarias

#### 1. Mecanismos de Responsabilidad

- **Comités de Ética:** Establecer comités de ética que supervisen el desarrollo y uso de los LLM, asegurando que se alineen con principios éticos y sociales.
- **Informes de Impacto Ético:** Requerir informes de impacto ético que evalúen cómo los LLM pueden afectar a diferentes grupos y contextos.

#### 2. Participación de la Comunidad

- **Participación de Stakeholders:** Involucrar a una variedad de stakeholders, incluyendo usuarios, reguladores, expertos en ética y representantes de comunidades afectadas, en el desarrollo y gobernanza de LLM.

- **Feedback y Mejora Continua:** Recopilar y actuar sobre el feedback de los usuarios y otros stakeholders para mejorar continuamente los LLM.

### 3. Tecnología de Supervisión

- **Supervisión Algorítmica:** Utilizar tecnologías de supervisión para rastrear y analizar las decisiones y salidas de los LLM, identificando y abordando comportamientos no deseados.
- **Regulación de Algoritmos:** Implementar regulaciones que requieran transparencia en los algoritmos y modelos utilizados, incluyendo explicabilidad y capacidad de auditoría.

El control y la gobernanza de los LLM requieren un enfoque multidimensional que combine políticas y normas claras, transparencia y rendición de cuentas, monitoreo y evaluación continuos, gestión de sesgos y equidad, y control de acceso y uso. La implementación de mecanismos de responsabilidad, la participación de la comunidad y el uso de tecnologías de supervisión son esenciales para garantizar que los LLM sean utilizados de manera ética, responsable y beneficiosa para la sociedad.

## 7. Ética y Autonomía

Abordar la ética y la autonomía en los Modelos de lenguaje es crucial para garantizar que estos modelos sean desarrollados y utilizados de manera responsable y alineada con los valores humanos. Los desafíos a los que se enfrenta son:

- **Decisiones Automatizadas:** La delegación de decisiones a LLM en áreas críticas, como la salud y la justicia, plantea cuestiones éticas sobre la autonomía humana y el control.
- **Manipulación del Comportamiento:** Los LLM pueden influir sutilmente en las decisiones y comportamientos humanos, planteando preguntas sobre la manipulación y la autonomía.



Algunas de las estrategias y prácticas clave para enfrentar estos desafíos son:

a) Estrategias para abordar la Ética y la Autonomía en LLM

1. Desarrollar Principios Éticos Claros

- **Directrices Éticas:** Establecer principios éticos claros y directrices para el desarrollo y uso de LLM. Estos principios deben abordar aspectos como la justicia, la no maleficencia, la beneficencia, la autonomía y la transparencia.
- **Código de Conducta:** Crear y aplicar un código de conducta para los desarrolladores y usuarios de LLM que enfatice la responsabilidad ética.

2. Incorporar la Ética en el Ciclo de Vida del Modelo

- **Diseño Ético:** Integrar consideraciones éticas en cada etapa del ciclo de vida del modelo, desde la recopilación de datos y el diseño del modelo hasta el despliegue y la evaluación.
- **Evaluaciones de Impacto Ético:** Realizar evaluaciones de impacto ético antes y después del despliegue de LLM para identificar y mitigar posibles daños.

3. Promover la Transparencia y la Explicabilidad

- **Modelos Explicables:** Desarrollar modelos que sean explicables, es decir, que puedan proporcionar justificaciones comprensibles para sus decisiones y salidas.
- **Transparencia en el Desarrollo:** Publicar información sobre los procesos de desarrollo, los datos de entrenamiento y los algoritmos utilizados para crear LLM.

4. Gestión de Sesgos y Equidad

- **Detección y Mitigación de Sesgos:** Utilizar herramientas y técnicas para detectar y mitigar sesgos en los datos de entrenamiento y en las salidas de los LLM.
- **Diversidad en Datos:** Asegurar que los datos de entrenamiento sean diversos y representen una amplia gama de perspectivas y grupos demográficos.

#### 5. Fomentar la Autonomía del Usuario

- **Control del Usuario:** Permitir que los usuarios tengan control sobre cómo se utilizan los LLM, incluyendo la capacidad de personalizar y ajustar las salidas del modelo.
- **Consentimiento Informado:** Asegurar que los usuarios comprendan cómo se recopilan, almacenan y utilizan sus datos, y obtener su consentimiento informado para estas prácticas.

#### b) Herramientas y Prácticas Complementarias

##### 1. Mecanismos de Responsabilidad

- **Comités de Ética:** Establecer comités de ética que supervisen el desarrollo y uso de los LLM, asegurando que se alineen con principios éticos y sociales.
- **Informes de Impacto Ético:** Requerir informes de impacto ético que evalúen cómo los LLM pueden afectar a diferentes grupos y contextos.

##### 2. Participación de la Comunidad

- **Participación de Stakeholders:** Involucrar a una variedad de stakeholders, incluyendo usuarios, reguladores, expertos en ética y representantes de comunidades afectadas, en el desarrollo y gobernanza de LLM.

- **Feedback y Mejora Continua:** Recopilar y actuar sobre el feedback de los usuarios y otros stakeholders para mejorar continuamente los LLM.

### 3. Supervisión y Auditoría

- **Auditorías Externas:** Permitir auditorías independientes de los modelos y sus operaciones para asegurar la transparencia y la rendición de cuentas.
- **Supervisión Algorítmica:** Utilizar tecnologías de supervisión para rastrear y analizar las decisiones y salidas de los LLM, identificando y abordando comportamientos no deseados.

Abordar la ética y la autonomía en los LLM requiere un enfoque proactivo y multidimensional que combine principios éticos claros, diseño ético, transparencia, gestión de sesgos, fomento de la autonomía del usuario y mecanismos de responsabilidad. La participación de la comunidad, la supervisión y la auditoría son esenciales para garantizar que los LLM sean desarrollados y utilizados de manera ética y responsable, beneficiando a la sociedad en su conjunto.

## 8. Acceso y democratización

Abordar el acceso y la democratización en los modelos de lenguaje implica garantizar que estas tecnologías sean accesibles y beneficiosas para una amplia gama de personas y comunidades, independientemente de sus recursos económicos, educativos o geográficos. Los principales desafíos son:

- **Brecha digital:** el acceso a la tecnología de LLM y sus beneficios puede estar limitado a países y comunidades con recursos, exacerbando las desigualdades existentes.
- **Inclusividad:** asegurar que los beneficios de los LLM sean accesibles a todos, incluyendo aquellos con discapacidades y barreras lingüísticas, es un desafío continuo.

Algunas estrategias y prácticas clave para lograr estos objetivos son:

a) Estrategias para abordar el acceso y la democratización en LLM

1. Acceso abierto y transparencia

- Modelos de código abierto: publicar modelos de LLM como código abierto para que investigadores, desarrolladores y organizaciones de todo el mundo puedan acceder, utilizar y mejorar estos modelos sin barreras económicas.
- Documentación detallada: proporcionar documentación completa y detallada sobre el diseño, el entrenamiento y el uso de los LLM para facilitar su comprensión y utilización por parte de una audiencia amplia.

2. Infraestructura y recursos asequibles

- Plataformas asequibles: ofrecer acceso a LLM a través de plataformas asequibles o gratuitas para desarrolladores, pequeñas empresas, organizaciones sin fines de lucro y comunidades académicas.
- Programas de becas y subsidios: crear programas de becas y subsidios para apoyar a investigadores y desarrolladores en regiones subrepresentadas o con menos recursos.

3. Educación y capacitación

- Cursos y tutoriales: desarrollar y distribuir cursos, tutoriales y materiales educativos gratuitos o de bajo costo sobre el uso y desarrollo de LLM.
- Capacitación en comunidades locales: realizar programas de capacitación y talleres en comunidades locales, especialmente en áreas con menos acceso a recursos tecnológicos.

#### 4. Colaboración internacional

- Alianzas globales: formar alianzas con instituciones académicas, gobiernos y organizaciones internacionales para promover la investigación y el desarrollo de LLM en todo el mundo.
- Redes de investigación: establecer redes de investigación que incluyan participantes de diversas regiones y contextos, fomentando la colaboración y el intercambio de conocimientos.

#### 5. Inclusión lingüística y cultural

- Soporte multilingüe: entrenar LLM en una amplia variedad de idiomas para asegurar que las comunidades no angloparlantes también se beneficien de estas tecnologías.
- Culturalmente sensibles: asegurar que los LLM sean culturalmente sensibles y respeten la diversidad cultural en sus respuestas y comportamientos.

#### 6. Infraestructura en la nube

- Acceso a la nube: ofrecer acceso a LLM a través de servicios en la nube, permitiendo a los usuarios con recursos limitados utilizar modelos avanzados sin necesidad de infraestructura propia.
- Créditos de computación en la nube: proporcionar créditos de computación en la nube a instituciones educativas y organizaciones sin fines de lucro para facilitar el acceso a LLM.

#### 7. Políticas de equidad digital

- Iniciativas de equidad digital: implementar políticas y programas que aborden la brecha digital, proporcionando acceso a internet de alta velocidad y dispositivos tecnológicos en áreas subrepresentadas.

- Incentivos gubernamentales: ofrecer incentivos gubernamentales para proyectos que promuevan la equidad digital y el acceso a tecnologías avanzadas como los LLM.

#### 8. Evaluación y monitoreo

- Evaluación continua: realizar evaluaciones continuas del impacto de los LLM en diferentes comunidades para identificar y abordar barreras de acceso y uso.
- Monitoreo de uso: monitorear el uso de los LLM para asegurar que las prácticas de acceso y democratización estén siendo efectivas y ajustarlas según sea necesario.

#### 9. Participación comunitaria

- Consulta comunitaria: involucrar a las comunidades locales en el diseño y desarrollo de programas y políticas relacionadas con los LLM para asegurar que sus necesidades y perspectivas sean consideradas.
- Feedback comunitario: recopilar feedback continuo de los usuarios de LLM para mejorar la accesibilidad y la relevancia de estas tecnologías.

Abordar el acceso y la democratización en los LLM es fundamental para asegurar que los beneficios de estas tecnologías sean equitativamente distribuidos y accesibles para todos. Esto requiere un enfoque multifacético que incluya el acceso abierto, infraestructura asequible, educación y capacitación, colaboración internacional, inclusión lingüística y cultural, y políticas de equidad digital. Al implementar estas estrategias, se puede asegurar que los LLM no solo sean una herramienta poderosa, sino también una fuerza para el bien social y la inclusión global.

Los LLM representan un avance significativo en la inteligencia artificial y el procesamiento del lenguaje natural, pero vienen acompañados de importantes desafíos sociales. Abordar estos retos requiere un enfoque multifacético que incluya regulaciones claras, prácticas de desarrollo éticas, transparencia en los modelos, y un compromiso con la equidad y la inclusión. Solo a través de un esfuerzo concertado y colaborativo se puede asegurar que los LLM se utilicen de manera que beneficien a la sociedad en su conjunto.

Realizar un modelo de lenguaje desde cero está al alcance de pocos, se necesita, además, de cantidad enorme de corpus y una capacidad de cómputo que no está al alcance de todo el mundo. Actualmente solo las grandes tecnológicas, como Google o Microsoft, están realizando modelos fundacionales. El caso de OpenAI es excepcional. Durante sus inicios contó con financiación de famosos inversores, como Elon Musk, aunque en 2018 se retiró del proyecto previo al lanzamiento de ChatGPT, en 2022. En este caso, no tuvo la visión o paciencia que la investigación requiere.

En España, la red de centros de ILENIA, que antes he mencionado, está trabajando en la creación de modelos fundacionales usando el Marenostrum 5 y la Red Española de Supercomputación (RES), que son los únicos que tienen la capacidad suficiente.

La restricción de acceso a grandes infraestructuras y el coste energético que supone millones de horas de cómputo han llevado al desarrollo de modelos de lenguaje más pequeños, de 1,5 billones de parámetros o de 6,3 billones, que permiten a los centros de investigación aplicar diferentes técnicas sobre los modelos de lenguaje y resolver tareas concretas de forma efectiva.

## 4. Empleabilidad en el PLN e IA

El empleo generado por el procesamiento del lenguaje natural (PLN) ha experimentado un incremento significativo en los últimos años, con una demanda creciente de profesionales especializados en este campo. El potencial del PLN radica en su capacidad para automatizar tareas lingüísticas complejas, lo que ha llevado a su adopción en una amplia variedad de sectores, desde la atención médica y jurídica hasta la industria tecnológica y el comercio electrónico. Los profesionales del PLN se dedican a desarrollar algoritmos y modelos de aprendizaje automático para mejorar la comprensión del lenguaje humano por parte de las máquinas, lo que incluye el procesamiento de texto, la generación de lenguaje natural, la detección de emociones y la traducción automática, entre otras áreas. Este campo multidisciplinario atrae a expertos en lingüística, informática, estadística y otras disciplinas relacionadas, con un énfasis en la capacidad para comprender y manipular grandes volúmenes de datos lingüísticos de manera eficiente y precisa. Con el avance continuo de la tecnología y el aumento en la adopción de soluciones basadas en PLN, se espera que la demanda de profesionales en este campo siga creciendo en el futuro.

El procesamiento del lenguaje natural (PLN) está adquiriendo cada día un mayor peso económico a nivel internacional. Por una parte, hay empresas e instituciones públicas y privadas que ofertan productos y servicios basados en el PLN (reconocimiento de voz, de imagen, chatbots, etc.). Además, los avances tecnológicos y la disponibilidad de datos, unido a un potencial de mercado muy amplio, están contribuyendo al crecimiento del mercado y del nivel de competencia, con un importante aumento en la oferta de este tipo de productos y servicios.



Por otra parte, el uso del PLN por parte de las empresas está también aumentando y todo apunta a que va a seguir esta tendencia. El PLN se está incorporando en las operaciones fabriles, administrativas y comerciales. El principal beneficio de las soluciones de PLN para las empresas es utilizar la automatización para reducir costes y mejorar las operaciones comerciales para maximizar la productividad y la rentabilidad. En este sentido, las empresas que llevan varios años empleando soluciones de PLN argumentan que esta tecnología mejora la eficacia, genera ventajas competitivas, reduce el tiempo de producción, reduce riesgos, costes y permite acceder a nuevas oportunidades de negocio<sup>16</sup>.

El potencial económico de las lenguas en España se pone de manifiesto con el impulso del “PERTE: La nueva economía de la lengua<sup>17</sup>”, que es un proyecto de carácter estratégico con gran capacidad de arrastre para el crecimiento económico, el empleo y la competitividad, en este caso, para la economía española. La nueva revolución industrial basada en la inteligencia artificial sitúa la lengua cómo uno de los factores clave para el desarrollo tecnológico y empresarial, ya que el tratamiento de la lengua es un parte fundamental de la digitalización y de las nuevas tecnologías y, por consiguiente, de la propia IA. El presupuesto inicial se estimó en 1.100 millones de euros para el período 2021 - 2026, anunciándose también que al menos 30 millones se iban a destinar en exclusiva a proyectos relacionados con lenguas cooficiales del Estado. A esta inversión se sumará la inversión privada que, en su caso, pudiera surgir.

A pesar del potencial económico-laboral atribuido a la aplicación del PLN, no existen apenas estudios que evalúen estas cuestiones. Los únicos informes disponibles han sido elaborados por consultoras que adoptan un enfoque internacional. En el ámbito nacional, solo conta-

.....  
16. <https://www.expert.ai/wp-content/uploads/2022/12/the-2023-expert-nlp-survey-report-trends-driving-nlp-investment-and-innovation.pdf>

17. <https://planderecuperacion.gob.es/como-acceder-a-los-fondos/ertes/perte/perte-nueva-economia-de-la-lengua>

mos con un estudio del año 2018<sup>18</sup> elaborado para el Observatorio Nacional de Tecnología y Sociedad (ONTSI), de ahí la necesidad de actualizar dichos datos. Por este motivo, analizaremos la situación actual y el potencial de las tecnologías lingüísticas, prestando especial atención al impacto en el empleo.

.....  
18. Estudio de Caracterización de él Sector de lanas Tecnologías de él Lenguaje en España 2018. Madrid : ACAP y OESIA, 2018.

## 4.1. Previsiones del mercado de procesamiento del lenguaje natural

Los datos aquí presentados son fruto del trabajo realizado por parte de los profesores de la Universidad de Santiago de Compostela Luis Ignacio Rodríguez y Luis Otero para la Sociedad Española del Procesamiento del Lenguaje Natural.

A nivel mundial, distintas fuentes cuantifican la importancia actual de las tecnologías de procesamiento del lenguaje y pronostican un crecimiento en el mercado muy intenso. En concreto, tal y como muestra la Figura 33, Mordor Intelligence (2020) estima que el mercado global crecerá desde los 10.720 millones de dólares facturados en 2020 hasta los 48.460 millones de 2026, registrando una tasa de crecimiento del 26,84% anual compuesto en el período de pronóstico (2021-2026). En la misma línea, el informe de Markets and Markets (2022) indica que la tasa de crecimiento del sector hasta 2027 será de un 25,7%, pasando de 15.700 a 49.400 millones en 2027. Adicionalmente, Fortune Business Insights (2022) estima un crecimiento próximo al 30% que situará el mercado en 160.000 millones en 2029.

Las empresas que están adoptando tecnologías de procesamiento del lenguaje natural están asignando una cantidad creciente de recursos a esta área. Según los datos recopilados en la encuesta realizada por Gradient Flow en 2021 a 655 líderes tecnológicos de 55 países, un 60% de los encuestados indicó su intención de incrementar al menos en un 19% el presupuesto dedicado al PLN. Además, el 33% reportó un aumento del 30% y el 15% señaló que duplicaron sus recursos. Las proyecciones futuras también apuntan a una tendencia positiva en la inversión en PLN. Según el informe “The 2023 Expert NLP Survey Report”, el 77% de los encuestados planea aumentar su presupuesto en PLN en los próximos 12 a 18 meses, como se ilustra la Figura 34 del informe. Este aumento en

la inversión en tecnologías lingüísticas contribuirá al crecimiento continuo del mercado, ya que la demanda de este tipo de soluciones sigue en aumento.

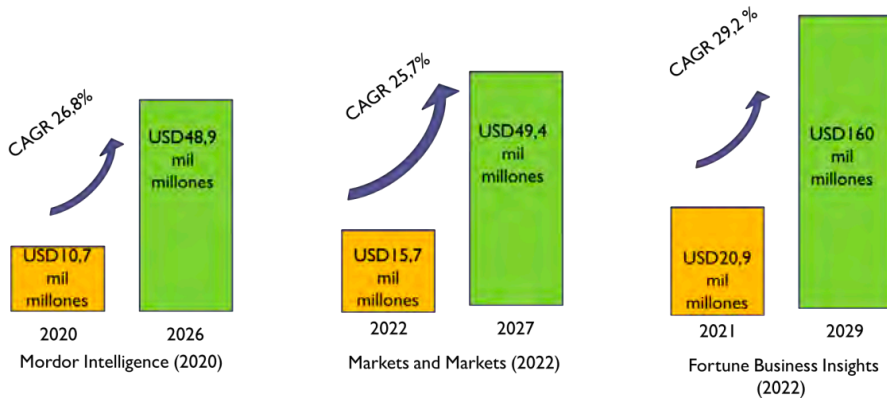
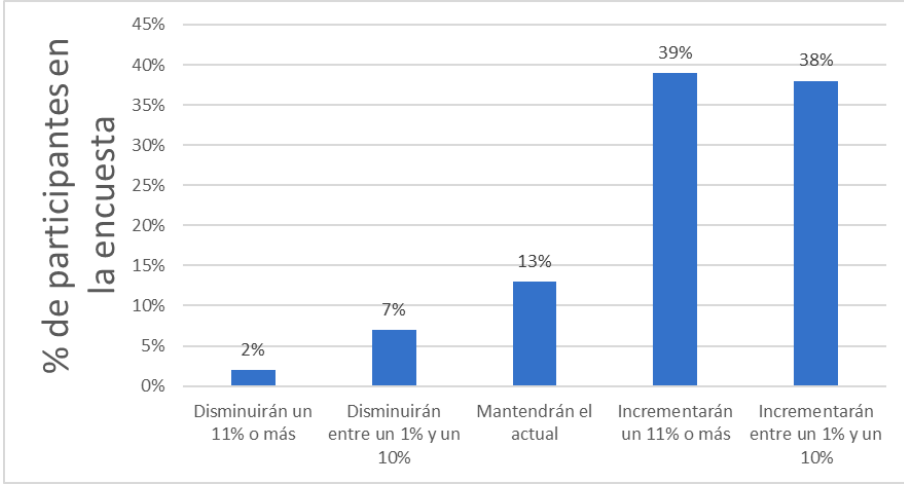


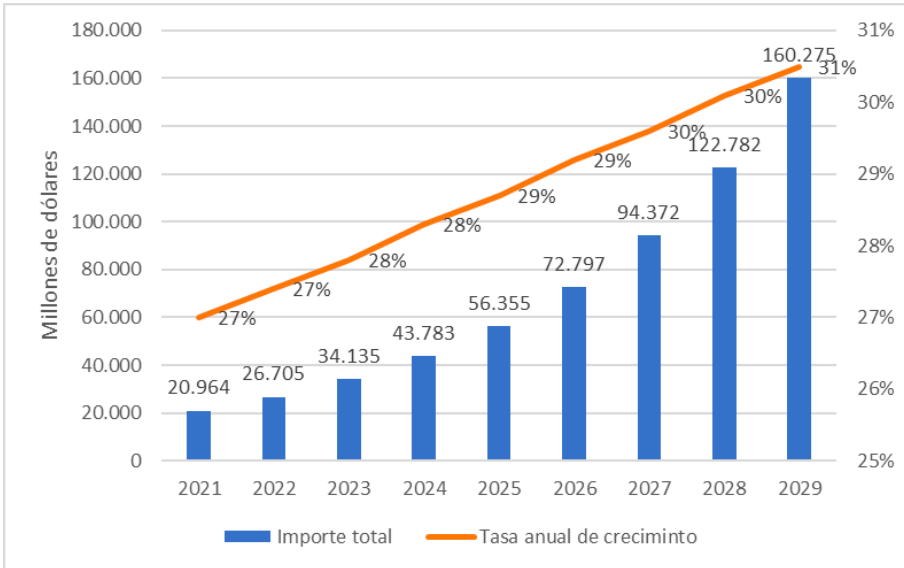
Figura 33: Previsiones del mercado PLN a nivel global

Las empresas en el sector del PLN pueden experimentar un crecimiento muy destacable, gracias a los avances tecnológicos en curso y a la abundancia de datos disponibles. Esto garantiza una oferta sostenida de productos y servicios en este ámbito, con potencial para un aumento adicional. Por otro lado, las empresas usuarias también pueden anticipar un panorama alentador, ya que la demanda de tecnologías de procesamiento del lenguaje está en alza debido a las ventajas inherentes que ofrecen.

Todas las fuentes consultadas pronostican un incremento mundial del mercado del PLN. En la Figura 35 se puede observar un crecimiento continuo a tasas muy elevadas y una estimación de más de 160.000 millones para el año 2029.



**Figura 34: Previsión sobre la variación en el presupuesto dedicado a PLN**  
 Fuente: The 2023 Expert NLP Survey Report (2022)



**Figura 35: Evolución del mercado mundial del PLN**  
 Fuente: Fortune Business Insights (2022)

Cuando abordamos la situación en el mercado europeo, al igual que lo que se observa para el mercado mundial, la tendencia esperada es de crecimiento sostenido durante el período 2020 -2028. Como se puede observar en la Figura 36, las distintas fuentes consultadas pronostican un aumento del volumen de mercado, si bien difieren de forma notoria. Una de las razones que puede explicar estas diferencias es la propia consideración de las tecnologías que se incluyen y si realmente se refieren al procesamiento del lenguaje natural o a las tecnologías lingüísticas.

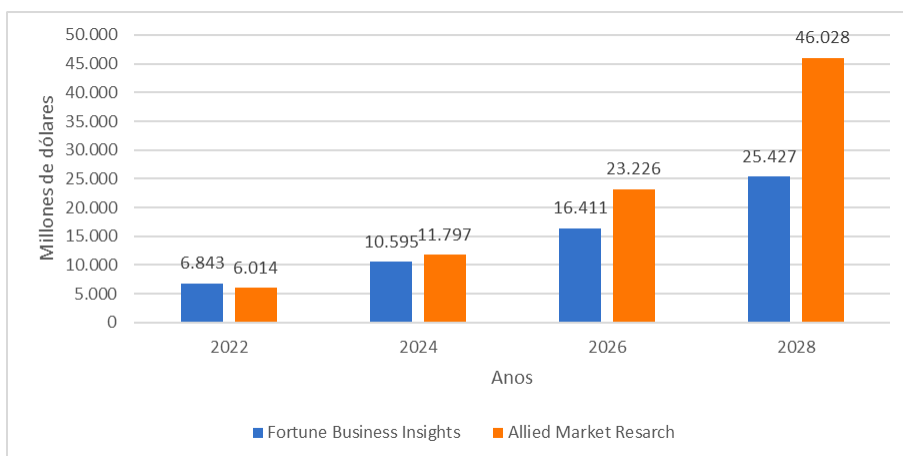


Figura 36: Evolución del mercado europeo del PLN

Fuente: Fortune Business Insights (2022) y Allied Market Reserarch (2022)

Para España, los datos referidos a 2022 de empresas de 10 o más personas empleadas reflejan que el uso de tecnologías de IA ha experimentado un incremento de casi el 40% con respecto al año anterior. Evolución aún más positiva es la que ha experimentado la comunidad de La Rioja, con un incremento del 130%.

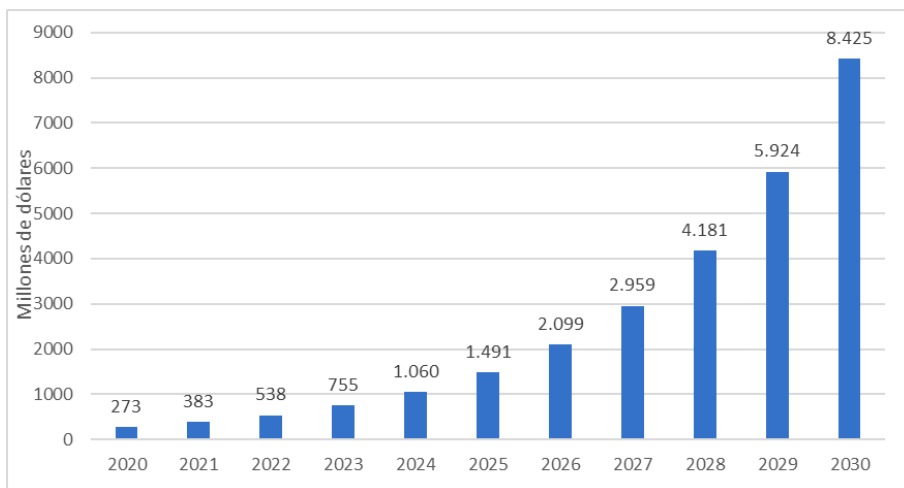


Figura 37: Evolución del mercado en España del PLN

Fuente: Allied Market Reserarch (2022)

Como se puede apreciar en la Figura 8, el peso de los agentes del PLN no es el mismo en todas las comunidades autónomas. La mayoría de los agentes del sector se concentran en Madrid, Cataluña, País Vasco, Valencia y Galicia, mientras que La Rioja representa un 2% del mercado.

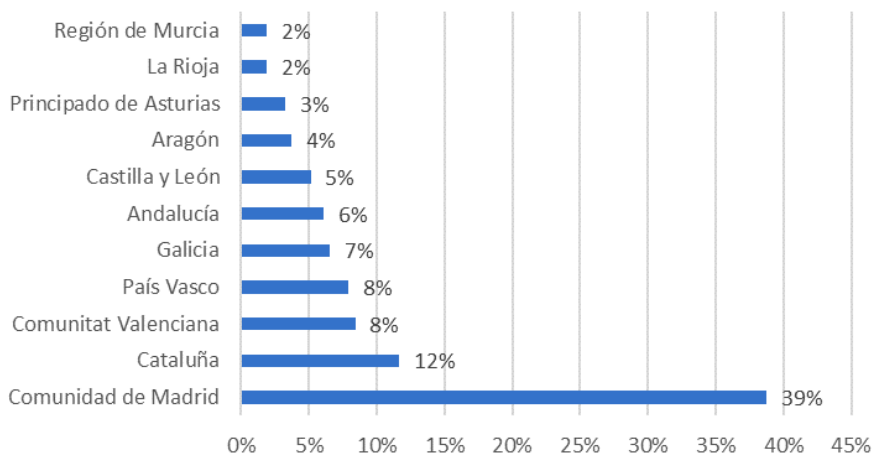


Figura 38: Agentes del PLN (%) por comunidad autónoma

Fuente: elaboración propia a partir del Registro de empresas y centros de investigación, Mapa de capacidades de tecnologías de IA y el Mapa de capacidades digitales de Galicia.

Para valorar la importancia económica del procesamiento del lenguaje natural se pueden considerar distintas variables, como son el volumen de ventas por productos y servicios y el personal que trabaja también de forma específica en este ámbito. Los datos mostrados hacen referencia al total de las empresas. De las 115 identificadas solo se obtuvieron datos para 103 a partir de la base de datos SABI. Por otra parte, no es posible saber a partir de la información pública el volumen de negocio específico del PLN, ya que las empresas no detallan el porcentaje que representa del total de ingresos. Lo mismo sucede para los datos referidos al personal. Con respecto a la facturación, las 103 empresas consideradas obtuvieron una cifra de negocio próxima a los 15.000 millones de euros en 2019, cuantía que disminuyó hasta los 13.300 millones de euros en 2021, cifra que se vio afectada por los efectos de la pandemia y que se recuperó de forma muy notable en 2022, al alcanzar los 15.400 millones de euros. No obstante, es difícil precisar el porcentaje de dicho mercado que corresponde al PLN, ya que la mayor parte de las empresas pertenecen al sector TIC y ofrecen otros productos y servicios.

La adopción del procesamiento del lenguaje natural (PLN) en actividades industriales, administrativas y comerciales ha experimentado un crecimiento muy destacable. Esta tendencia se debe, en parte, a la creciente digitalización de las empresas, que buscan soluciones PLN para mejorar su eficiencia, reducir costes y obtener ventajas competitivas significativas. Por otro lado, las empresas dedicadas al sector de las tecnologías del lenguaje (TL) están ofreciendo soluciones de PLN cada vez más avanzadas, impulsadas tanto por avances tecnológicos como por la disponibilidad de datos.

Este panorama prometedor sugiere un crecimiento significativo en el uso de estas tecnologías en los próximos años. Aunque los estudios son escasos y generalmente llevados a cabo por grandes consultoras a nivel internacional, todos coinciden en pronosticar un crecimiento muy destacable, hasta alcanzar los 160.000 millones de dólares para el año 2029, con un aumento especialmente notable en Asia, impulsado por China e India.



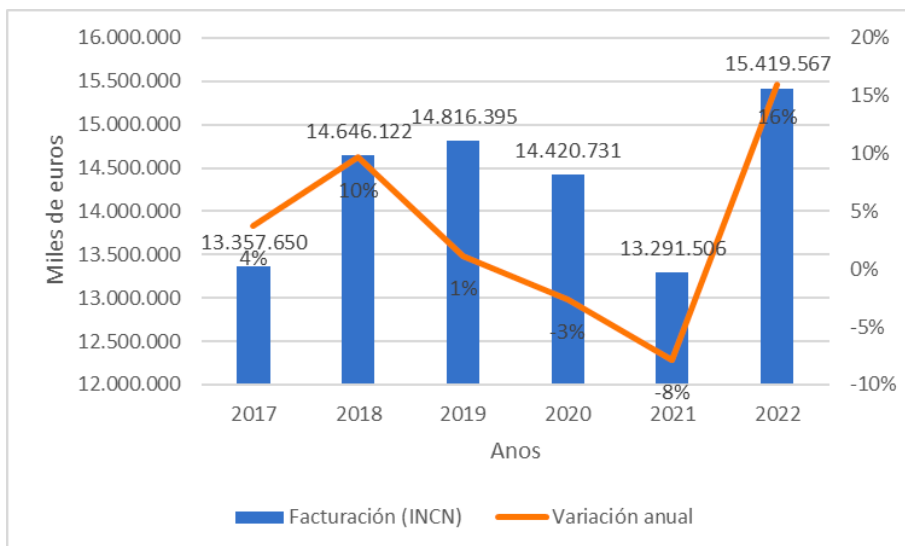


Figura 39: Facturación de las empresas del sector del PLN

Identificar las empresas españolas dedicadas al PLN no es una tarea sencilla, ya que muchas de ellas también realizan otras actividades. Sin embargo, en España se han identificado un total de 214 agentes distribuidos por todo el territorio, con una concentración significativa en la Comunidad de Madrid, seguida por Cataluña, la Comunidad Valenciana, el País Vasco y Galicia. Estas empresas tienen un impacto económico considerable, representando aproximadamente el 1% del PIB nacional, con una facturación que superó los 15.400 millones de euros en 2022.

Además, estas empresas generan un valor añadido significativo, ya que representan el 67% de los ingresos, en comparación con el 24% de otros sectores de la economía. Es importante destacar que el empleo que generan es altamente cualificado, con un gasto salarial anual por trabajador en el sector del PLN de 63.000 euros, en contraste con los 37.000 euros en otros sectores.

Según datos recientes, el 76% de las empresas dedicadas al PLN pertenecen a sectores que conforman el sector TIC. Entre estos sectores desta-

can: otros servicios relacionados con las tecnologías de la información y la informática (6.209); actividades de programación informática (6.201); Actividades de consultoría informática (6.202); otras actividades de telecomunicaciones (6.190); comercio al por mayor de ordenadores y equipos periféricos (4.651); gestión de recursos informáticos (6.203); edición de otros programas informáticos (5.829); telecomunicaciones por cable (6.110) y fabricación de componentes electrónicos (2.611).

En términos de empleo, se observa un crecimiento constante en el número de personas empleadas en estas empresas en los últimos años. Según datos del año 2021, el número total de personas empleadas en el subsector de Programación, consultoría y otras actividades relacionadas con la informática (CNAE 6201, 6202, 6203 y 6209), al que pertenecen el 63% de las empresas de PLN, alcanzó la cifra de 343.812 empleados.

Al analizar el empleo en España según el tamaño de las empresas dedicadas al PLN, se observa que aproximadamente el 54% de ellas son pymes o microempresas, mientras que un 33% corresponden a empresas con más de 250 empleados. Este panorama refleja una cierta polarización en el sector, donde las pymes y microempresas predominan, aunque también se evidencia la presencia significativa de empresas más grandes.

Determinar el número exacto de personas dedicadas al PLN es un desafío debido a la escasez de información, pero se estima que hay entre 1.327 y 1.793 personas trabajando en PLN en las empresas identificadas en 2023, y entre 2.773 y 3.662 si se consideran todas las entidades dedicadas al PLN. Es importante destacar que el personal dedicado a actividades de PLN, en su mayoría (aproximadamente el 71%) posee títulos superiores, lo que subraya la importancia de la formación especializada en este campo.

El empleo relacionado con el PLN abarca una amplia gama de perfiles profesionales, desde ingenieros de sistemas y científicos de datos hasta lingüistas y auditores de algoritmos. Se observa una tendencia hacia la

contratación de profesionales con habilidades tanto técnicas como humanísticas, lo que refleja la multidisciplinariedad de este campo y la necesidad de abordar aspectos críticos relacionados con la ética y la filosofía.

Las perspectivas de empleo en el sector del PLN son muy alentadoras, con un crecimiento continuo proyectado en los próximos años. Según el Informe de prospección y detección de necesidades formativas de 2021, la tendencia del empleo en ocupaciones relacionadas con la familia profesional de Informática y Comunicaciones, donde se incluye el PLN, es claramente alcista.

Además, la demanda de profesionales en el campo de la IA, que incluye áreas como el aprendizaje automático, el procesamiento del lenguaje natural y los sistemas expertos, está en aumento. Este crecimiento se refleja en el incremento de ofertas laborales relacionadas con la IA, que han experimentado un aumento significativo en los últimos años tanto en España como en Europa.

Según datos de Eurostat, el porcentaje de empresas en Europa que utilizan tecnologías de PLN está en aumento, lo que generará una mayor demanda de profesionales en este campo para satisfacer las necesidades del mercado laboral en constante evolución.

En resumen, el sector del procesamiento del lenguaje natural ofrece un panorama prometedor en términos de empleo, con un crecimiento continuo proyectado en los próximos años. La demanda de profesionales especializados en PLN está en aumento, lo que ofrece oportunidades significativas para aquellos interesados en ingresar a este campo multidisciplinario en constante evolución.

## 5. Conclusión

Para finalizar, destacaría el enorme potencial que tienen la IA generativa y el PLN para transformar positivamente nuestra sociedad. Ambas tecnologías han evolucionado enormemente desde el aumento de la capacidad de cómputo por parte de los supercomputadores y el desarrollo de algoritmos como los Transformers.

Debemos abordar la complejidad del lenguaje humano, la falta de datos, la necesidad de explicabilidad y transparencia en los modelos de IA, y la ética y responsabilidad en el desarrollo de los sistemas. Nos espera un futuro donde los sistemas de IA puedan entender y responder a nuestras preguntas de manera natural y precisa, donde los robots puedan ayudarnos en nuestras tareas diarias y donde la IA pueda mejorar la salud, la educación y la calidad de vida de las personas.

Debemos ser éticos y responsables en el desarrollo de la IA y debemos asegurarnos de que la IA no incremente las desigualdades o cree nuevos problemas y de que sea accesible para todos, sin importar la edad, el género, la raza o la capacidad económica.

Es necesario establecer marcos legales y éticos claros para el desarrollo y uso de la IA, promover la transparencia y fomentar la educación y la comprensión de estas tecnologías. La colaboración entre gobiernos, empresas, academia y sociedad civil será fundamental para desarrollar de forma efectiva este nuevo panorama tecnológico en pro de un futuro equitativo y beneficioso para todos.

Por otro lado, se debe garantizar la formación en estas tecnologías, formar profesionales en PLN que sepan integrar las tecnologías lingüísticas con los algoritmos de IA. Felicito a la Universidad de La Rioja por la propuesta de grado en Lingüística Computacional que estoy conven-

cido, será todo un éxito, dada las enormes perspectivas de empleabilidad que existen.

Si bien no hay que temer a la IA, no es menos cierto que se deben tener en cuenta los riesgos. En mi opinión, la IA debe servir al ser humano a la toma de decisiones que mejoren nuestras vidas, pero no hay que dejarla sin supervisión, que tome decisiones importantes que afecten a las personas.

## 6. Referencias

- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1996). *The CELEX lexical database (Release 2)*. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.
- Backus, J.; Beeber, R. J.; Best, Sheldon F.; Goldberg, Richard; Herrick, Harlan L.; Hughes, R. A.; Mitchell, L. B.; Nelson, R. A.; Nutt, Roy; Sayre, D.; Sheridan, P. B.; Stern, H.; Ziller, I. (1956). Sayre, David, ed. *The FORTRAN Automatic Coding System for the IBM 704 EDPM: Programmer's Reference Manual*. New York, USA: Applied Science Division and Programming Research Department, International Business.
- Baker, C. F., Fillmore, C. J., & Lowe, J. B. (1998). *The Berkeley FrameNet Project*. Proceedings of the 17th International Conference on Computational Linguistics (COLING) and 36th Annual Meeting of the Association for Computational Linguistics (ACL), Volume 1. 86-90.10.3115/980845.980860.
- Brown, G., & Yule, G. (1983). *Discourse Analysis*. Cambridge University Press.
- Chomsky, N. (1986). *Knowledge of language*. New York: Praeger.
- Chomsky, N., & Bouallal, K. A. R. I. M. A. (1968). *Gramática generativa y transformacional*. México FCE.
- Conneau, A. and Kiela, D. *SentEval: An Evaluation Toolkit for Universal Sentence Representations*. (21018). Facebook Artificial Intelligence Research. <https://arxiv.org/pdf/1803.05449>.
- Conneau, A., Schwenk, H., Barrault, L. and Lecun, Y. "Very deep convolutional networks for text classification," in European ACL, vol. 1, 2017, pp. 1107-1116.
- Fellbaum, C. (Ed.). (1998). *WordNet: An Electronic Lexical Database*. MIT Press.
- Haspelmath, M., & Sims, A. D. (2013). *Understanding Morphology (2nd ed.)*. Routledge.
- Francis, W. N., & Kucera, H. (1964). *The Brown Corpus of Standard American English*. Brown University.

- Franco, M., Romero, M.A., Palomar, M.E. Cobos, J.A., Álvarez, A. , Hernández, D. *De neuronas biológicas a neuronas artificiales, el fascinante mundo de las redes neuronales*. <https://www.uaeh.edu.mx/divulgacion-ciencia/redes-neuronales/>.
- Garofolo, J. S., et al. (1993). TIMIT Acoustic-Phonetic Continuous Speech Corpus. *Linguistic Data Consortium*.
- Hochreiter, S., & Schmidhuber, J. (1996). LSTM can solve hard long time lag problems. *Advances in neural information processing systems*, 9.
- Jurafsky, D. and Martin, J. *Speech & language processing*. Pearson Education, 2000.
- Kim, Y. “*Convolutional neural networks for sentence classification*,” arXiv preprint arXiv:1408.5882, 2014.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
- LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. (1998). “*Gradient-based learning applied to document recognition*” (PDF). *Proceedings of the IEEE*. 86 (11): 2278-2324. doi:10.1109/5.726791. S2CID 14542261.
- Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., ... & Bizer, C. (2015). DBpedia - A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. *Semantic Web Journal*, 6(2), 167-195. 10.3233/SW-140134.
- Manning, C. D., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press.
- Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a Large Annotated Corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2), 313-330.
- Martí, M. A., et al. (2007). The CESS-CAST Corpus: A Spanish Corpus with Rich Morphosyntactic Annotation. *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*.
- Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Pearson/Prentice Hall.

- Michel, J.-B., et al. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, 331(6014), 176-182.
- Miller, G. A. (1990) «Wordnet: An On-line Lexical Database», *International Journal of Lexicography* 3(4).
- Navarro, B. (2024). *Minería de textos. Primera parte. Fundamentos de Procesamiento del Lenguaje Natural*. Máster en Ciencia de Datos Universidad de Alicante Curso 2023-2024
- Real Academia Española (2005). *Corpus de Referencia del Español Actual (CREA)*.
- Rajpurkar, P., Zhang, J., & Liu, J. (2016). SQuAD: 100,000+ Questions for Machine Comprehension of Text. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP 2016)*.
- Rodríguez, F. J., & Gómez, J. (2014). Spanish Billion Words Corpus. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014)*.
- Rosenblatt, F. *The perceptron: A probabilistic model for information storage and organization in the brain*. *Psychological Review* 65 (6):386-408 (1958)
- Sánchez-Gutiérrez, C.H., Mailhot, H., Deacon, S.H. et al. MorphoLex: A derivational morphological database for 70,000 English words. *Behav Res* 50, 1568-1580 (2018). <https://doi.org/10.3758/s13428-017-0981-8>
- Sang, E. F., & Veenstra, J. (1999). Representing Texts for Machine Learning: A Case Study of the CoNLL-2003 Shared Task. *Proceedings of the CoNLL-2003 Shared Task*.
- Schuler, K. K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. University of Pennsylvania.
- Sowa, J. F. (1999). *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co.
- Suchanek, F. M., Kasneci, G., & Weikum, G. (2007). YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia. *Proceedings of the 16th International Conference on World Wide Web (WWW 2007)*.: 697-706. 10.1145/1242572.1242667.



- Taboada, M., et al. (2016). TASS 2016: Análisis de Sentimientos en Español. *Proceedings of the 2nd Workshop on Spanish Sentiment Analysis (TASS 2016)*.
- Taule, M., et al. (2008). *AnCora: A Multilingual Corpus for NLP*. Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008).
- Turing, A. M. (1950). *Mind*. *Mind* 59(236), 433-460.
- Van Hee, C., Jacobs, G., Emmery, C., Desmet, B., Lefever, E., Verhoeven, B., ... & Hoste, V. (2018). *Automatic detection of cyberbullying in social media text*. *PloS one*, 13(10).
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). *Attention is all you need*. *Advances in neural information processing systems*, 30.
- Vossen, P. *EuroWordNet: A multilingual database with lexical semantic networks. Computers and the humanities*. <https://books.google.com.pe/books?id=-qEep-1ib8UC> (1998). Springer Netherlands.
- Williams, R. J., & Zipser, D. (1989). *A Learning Algorithm for Continually Running Fully Recurrent Neural Networks*. *Neural Computation*, 1(2), 270-280.

**2024**  

---

**2025**