

TÉCNICAS DE PROCESAMIENTO DEL HABLA

Sonido, Transformada de Fourier y Análisis de Sentimientos con NLTK

BREVE RESUMEN DE LOS TEMAS A CUBRIR:

- - Definición básica de sonido y características relevantes para el procesamiento.
 - Uso de la Transformada de Fourier para analizar frecuencias en señales de audio.
 - Análisis de sentimientos aplicado al lenguaje con la librería NLTK de Python.
-

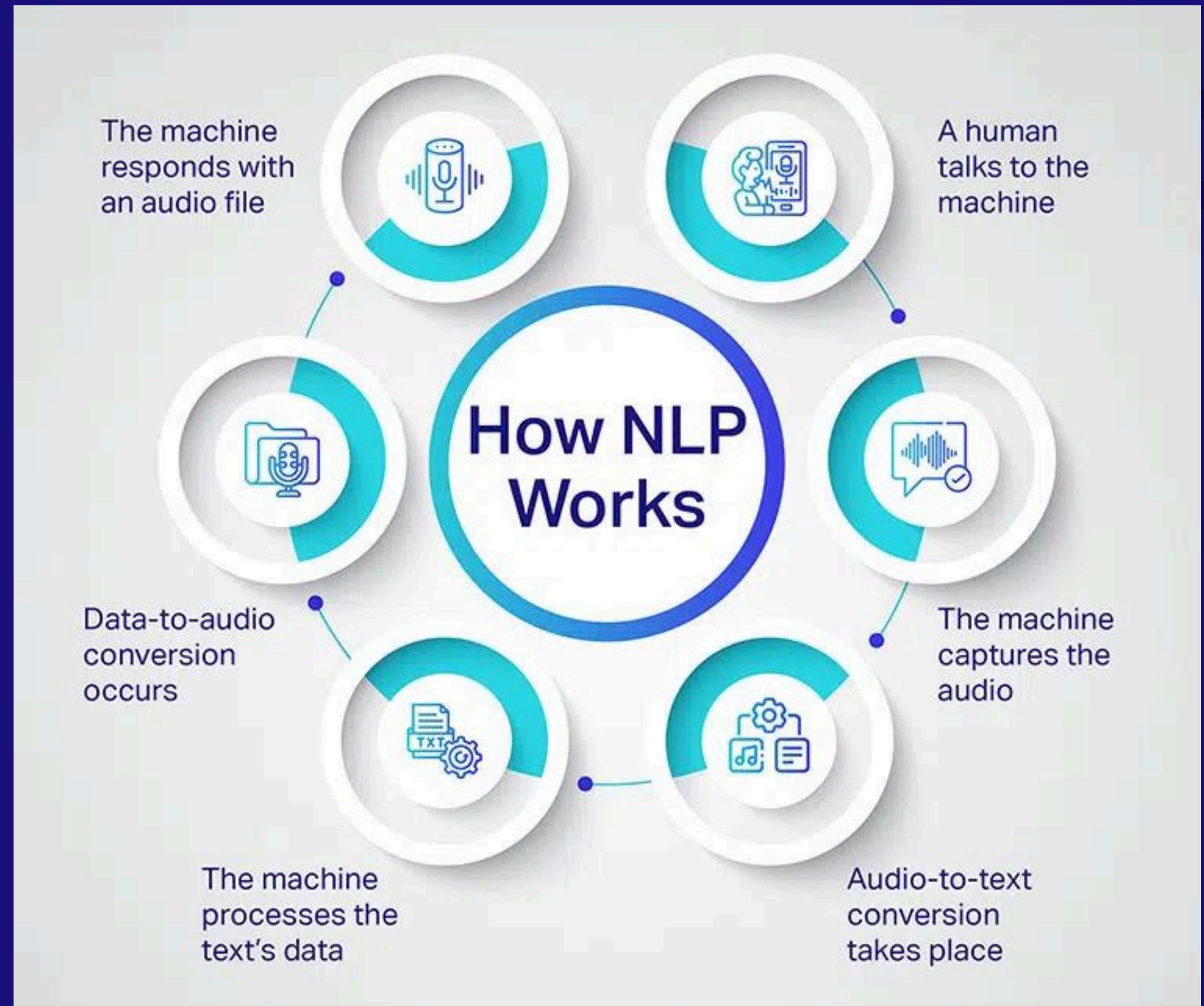
Sonido

- Definición: El sonido es una onda mecánica que viaja a través de un medio (aire, agua, etc.) y que puede ser percibida por el oído humano. Es energía que se desplaza en el medio.
- Características principales:
 - Frecuencia: Medida en Hertz (Hz), determina el tono del sonido.
 - Amplitud: Relacionada con el volumen o la intensidad del sonido.
 - Timbre: Calidad o color del sonido, diferencia entre instrumentos o voces.

REPRESENTACION DEL SONIDO

- Muestreo: Convertir un sonido continuo en una señal discreta mediante la captura de muestras a intervalos regulares.
- Frecuencia de muestreo: La tasa a la que se toman las muestras, por ejemplo, 44.1 kHz para audio de alta calidad.
- Cuantización: Representación digital de la amplitud del sonido en un rango de valores discretos.

APLICACIONES



APLICACIONES

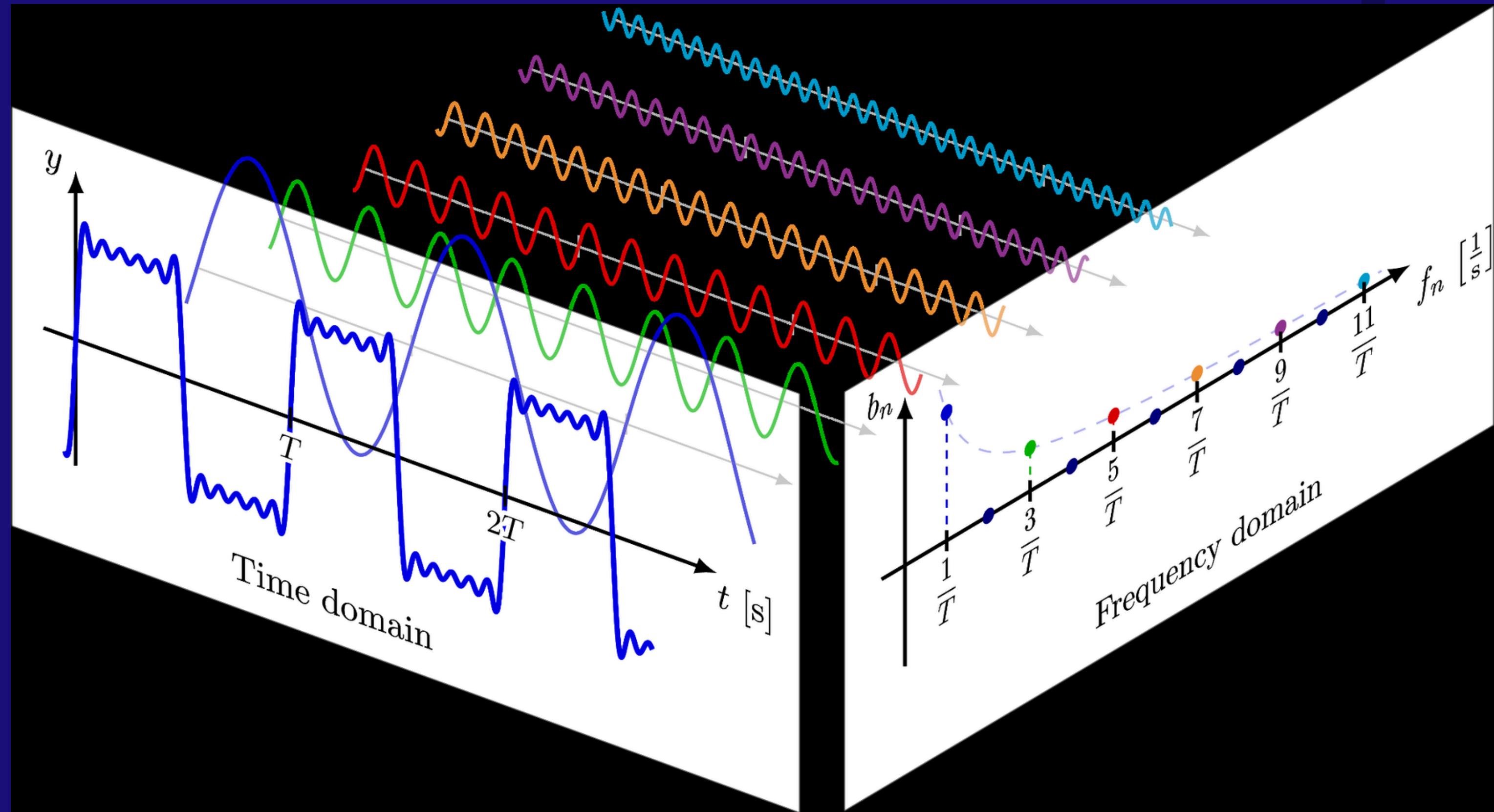
Algunas de las aplicaciones más básicas de la PNL incluyen:

- Traducción de idiomas en tiempo real
- Filtros de spam en servicios de correo electrónico
- Asistentes de voz y chatbots
- Resumen de texto
- Funciones de autocorrección
- Análisis de sentimiento y más

TRANSFORMADA DE FOURIER

- Concepto: Herramienta matemática que descompone una señal en sus componentes de frecuencia.
- Importancia en el procesamiento de audio:
- Permite analizar cómo se distribuyen las frecuencias en una señal de sonido.
- Utilizada para la reducción de ruido, la síntesis de sonido y el reconocimiento de patrones de voz.
- Transformada de Fourier Rápida (FFT): Algoritmo eficiente para calcular la transformada de Fourier.

ENFOQUES



ANÁLISIS DE SENTIMIENTOS CON NLTK

- ¿Qué es el análisis de sentimientos?: Técnica para determinar la polaridad emocional (positiva, negativa o neutral) en textos.
- Procesos clave:
 - Tokenización: Separar el texto en palabras o frases.
 - Limpieza de texto: Eliminar caracteres no deseados, como puntuaciones o palabras vacías (stopwords).
 - Modelos de clasificación: Asignar etiquetas de sentimiento a fragmentos de texto.

QUE ES NLTK?

Herramienta NLTK

- NLTK (Natural Language Toolkit): Librería de Python para el procesamiento del lenguaje natural.
- Componentes clave:
 - Tokenizadores.
 - Diccionarios de stopwords.
 - Clasificadores de sentimiento.

FOURIER

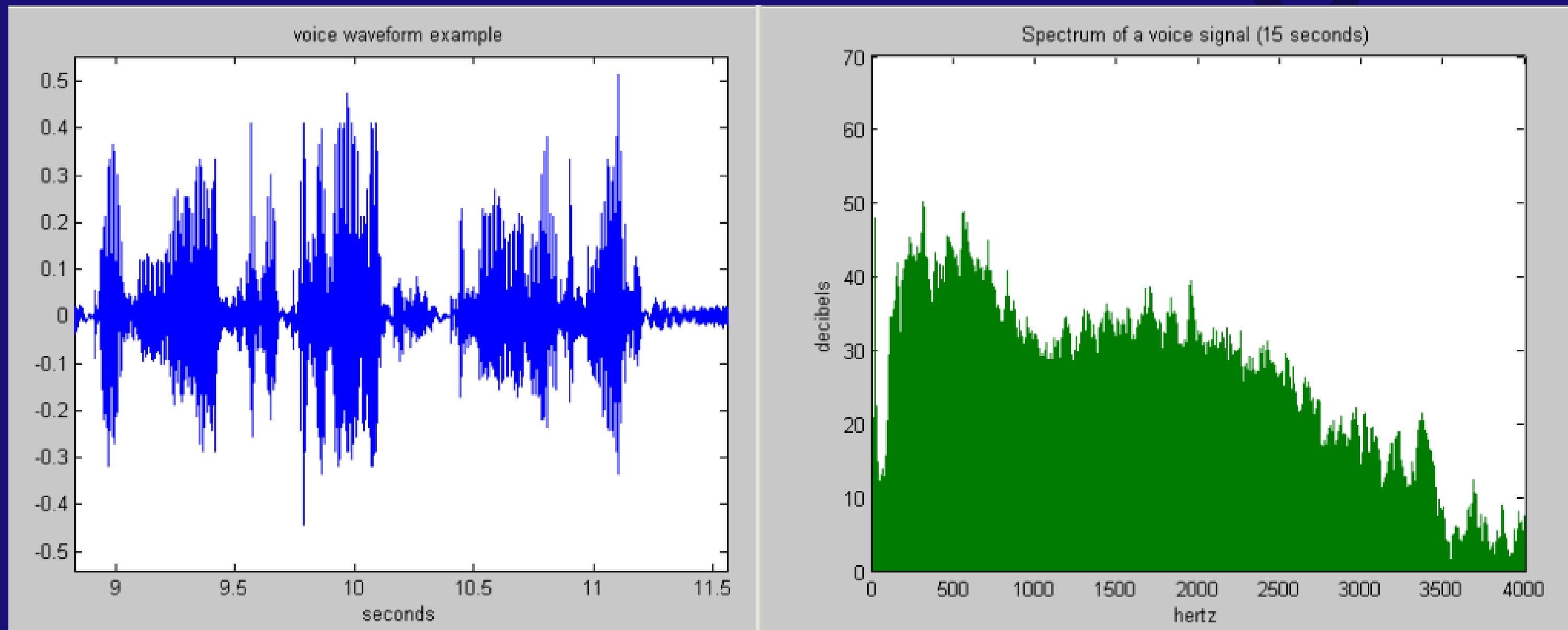


el dominio de la frecuencia y cubre series de Fourier, transformada de Fourier, propiedades de Fourier, FFT, ventanas y espectrogramas, utilizando ejemplos de Python.

La transformada de Fourier es una transformación matemática usada para transformar señales entre el dominio del tiempo o espacio al dominio de la frecuencia, y viceversa. El concepto de 'transformada de Fourier' se refiere a varios conceptos de forma simultánea:

- Operación de transformación de una función.
- Función resultado de la operación.
- Espectro de frecuencias de una función.

EL ESPECTRO DEL SONIDO



FOURIER

A pesar de que calcular la transformada de Fourier de una función es bastante sencillo, es necesario aprender a **integrar múltiples de funciones complejas** con anterioridad. Si $x(t)$ es la función original, su transformada $X(f)$ será:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt$$

EL ESPECTRO DE SONIDO



El espectro de sonido se mide en hertzios (Hz), que representan la cantidad de ciclos por segundo de una onda sonora. Las frecuencias más bajas, en el rango de 20 Hz a aproximadamente 200 Hz, se perciben como sonidos graves. Estos sonidos son producidos por vibraciones lentas, como las de un bajo en una canción o el trueno de una tormenta. El espectro audible,¹ también denominado campo tonal, se encuentra conformado por las audiofrecuencias, es decir, toda la gama de frecuencias que pueden ser percibidas por el oído humano.²

EL ESPECTRO DE SONIDO

Un oído sano y joven es sensible a las frecuencias comprendidas entre los 20 Hz y los 20 kHz. No obstante, este margen varía según cada persona y se reduce con la edad (llamamos presbiacusia a la pérdida de audición con la edad). Este rango equivale muy aproximadamente a diez octavas completas ($2^{10}=1024$). Frecuencias más graves incluso de hasta 4 ciclos por segundo son perceptibles a través del tacto, cuando la amplitud del sonido genera una presión suficiente.

Fuera del espectro audible:

- Por encima estarían los ultrasonidos (Ondas acústicas de frecuencias superiores a los 20 kHz).
- Por debajo, los infrasonidos (Ondas acústicas inferiores a los 20 Hz).

BIOMETRIA DE LA VOZ

Además, en los últimos años ha surgido una nueva aplicación: el **reconocimiento de voz**, una tecnología que abre las puertas a la **identificación biométrica** por sonido.

El fraude de identidad y la suplantación han compartido espacio en las mismas conversaciones y foros en los que lo ha hecho la biometría de voz, al igual que en aquellos dedicados a la experiencia de cliente y la productividad de los equipos de ventas. Compañías de sectores como la banca, los servicios financieros, las telecomunicaciones o los servicios públicos están ya haciendo uso de esta tecnología que ha demostrado ser fiable, segura y generadora de oportunidades.

Analisis de sentimientos



ANALISIS DE SENTIMIENTOS

El análisis de sentimientos es una técnica utilizada dentro del campo del Procesamiento del Lenguaje Natural (PNL) para determinar la emoción o actitud expresada en un fragmento de texto. Este análisis puede ayudar a detectar si el sentimiento general es positivo, negativo o neutral. Algunas de las principales aplicaciones incluyen:

- Opiniones en redes sociales: Para entender la percepción del público sobre marcas, productos, o eventos.
- Atención al cliente: Clasificar las interacciones en función del tono para mejorar la experiencia.
- Encuestas y retroalimentación: Para analizar automáticamente respuestas y agruparlas en categorías.

ANALISIS DE SENTIMIENTOS

Flujo del Análisis de Sentimientos

1. Preprocesamiento: El texto debe ser limpiado eliminando ruido como puntuación, números o palabras irrelevantes.
2. Tokenización: Dividir el texto en palabras o frases clave.
3. Lematización/Stemming: Reducir las palabras a su forma base o raíz.
4. Vectorización: Convertir el texto en datos numéricos para que los modelos puedan procesarlo (por ejemplo, usando TF-IDF o word embeddings).
5. Clasificación: Usar un modelo de machine learning o una técnica de diccionario para etiquetar los sentimientos.
6. Evaluación: Verificar la precisión y efectividad del modelo.

NLTK

Implementación con NLTK

Una de las bibliotecas más utilizadas en Python para análisis de sentimientos es NLTK (Natural Language Toolkit), que provee herramientas para preprocesamiento de texto y acceso a recursos como lexicon de sentimientos (por ejemplo, VADER, que es un lexicon específico para redes sociales).

Para un trabajo práctico, se podría implementar un pipeline con las siguientes etapas:

1. Importar y preprocesar el texto con NLTK.
2. Aplicar el análisis de sentimientos utilizando el clasificador VADER u otro modelo de ML (como Naive Bayes).
3. Visualizar los resultados usando gráficos de distribución de sentimientos.

GES C.



VADER

Valence Aware Dictionary and sEntiment Reasoner” y es la librería que usa Python para el análisis de sentimientos.

Esta librería se basa en reglas y un poderoso diccionario para realizar sus análisis y determinar si una noticia es positiva, negativa y neutra. El resultado que nos devolverá Vader va desde -1 a 1, siendo -1 una noticia muy negativa, 0 una noticia neutra, y 1 una noticia muy positiva.

En términos de precisión, cuando se han realizado las pruebas Vader ha demostrado incluso ser más preciso que un humano. Esta librería tiene un porcentaje de acierto con la métrica F1 de un 96%, mientras que los evaluadores humanos tienen un 84% al intentar etiquetar una noticia como positiva, negativa o neutra.

VADER

Valence Aware Dictionary and sEntiment Reasoner” y es la librería que usa Python para el análisis de sentimientos.

Esta librería se basa en reglas y un poderoso diccionario para realizar sus análisis y determinar si una noticia es positiva, negativa y neutra. El resultado que nos devolverá Vader va desde -1 a 1, siendo -1 una noticia muy negativa, 0 una noticia neutra, y 1 una noticia muy positiva.

En términos de precisión, cuando se han realizado las pruebas Vader ha demostrado incluso ser más preciso que un humano. Esta librería tiene un porcentaje de acierto con la métrica F1 de un 96%, mientras que los evaluadores humanos tienen un 84% al intentar etiquetar una noticia como positiva, negativa o neutra.



BIOMETRÍA DE LA VOZ

Definición:

- La biometría de la voz es una tecnología que utiliza características únicas de la voz de una persona para su identificación y autenticación.
- Diferencia clave con otras técnicas biométricas:
- Mientras que la biometría de huellas dactilares o el reconocimiento facial se basan en rasgos físicos, la biometría de la voz utiliza características fisiológicas (forma del tracto vocal) y conductuales (forma de hablar).

BIOMETRÍA DE LA VOZ

Componentes que definen la biometría de la voz:

- Frecuencia fundamental (tono): Relacionada con la vibración de las cuerdas vocales.
- Timbre y resonancia: Determinados por la forma del tracto vocal (garganta, boca, nariz).
- Pronunciación y acento: Factores conductuales que varían según el contexto cultural y social.
- Duración y ritmo del habla: Forman patrones únicos de cada persona.
-



BIOMETRÍA DE LA VOZ

Seguridad y autenticación:

- Autenticación en servicios bancarios o aplicaciones móviles (Ej. reconocimiento de voz en bancos).
- Control de acceso a sistemas sensibles (Ej. en edificios corporativos o sistemas de telecomunicaciones).

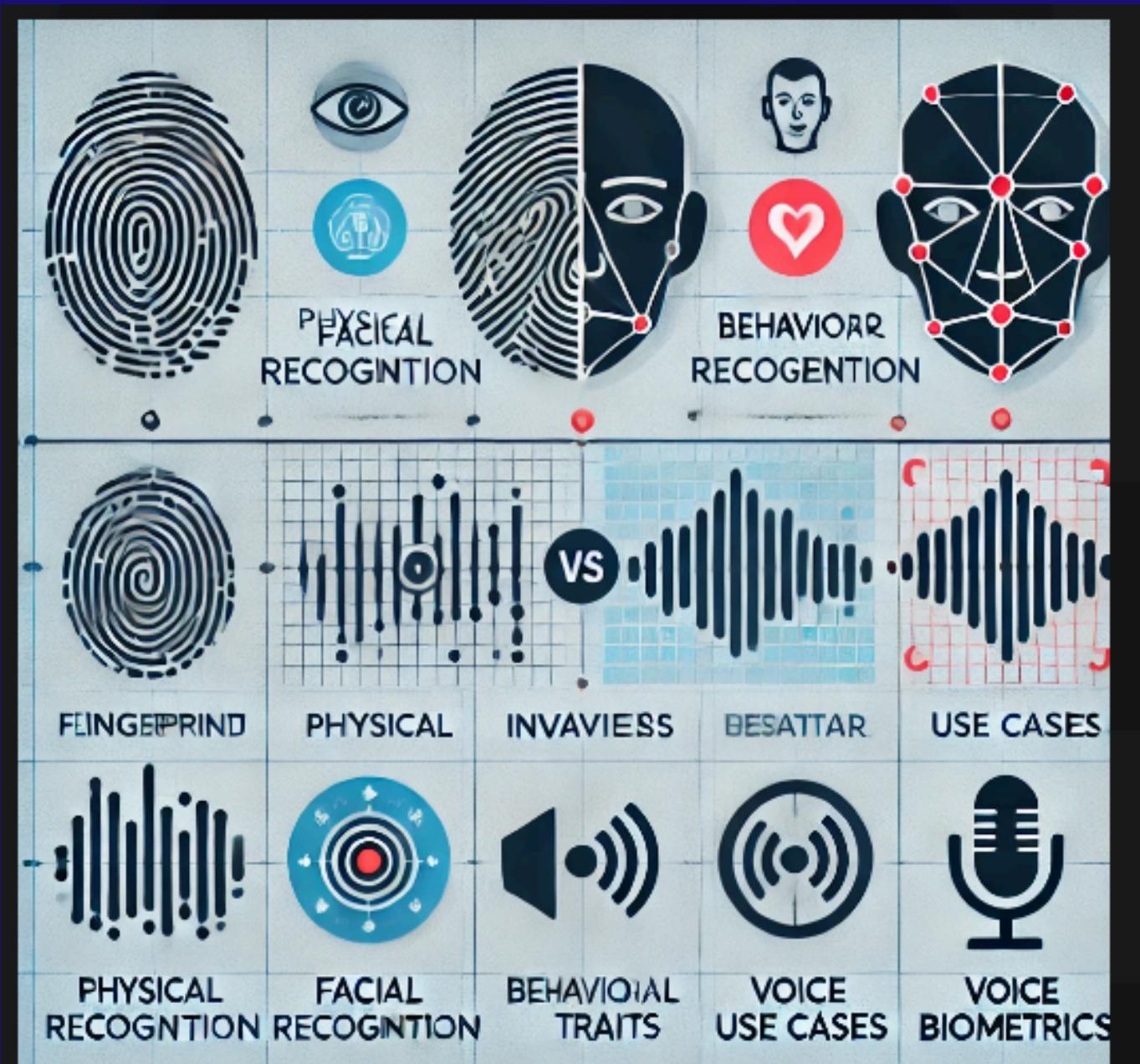
Forense:

- Identificación de personas a partir de grabaciones de voz en investigaciones criminales.

Asistentes virtuales:

- Personalización de asistentes como Alexa, Google Assistant o Siri, que pueden reconocer voces de diferentes usuarios.
- Telemedicina y bienestar:
- Monitoreo de emociones y estados de salud mental a través del análisis de la voz.

BIOMETRÍA DE LA VOZ



ALGORITMOS Y TÉCNICAS – TOKENIZACIÓN

- Consiste en segmentar los datos de entrada en una lista de palabras, a fin de simplificar el uso de las mismas en las distintas operaciones a realizar en el procesamiento.



ALGORITMOS Y TÉCNICAS – TOKENIZACIÓN

- Consiste en segmentar los datos de entrada en una lista de palabras, a fin de simplificar el uso de las mismas en las distintas operaciones a realizar en el procesamiento.



ALGORITMOS Y TÉCNICAS – TF IDF

- Term Frequency – Inverse Document Frequency es un mecanismo que permite asignar un valor de peso (relevancia) a las palabras que componen un texto, a fin de poder establecer la relación de esta con las demás y poder determinar el contexto mediante otros algoritmos.
- Para obtener este valor, se multiplica la cantidad de veces que se repite un término dentro del documento por el logaritmo de la relación entre la cantidad de documentos dividido entre la cantidad de documentos que contienen la palabra en cuestión.

$$w_{x,y} = tf_{x,y} \times \log\left(\frac{N}{df_x}\right)$$

TF-IDF

Term x within document y

$tf_{x,y}$ = frequency of x in y

df_x = number of documents containing x

N = total number of documents

	tfidf
had	0.493562
little	0.493562
tiny	0.493562
house	0.398203

ALGORITMOS Y TÉCNICAS – TF IDF

- Term Frequency – Inverse Document Frequency es un mecanismo que permite asignar un valor de peso (relevancia) a las palabras que componen un texto, a fin de poder establecer la relación de esta con las demás y poder determinar el contexto mediante otros algoritmos.
- Para obtener este valor, se multiplica la cantidad de veces que se repite un término dentro del documento por el logaritmo de la relación entre la cantidad de documentos dividido entre la cantidad de documentos que contienen la palabra en cuestión.

$$w_{x,y} = tf_{x,y} \times \log\left(\frac{N}{df_x}\right)$$

TF-IDF

Term x within document y

$tf_{x,y}$ = frequency of x in y

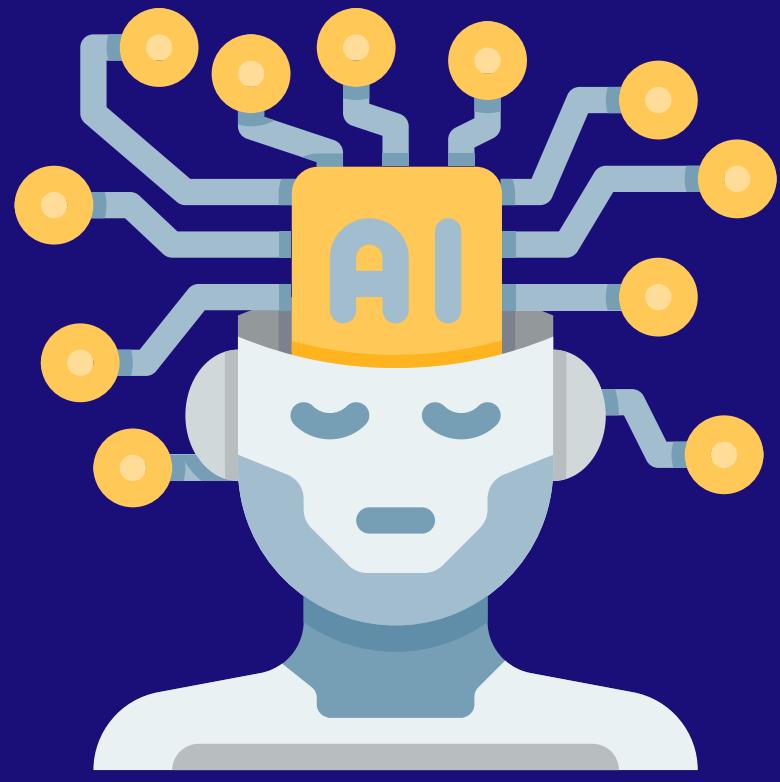
df_x = number of documents containing x

N = total number of documents

	tfidf
had	0.493562
little	0.493562
tiny	0.493562
house	0.398203

LIBRERÍAS DE NLP

The spacy logo is written in a large, bold, blue sans-serif font.The flair logo is written in a large, bold, black and orange sans-serif font.



Elementos gráficos

