



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA



DEPARTAMENTO
DE CIENCIAS DE LA
COMPUTACIÓN

Trabajo de Titulación previo a la obtención del título de Ingeniero en Tecnologías de la Información

Tema: "Framework de Reconocimiento de Emociones para el análisis de salud ocupacional en tiempo real, aplicando un Método Multimodal basado en Deep Learning".

Tutor:
Dr. Freddy Tapia León

Autores:
Mercedes Elizabeth Morejón Cevallos
Andrés Esteban Noboa Villacís



ANTECEDENTES

01



Puntos Críticos

01

La Inteligencia Artificial ha logrado interés en cuanto al reconocimiento de emociones de manera automatizada.

02

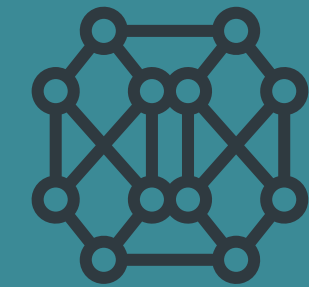
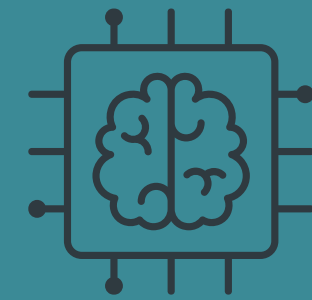
El comportamiento usual del ser humano permite considerar 6 emociones básicas como felicidad, sorpresa, enojo, asco, miedo y tristeza

03

Extracción de características basadas en: Redes de Atención Multimodal, Redes Neuronales Convolucionales, Memoria a Largo Plazo y Transformers.

04

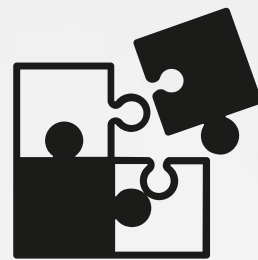
Determinación del estado afectivo y tendencia conductual de usuarios quien usan tecnología constante y que pueden llegar a presentar cuadros críticos de tecnoestrés.



PROBLEMA

02





Identificación de estado emocional a través de señales fisiológicas.



Uso excesivo de dispositivos tecnológicos



Cambios emocionales



Afectación en el ámbito de salud psicológica

¿Es posible determinar el estado de sensibilidad mediante reconocimiento de emociones en tiempo real aplicando un método multimodal y la utilización de técnicas de Aprendizaje Profundo (Deep Learning)?



OBJETIVOS

03



Objetivo General

■ Desarrollar un prototipo capaz de identificar y analizar las emociones de un individuo dentro del ámbito de salud ocupacional a través de un método multimodal, aplicando expresiones faciales, aspectos gestuales, no verbales y vocales logrando detectar el nivel de tecnoestrés con mayor precisión y fiabilidad en tiempo real bajo el uso de herramientas de Aprendizaje Profundo (Deep Learning).



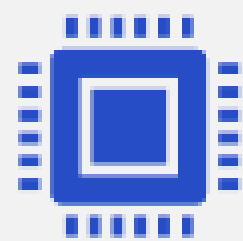
Objetivos Específicos



Realizar la revisión de literatura sobre el reconocimiento multimodal de emociones, identificando factores de riesgo que contribuyan al tecnoestrés a través de la identificación y clasificación de fuentes primarias.



Desarrollar un prototipo funcional bajo un método multimodal capaz de reconocer y autenticar el comportamiento de un individuo a través de la combinación de múltiples modalidades como reconocimiento facial, gestual, no verbal y vocal.



Implementar un sistema basado en computación afectiva en tiempo real, usando técnicas de Aprendizaje Profundo (Deep Learning) y procesamiento de señales fisiológicas mediante el cual se permita detectar, interpretar, procesar y responder ante la toma de decisiones y acciones de un individuo con la manipulación constante de medios tecnológicos.

Objetivos Específicos



Determinar arquitecturas de reconocimiento de Aprendizaje Profundo (Deep Learning) adecuadas para la inferencia de emociones faciales mediante técnicas de Hold-out y Cross Validation , que permitan comprender y modelar patrones de comportamiento humano bajo el uso de la herramienta colaborativa Google Colab .



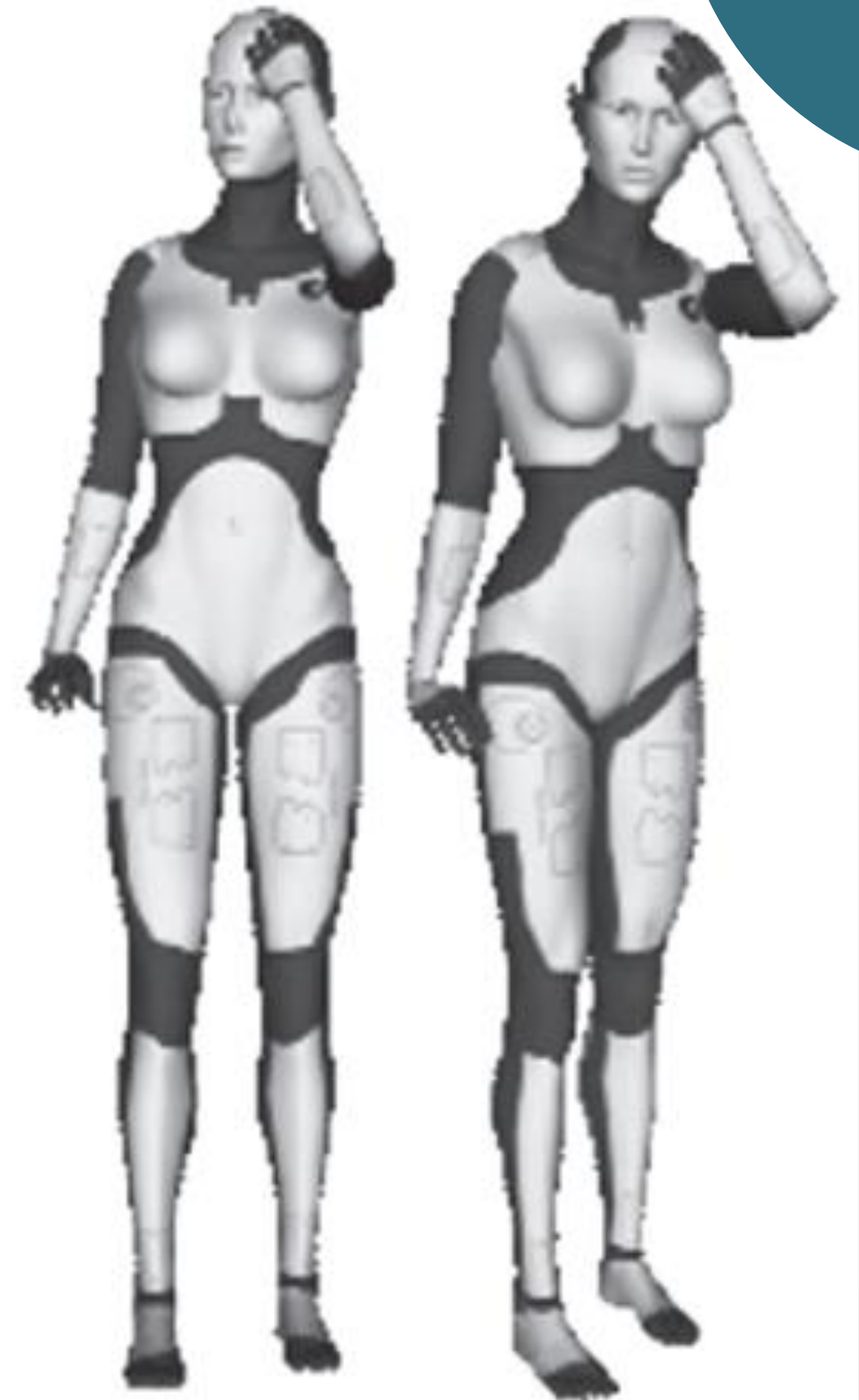
Evaluar el prototipo en un entorno laboral no controlado que permita mejorar la calidad de vida digital del personal en cuanto al uso equilibrado de la tecnología bajo contribución de un experto en salud ocupacional.

MODELO EXPERIMENTAL

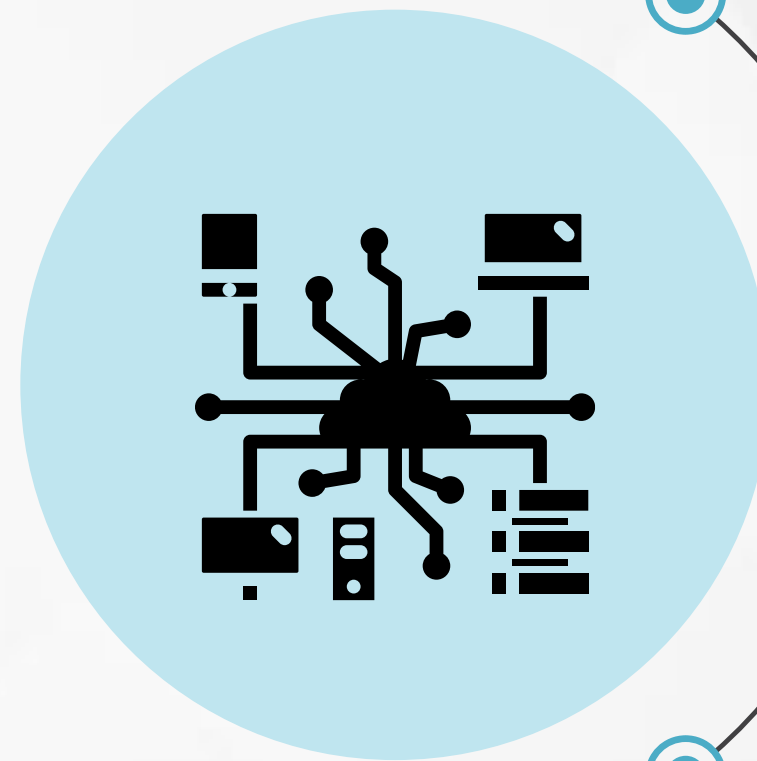
04


Computación Afectiva

Reconocer, procesar y simular el comportamiento humano con el uso de dispositivos tecnológicos



Computación Afectiva



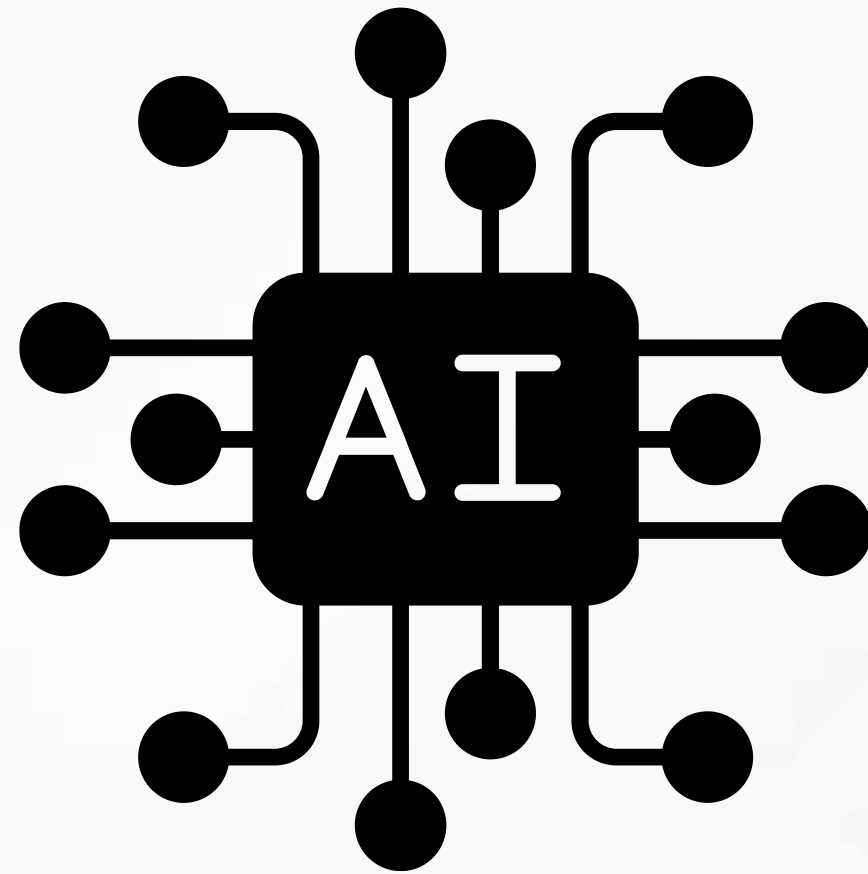
 Interacción
Humano
Computador

 Educación

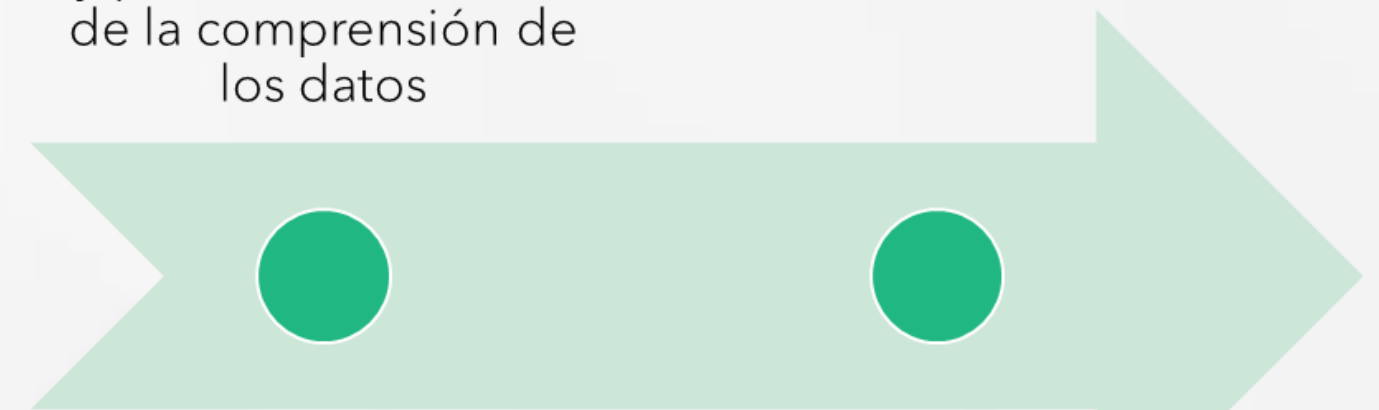
 Atención
Médica

Inteligencia Artificial - AI

“SE INTENTARÁ ENCONTRAR CÓMO HACER QUE LAS MÁQUINAS UTILICEN EL LENGUAJE, FORMEN ABSTRACCIONES Y CONCEPTOS, RESUELVAN TIPOS DE PROBLEMAS AHORA RESERVADOS A LOS HUMANOS Y SE MEJOREN A SÍ MISMAS” JOHN MCCARTHY

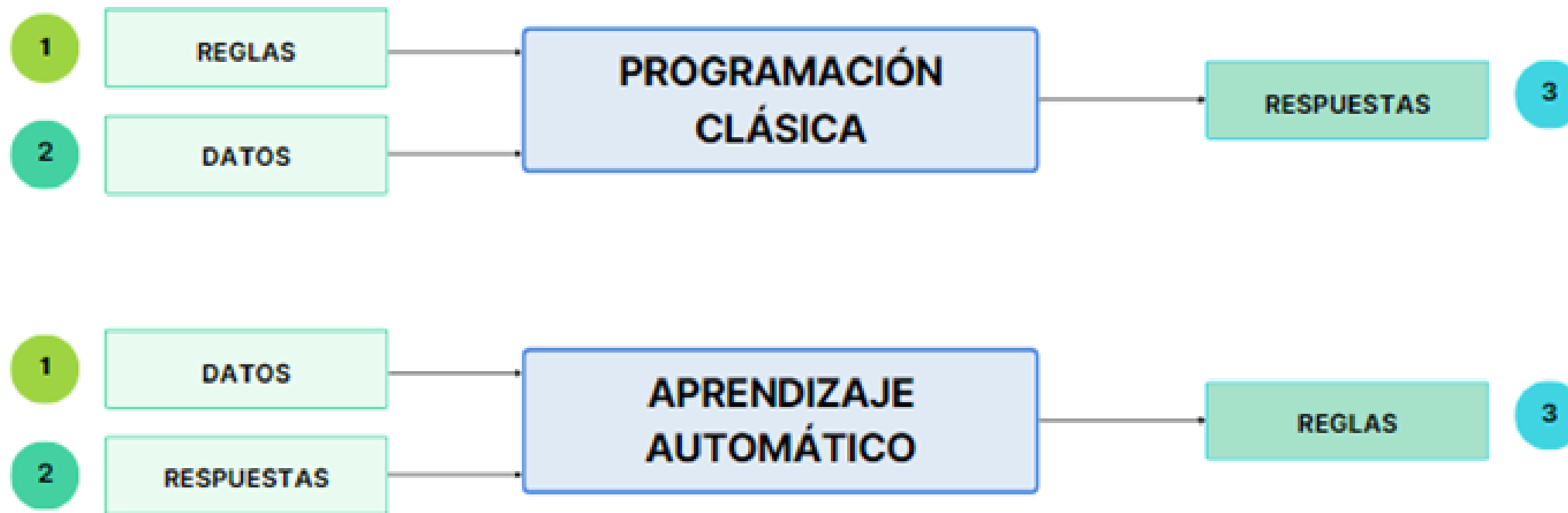


Consiste en la habilidad de los ordenadores de usar algoritmos, mejorarlos y procesarlos a través de la comprensión de los datos



Para la toma de decisiones tal cual el comportamiento humano

Machine Learning



“El estudio que otorga a las computadoras la capacidad de aprender sin ser programadas explícitamente” - Arthur Lee Samuel

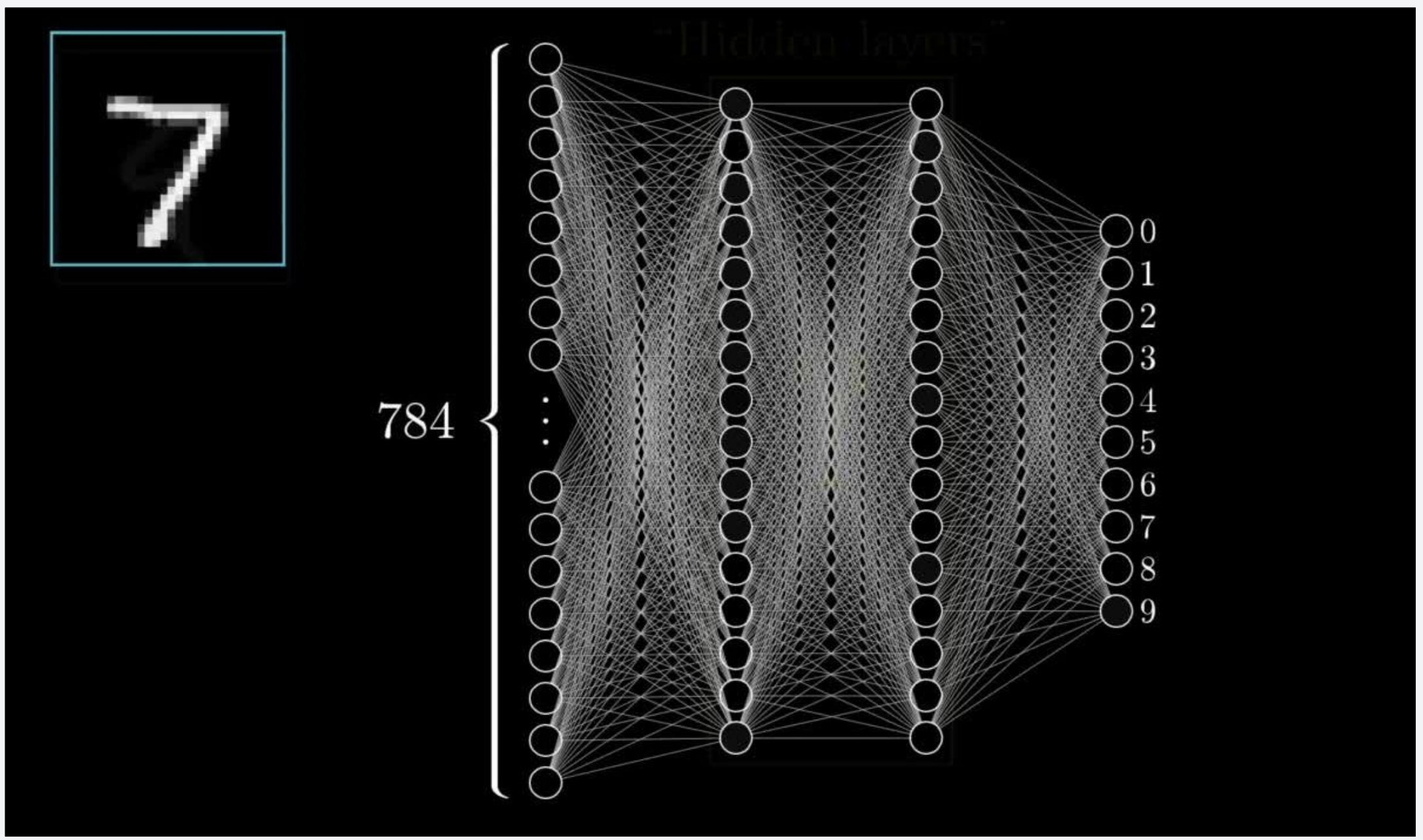
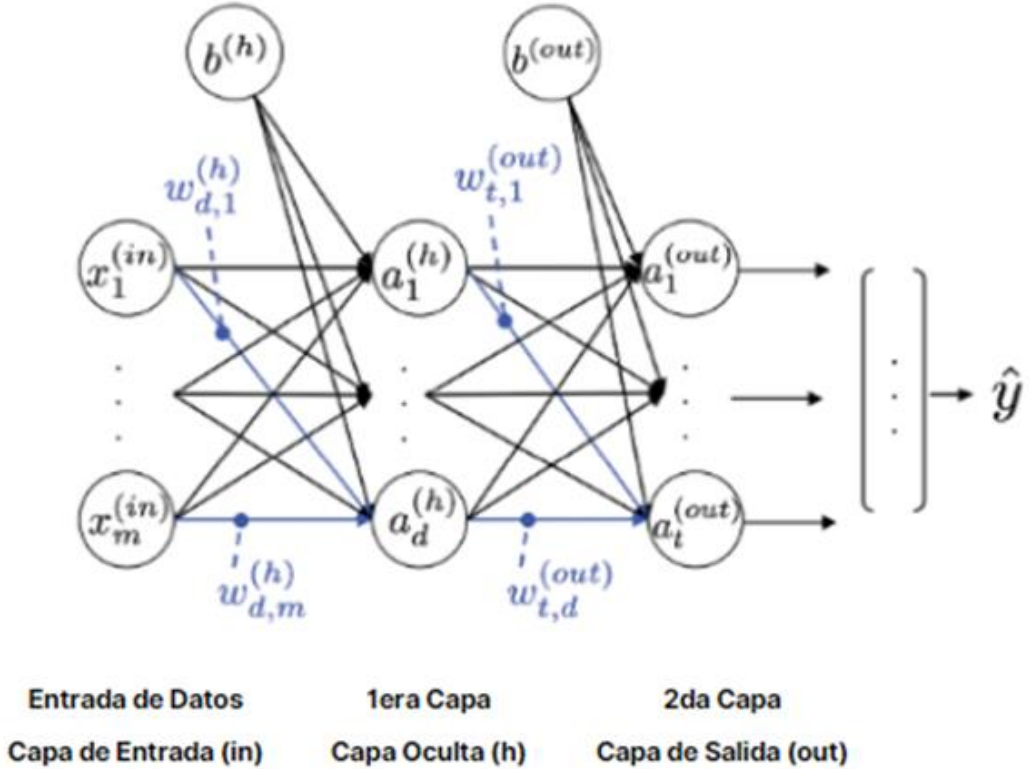
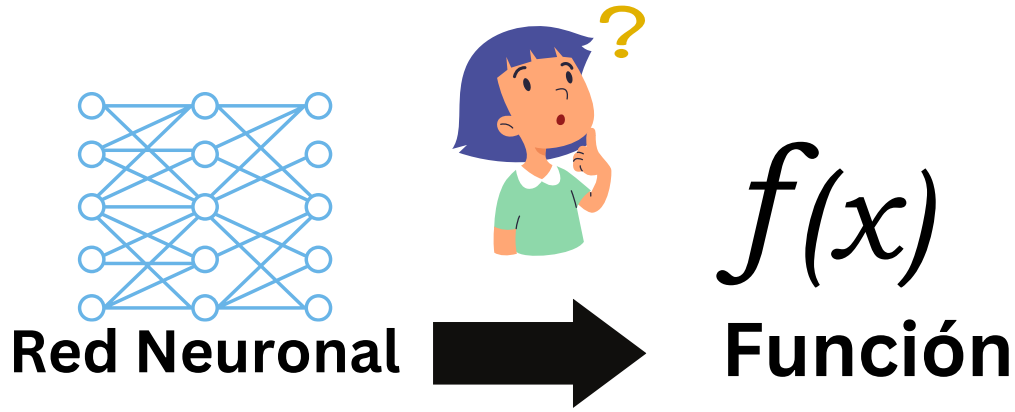
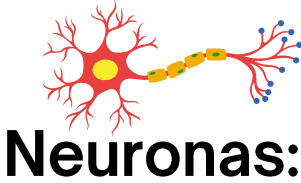
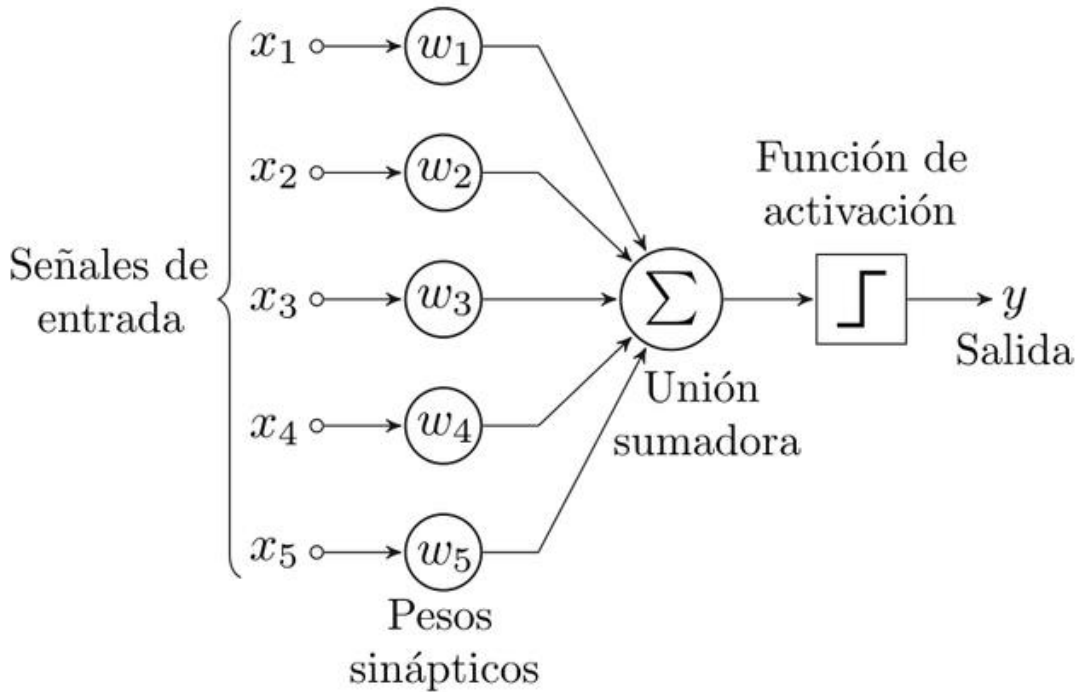


Un sistema de aprendizaje automático se “entrena” en lugar de programarse explícitamente.

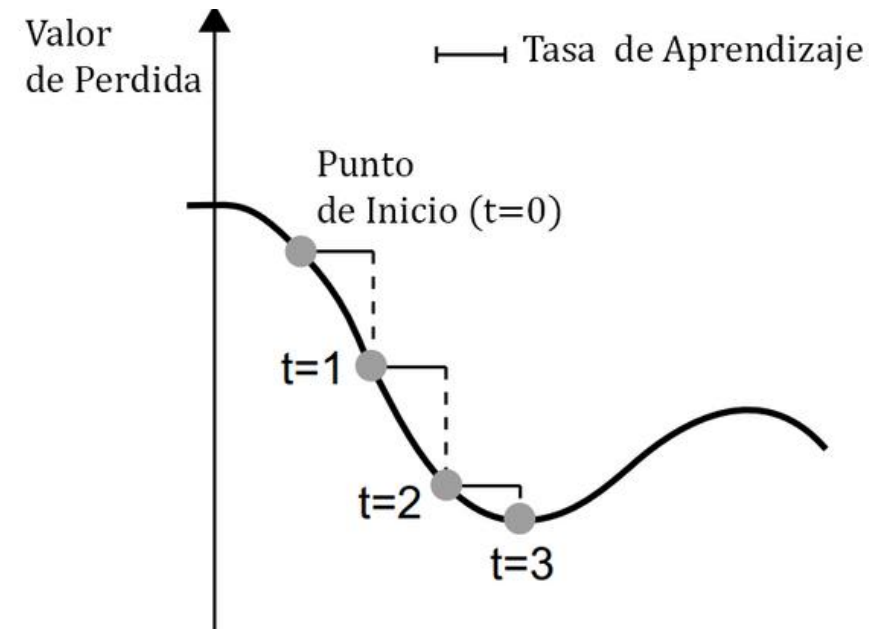


Se presenta con muchos ejemplos relevantes para una tarea y encuentra una estructura estadística en estos ejemplos que eventualmente permite al sistema generar reglas para automatizar las tareas.

REDES NEURONALES

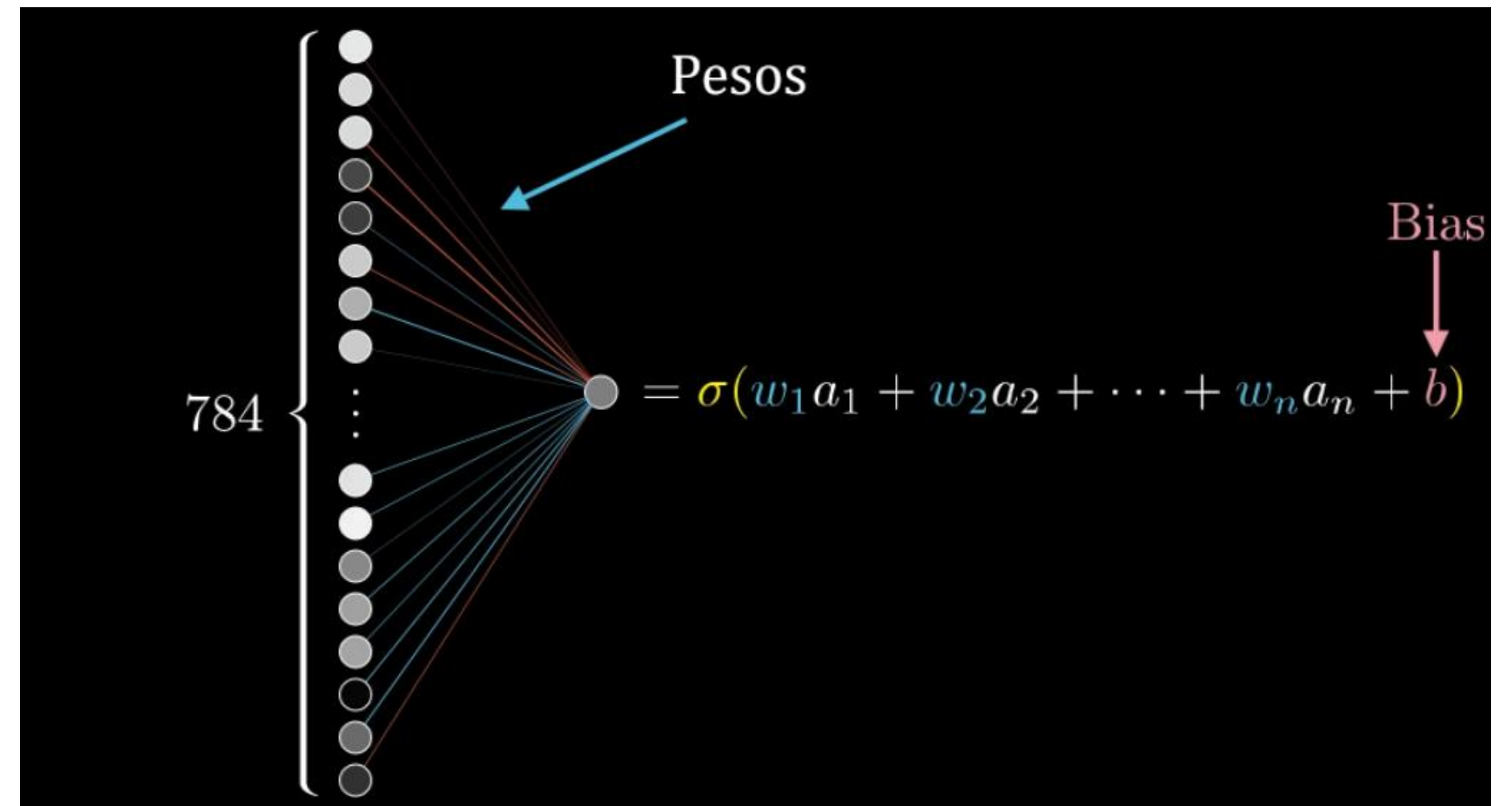
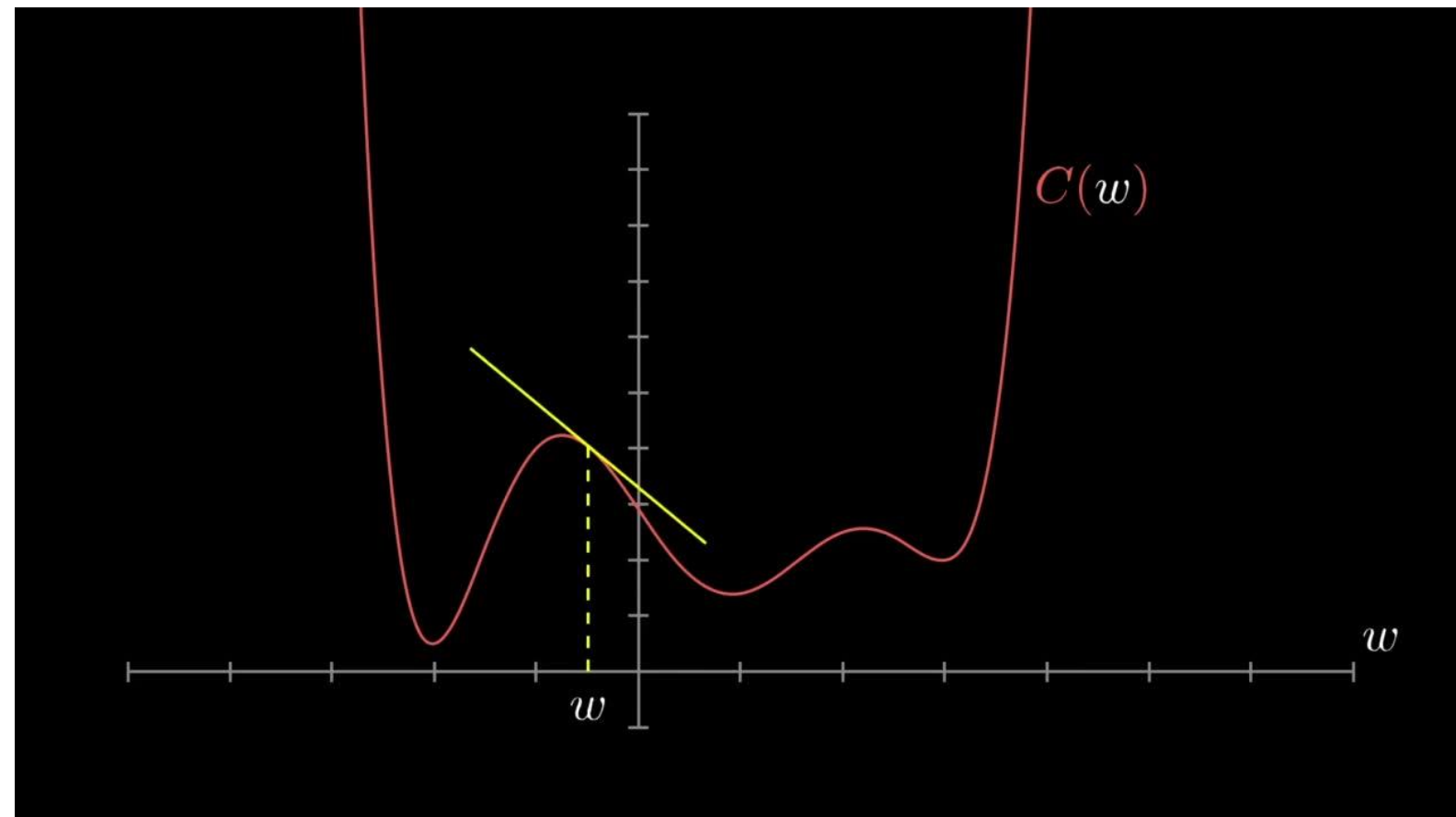
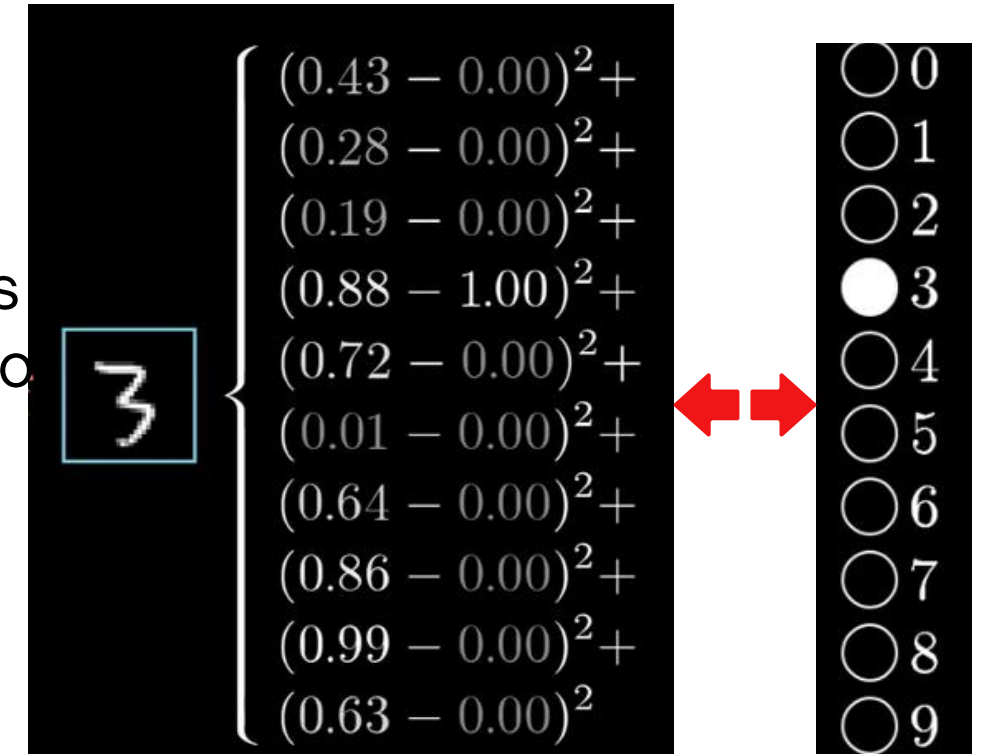


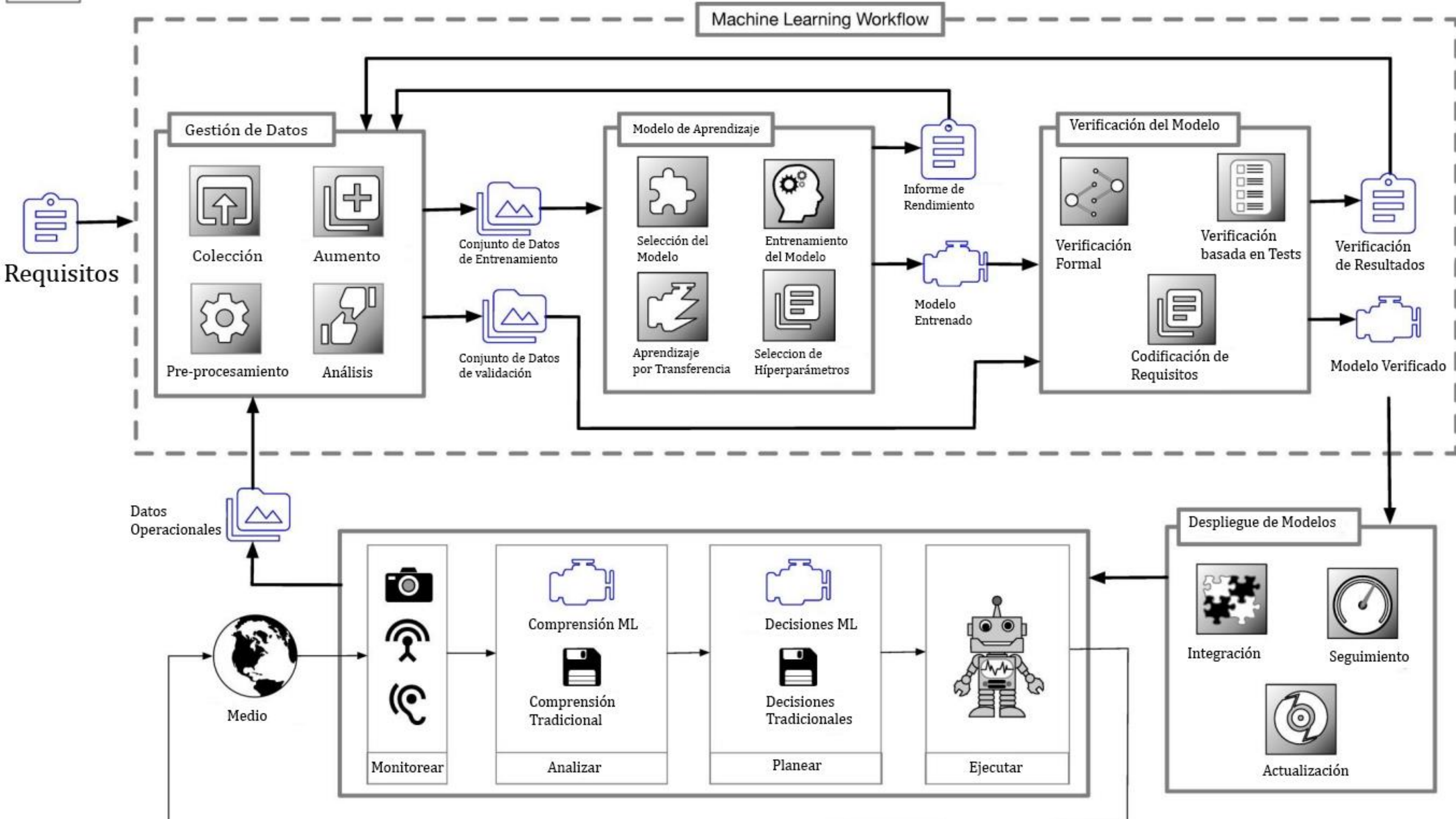
¿Como Aprenden realmente las Redes Neuronales? - Descenso del Gradiente



La función de costo: El "coste" se calcula sumando los cuadrados de las diferencias entre lo que obtuvimos y lo que queremos.

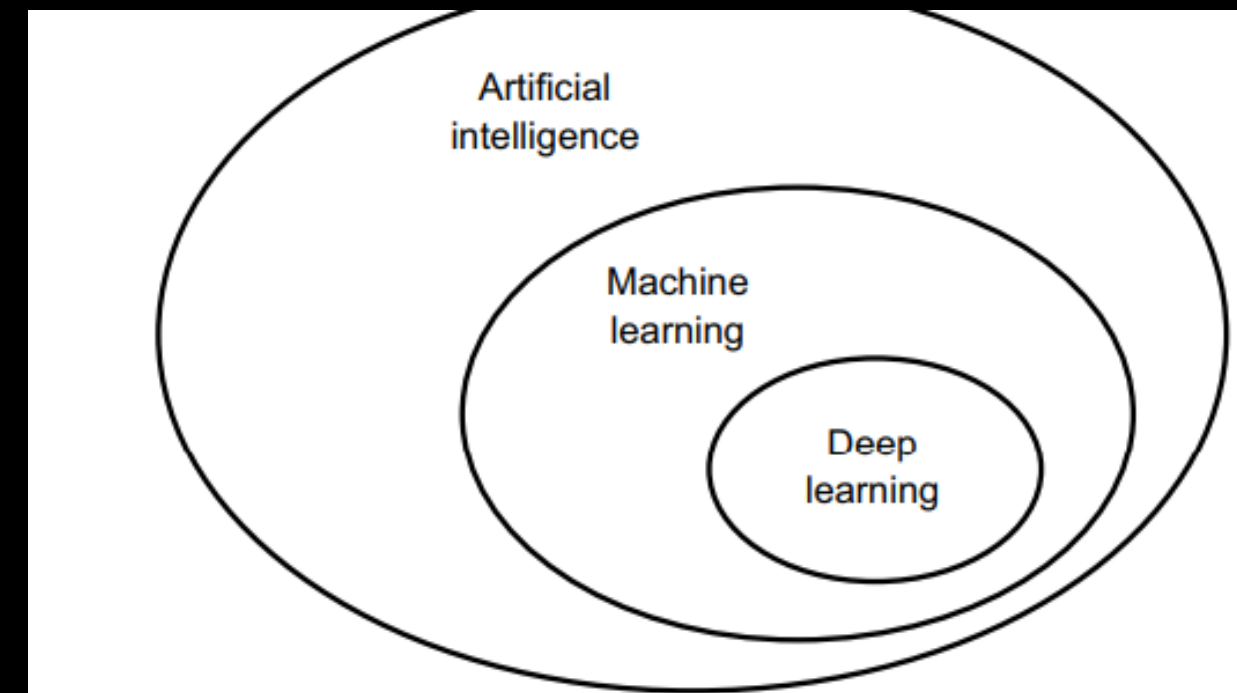
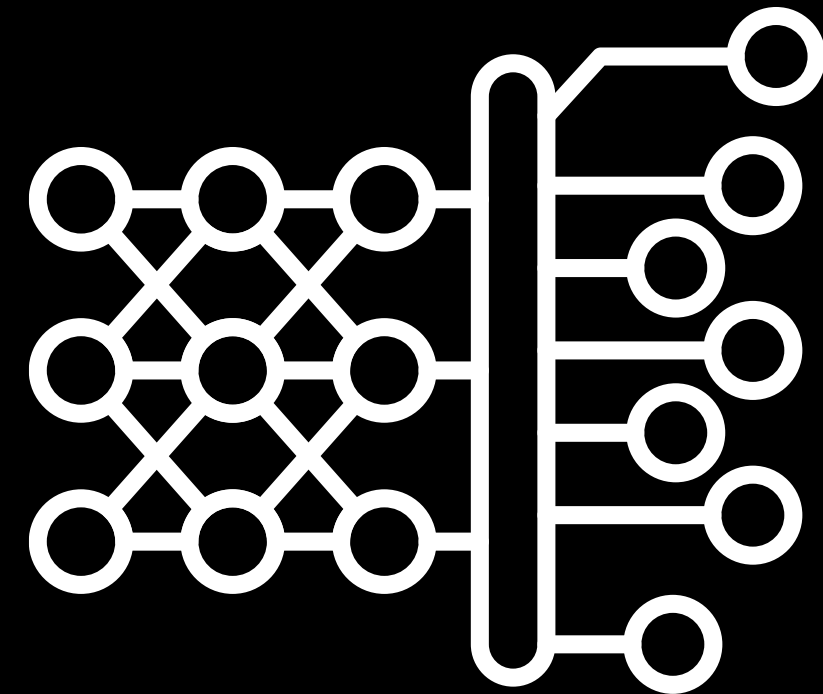
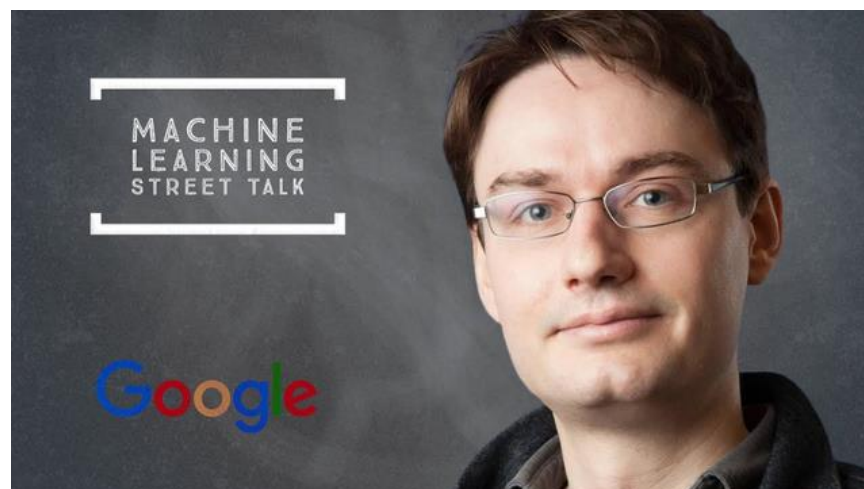
$$C = \frac{1}{m} \sum_{n=1}^m (y - y')^2$$



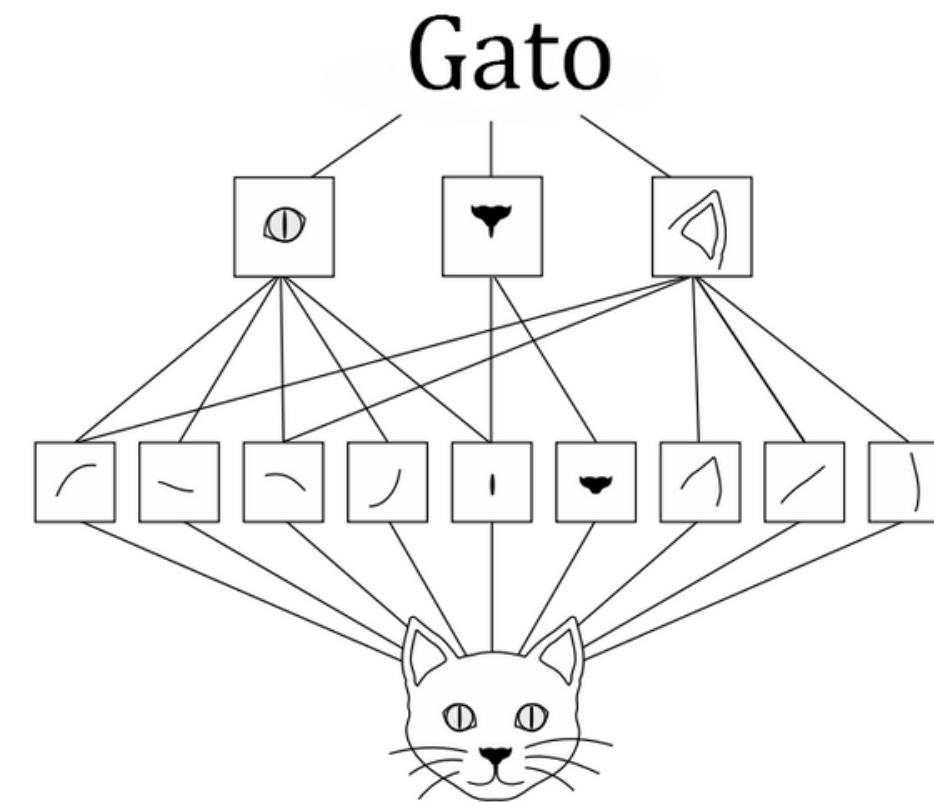
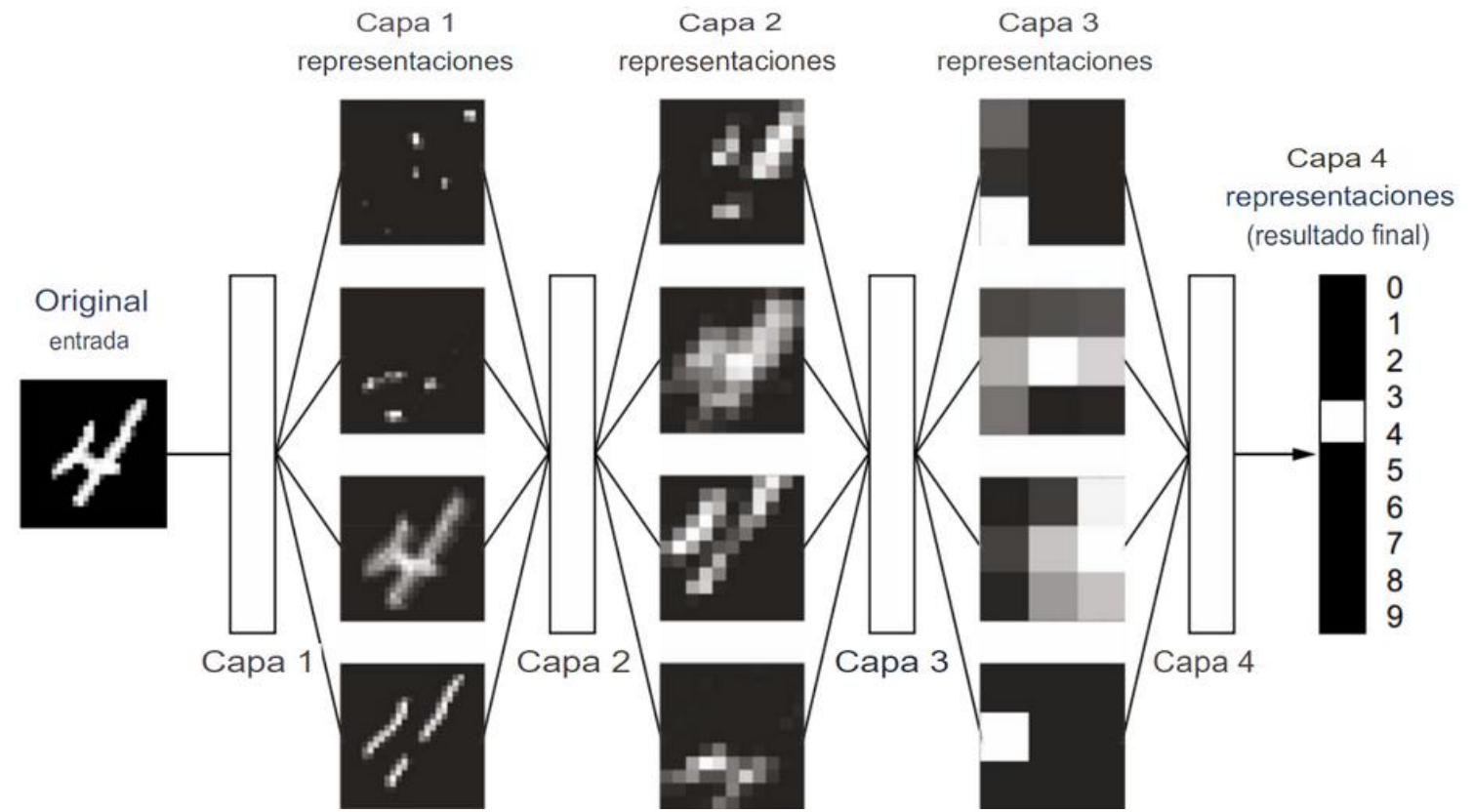
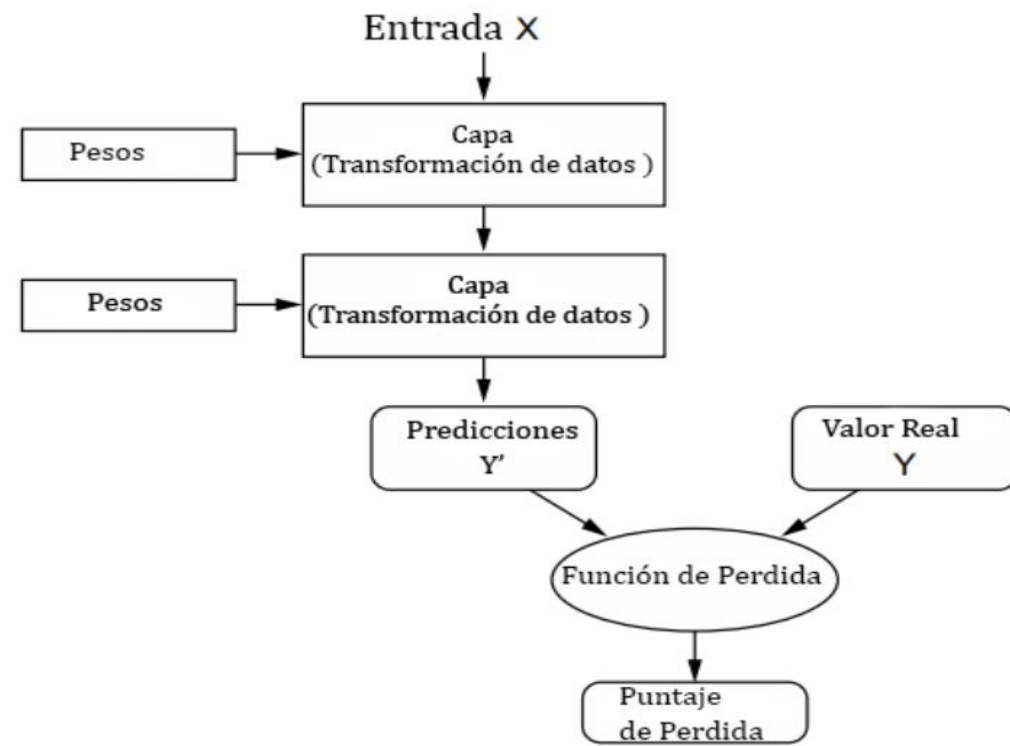


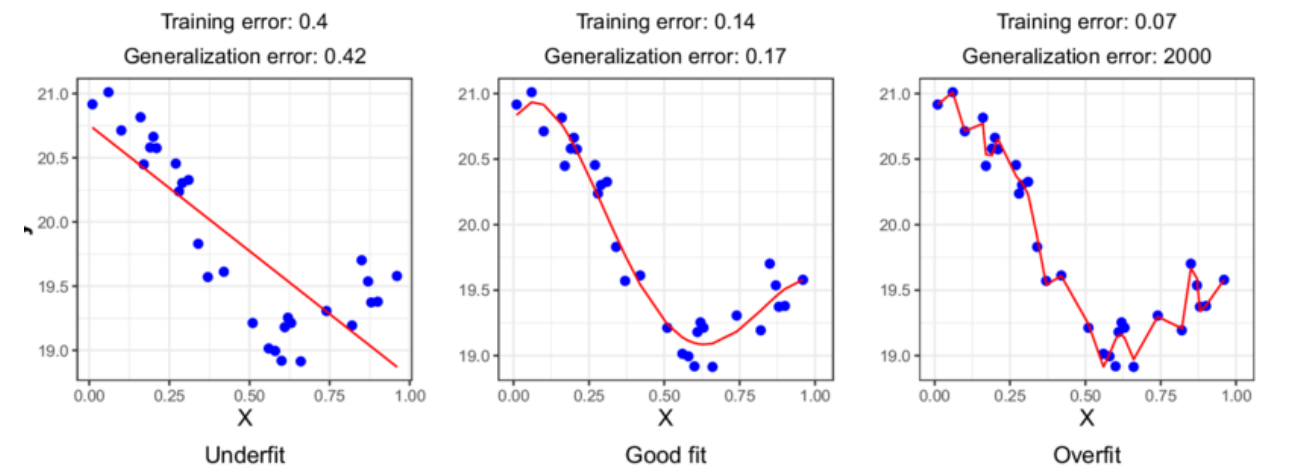
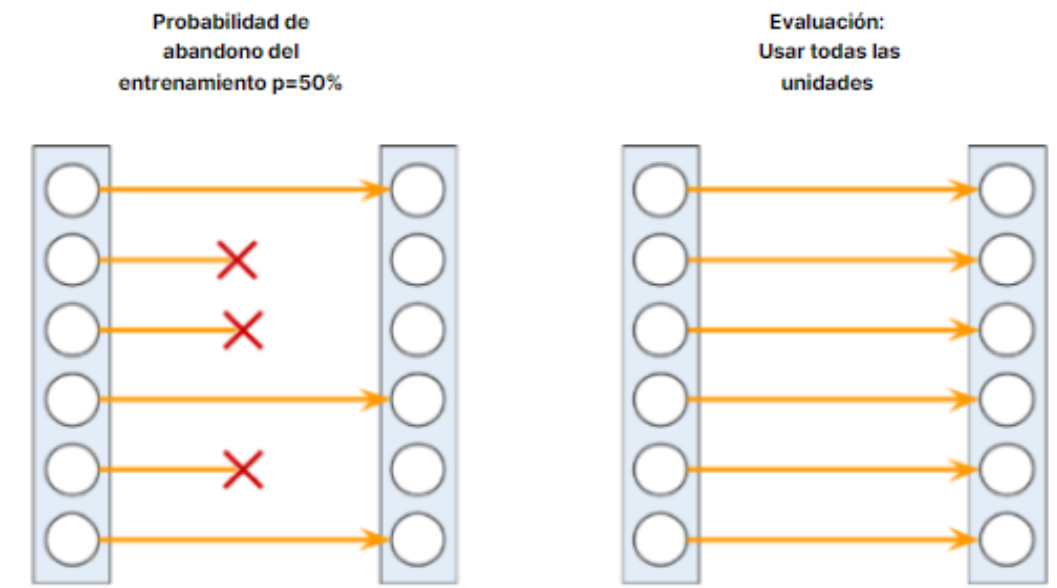
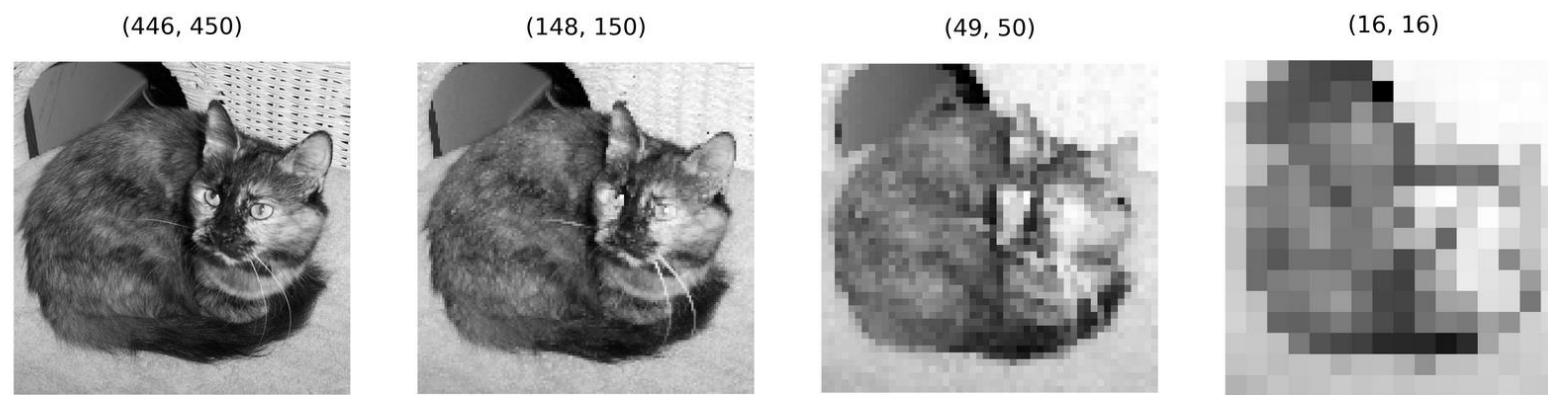
Deep Learning - DL

“El aprendizaje profundo es un subcampo específico del Machine Learning: una nueva visión del aprendizaje de representaciones a partir de datos que pone énfasis en el aprendizaje de capas sucesivas de representaciones cada vez más significativas.” - Francis Collet

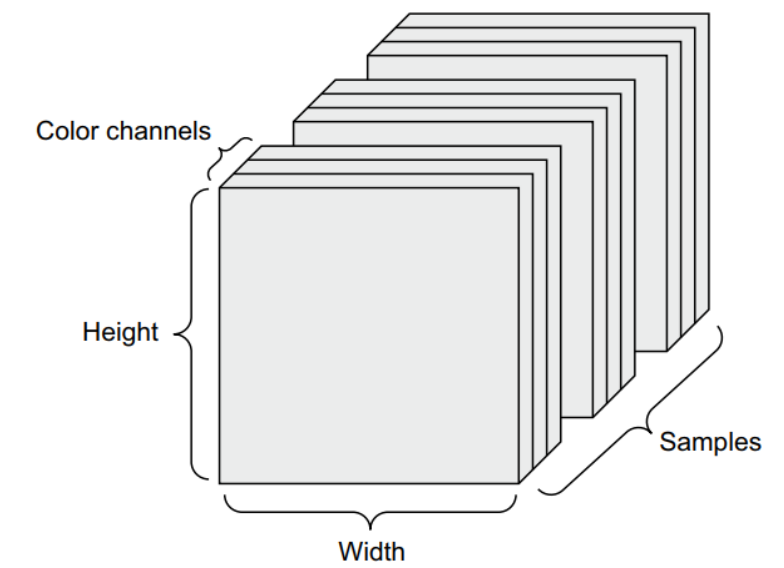
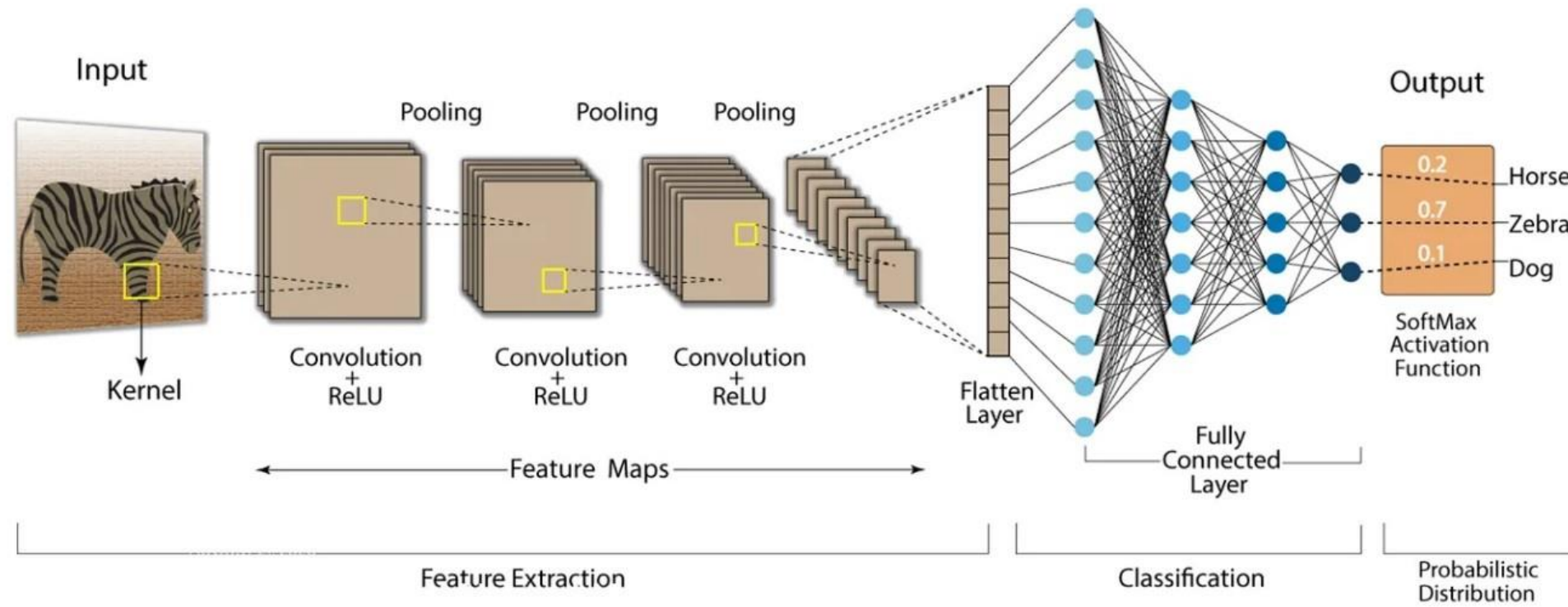


Deep Learning - DL

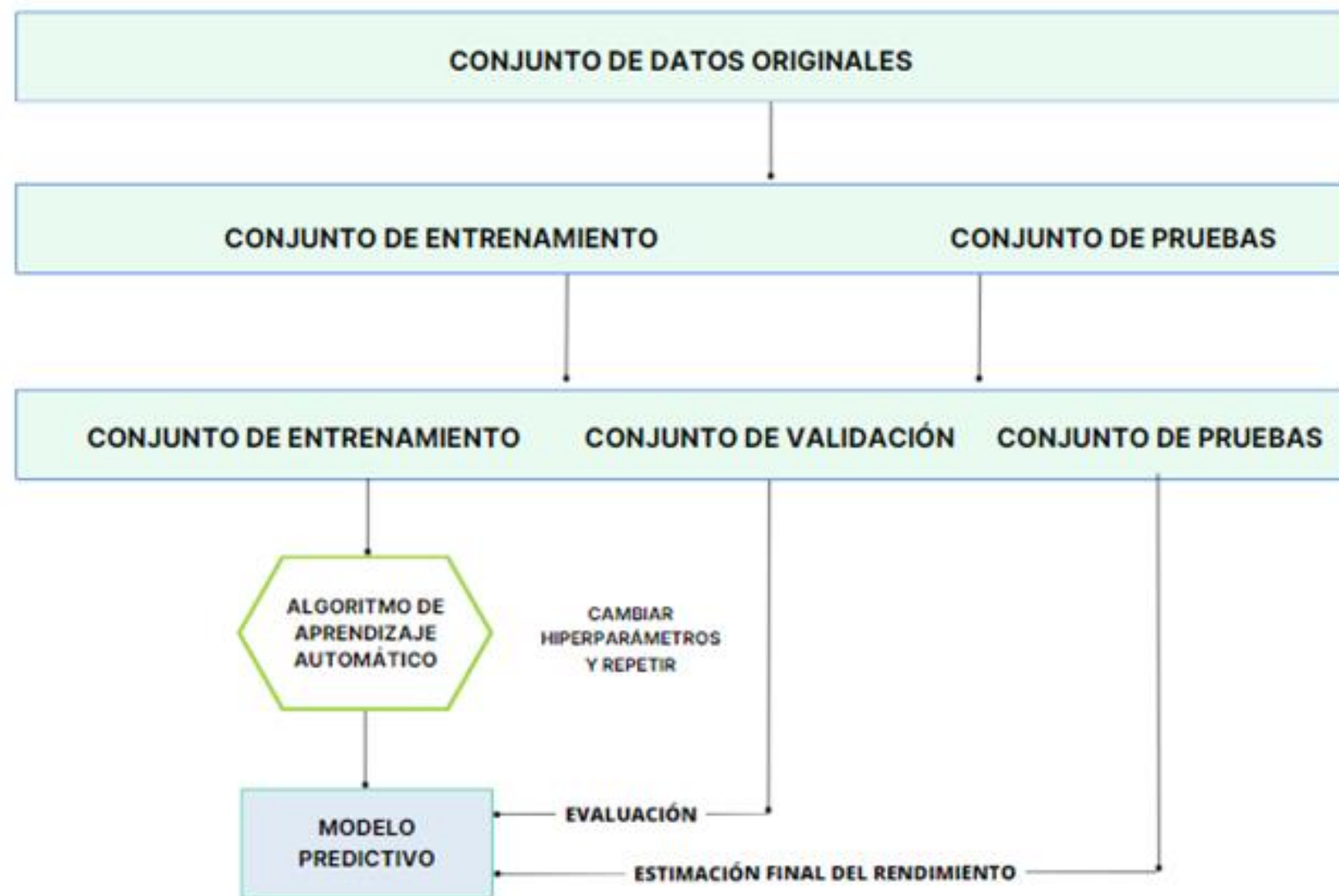




Convolution Neural Network (CNN)



Métricas de evaluación de modelos Machine Learning (Hold Out)



Estimación

		$C_0 \dots C_{k-1}$	C_k	$C_{k+1} \dots C_n$
Verdad Fundamental	$C_0 \dots C_{k-1}$	TN	FP	TN
	C_k	FN	TP	FN
	$C_{k+1} \dots C_n$	TN	FP	TN

TN Verdadero Negativo

TP Verdadero Positivo

FN Falso Negativo

FP Falso Positivo

$$Presición = \frac{TP}{TP + FP}$$

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

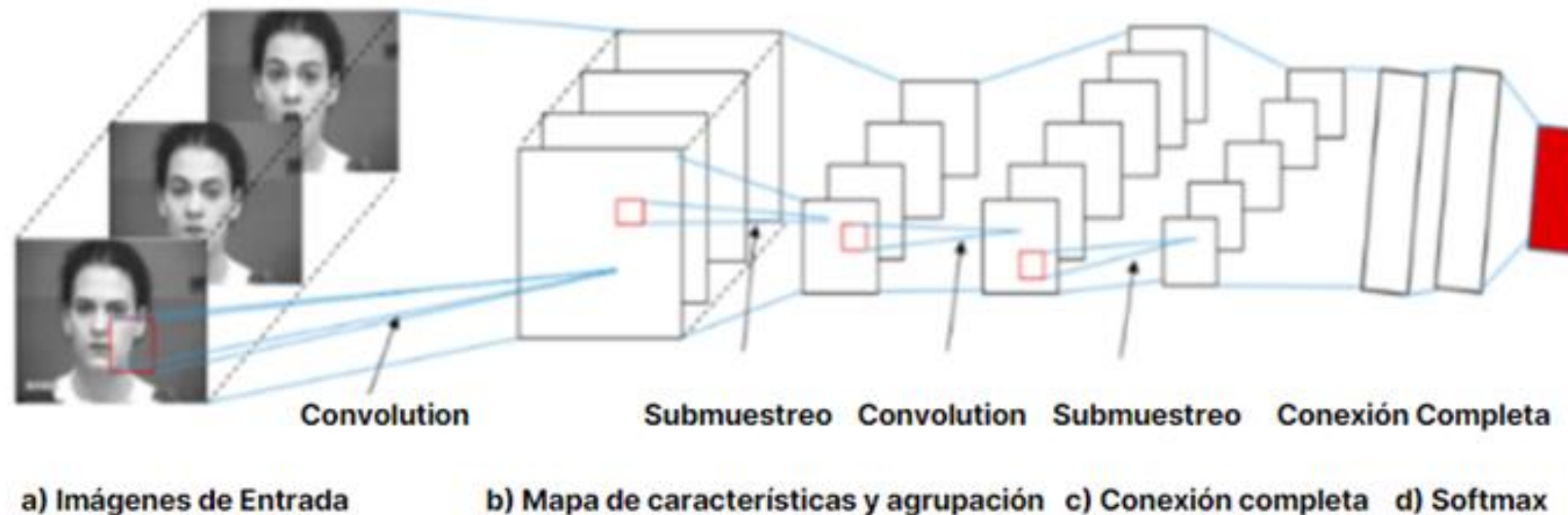
$$PuntajeF1 = \frac{2 * Precision * Sensibilidad}{Precision + Sensibilidad}$$

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

Reconocimiento de Emociones Faciales - FER

- El Reconocimiento de Emociones (FER) es usado para interpretar las emociones emitidas por un individuo.
- Es importante considerar que el ser humano es un ser cambiante en cuanto a sus expresiones emocionales independientemente de su entorno

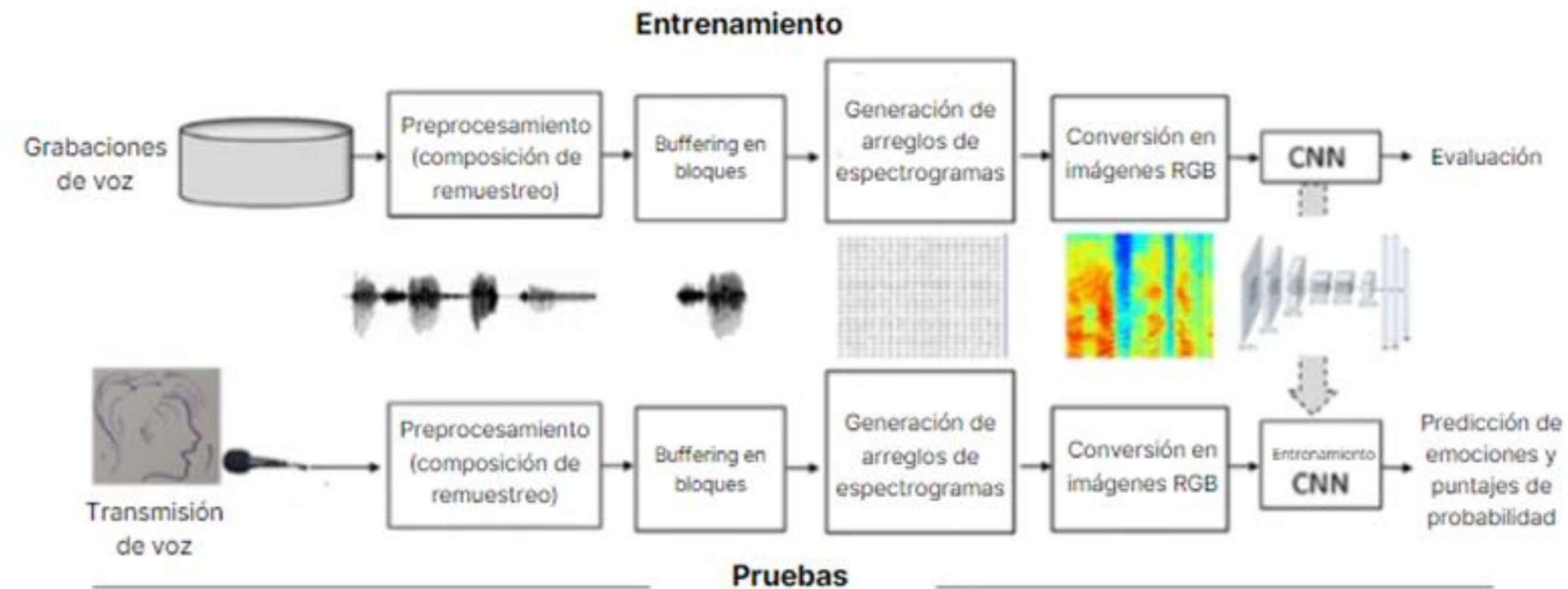
- La CNN ha logrado cubrir áreas que incluyen procesamiento de imágenes y reconocimiento de patrones que ayuden a determinar el estado de un individuo y su condición en un ambiente determinado.



Reconocimiento de emociones del habla - SER



Un sistema SER general permite la clasificación de las emociones mediante distintas partes de un individuo facilitando la comprensión del contenido verbal y la respuesta de manera más eficiente.



Procesamiento de Lenguaje Natural - NLP

1

El Procesamiento de Lenguaje Natural no obliga al sistema a aprender

2

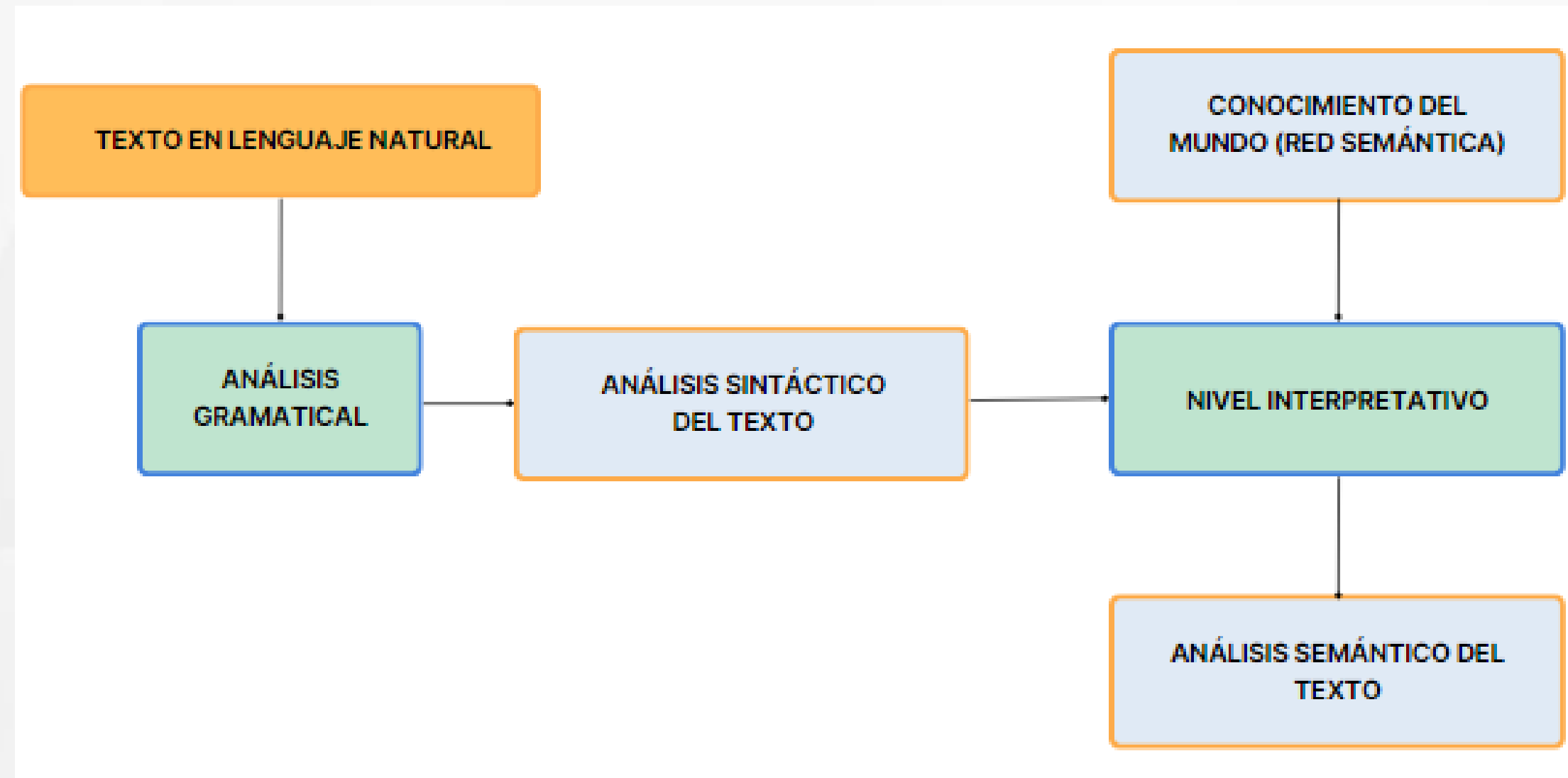
NLP incluye una interfaz amigable, intuitiva y sencilla de manejar.

3

Permite resolver las necesidades de los usuarios de forma automatizada

4

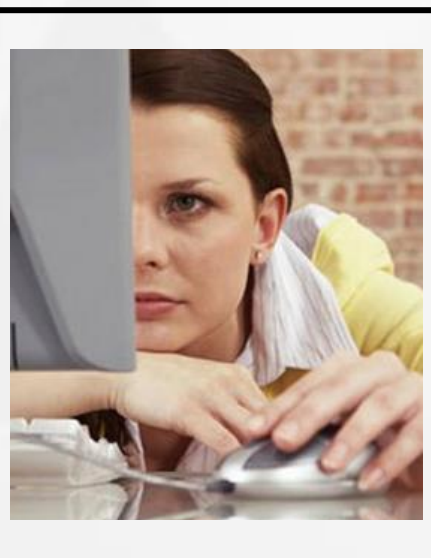
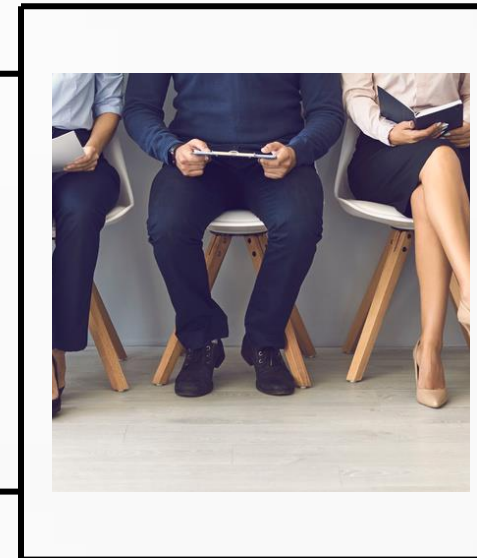
Comprende el proceso humano - máquina a fin de interpretar su comportamiento y automatizarlo.



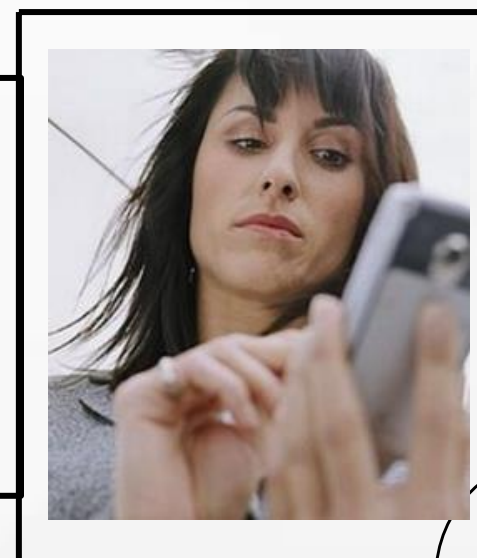
Tecnoestrés



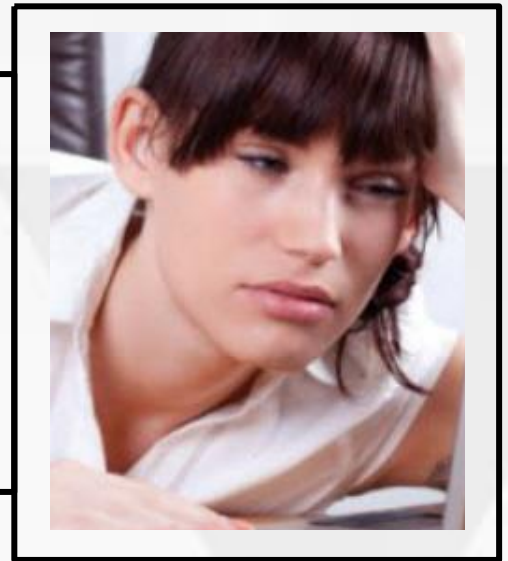
Tecnoestrés, concepto ligado directamente a efectos negativos comprendidos por el uso de TICs.



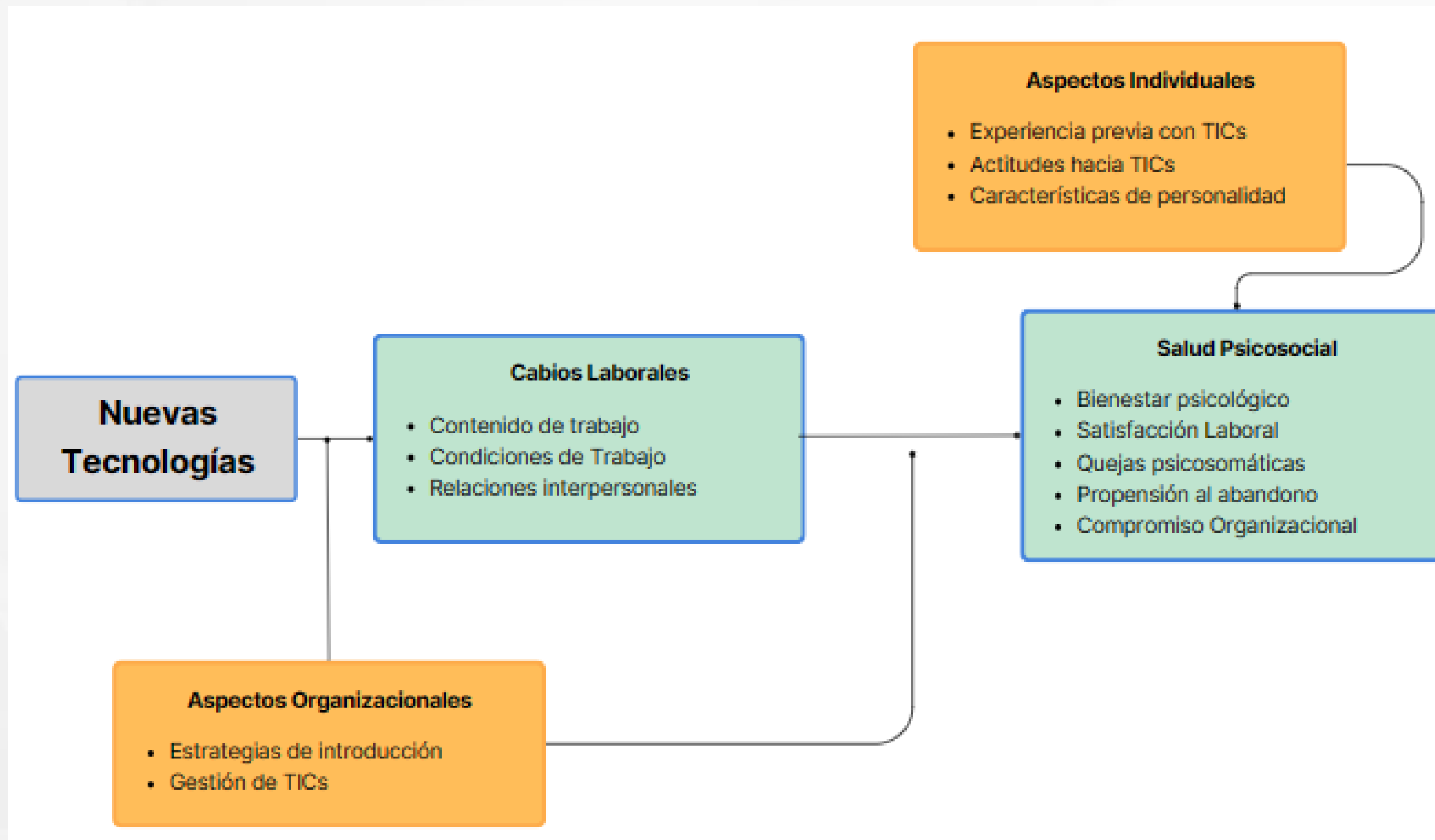
El Tecnoestrés está definido como el proceso ligado a la necesidad en niveles extremistas de estar conectado con el mundo tecnológico, logrando mayor impacto en relación con el cuadro de salud



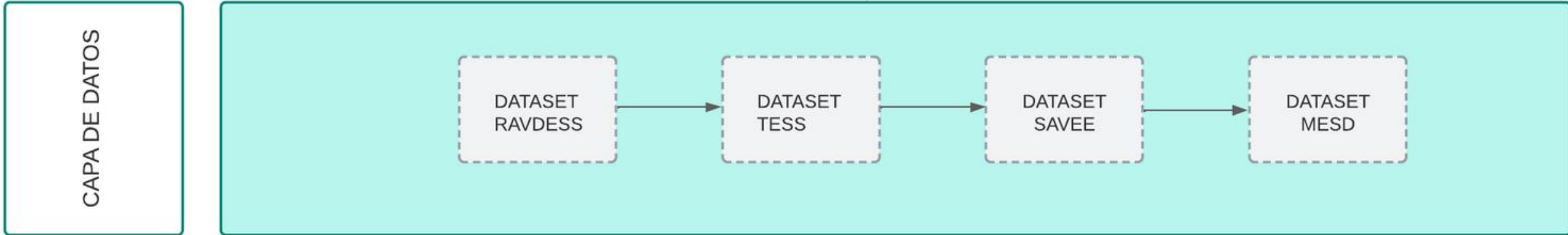
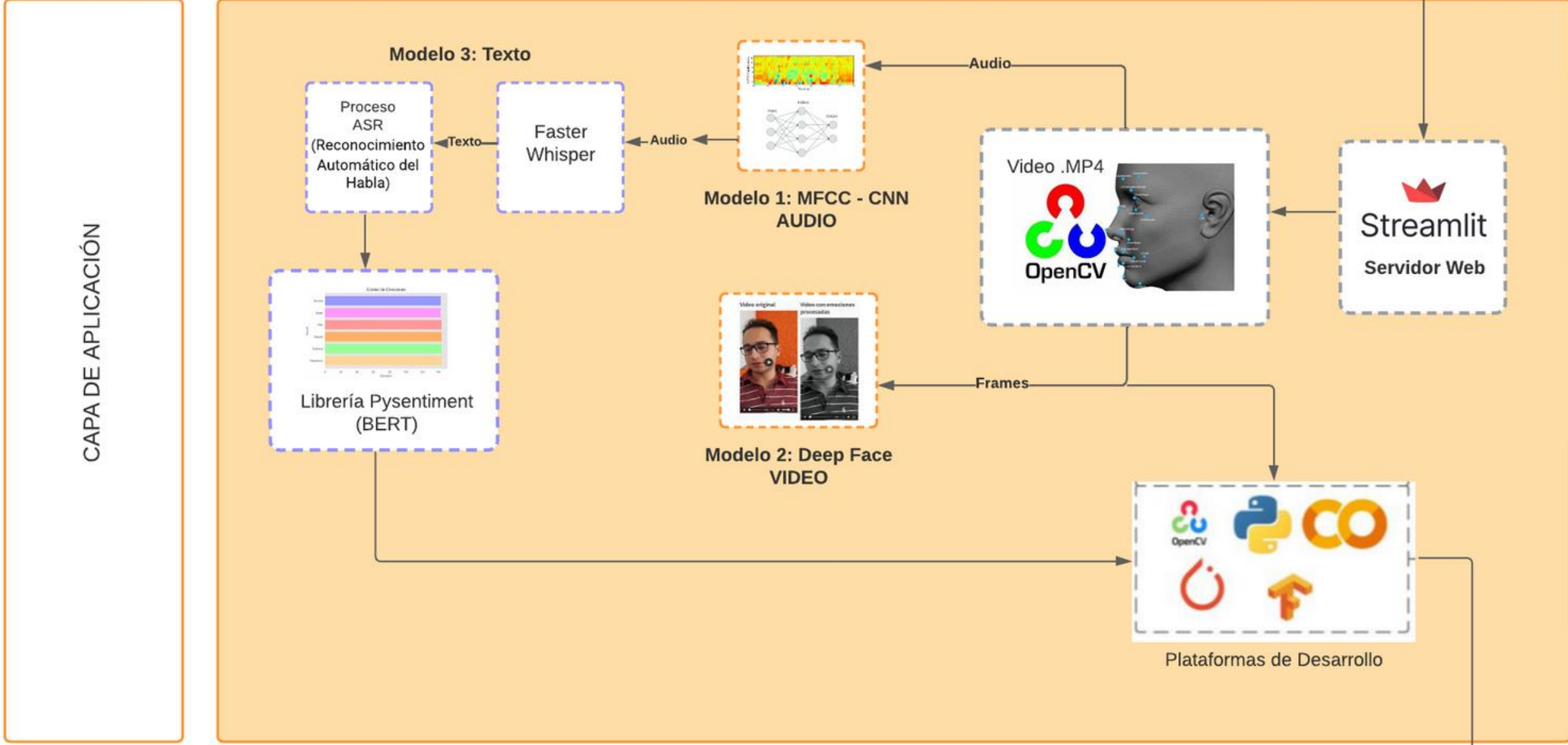
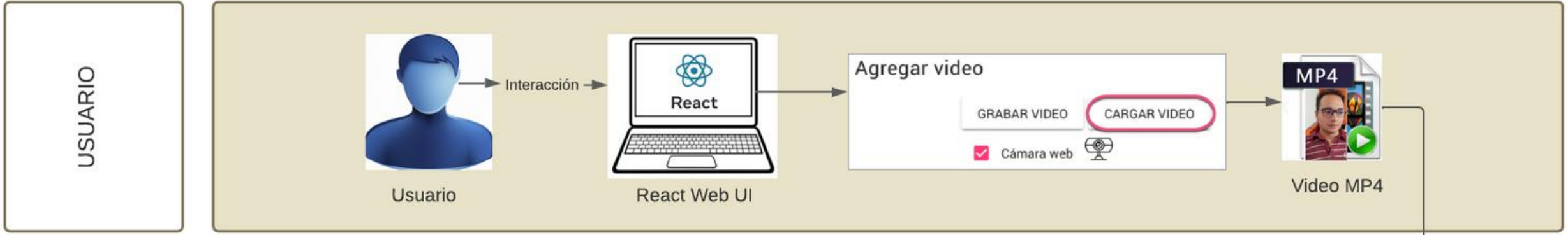
- **Dimensión Afectiva:** Comprende emociones y sentimientos.
- **Dimensión Actitudinal:** Cambios actitudinales del usuario.
- **Dimensión Cognitiva:** Eficiencia o Ineficiencia de un individuo.



Proceso de Tecnoestrés

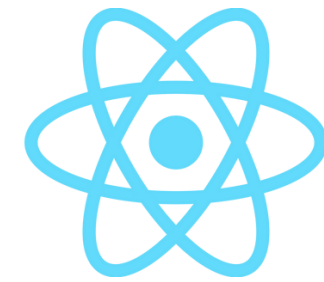
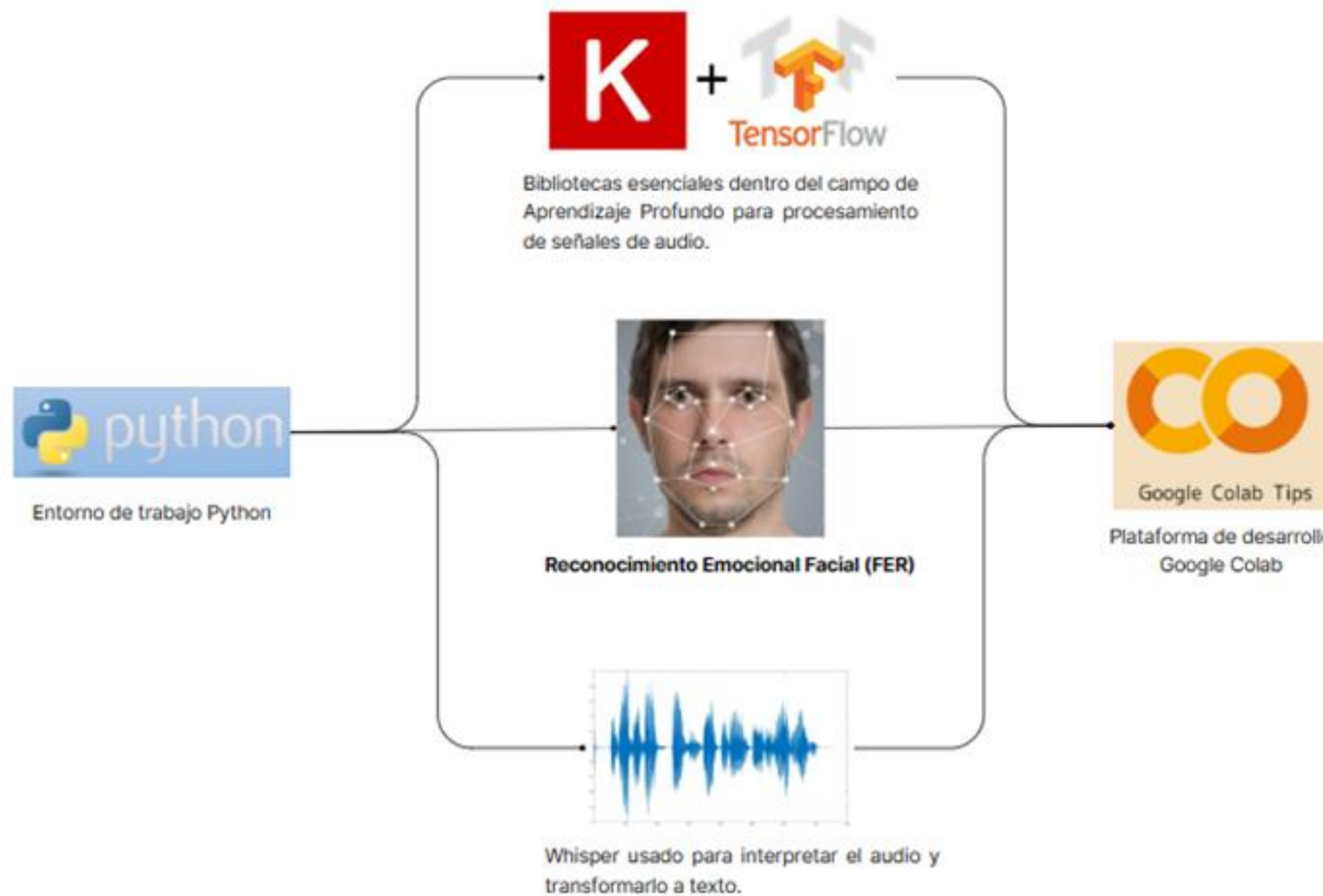


FUNCIONAMIENTO DEL PROTOTIPO.



ARQUITECTURA DEL SISTEMA PROPUESTO

Tecnologías Utilizadas



React



Python 3.9.7



Pytorch



OpenCV



Tensorflow

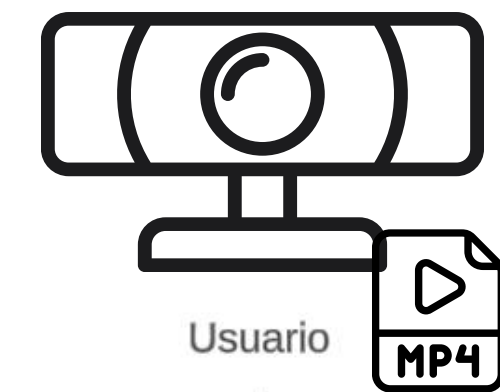


librosa

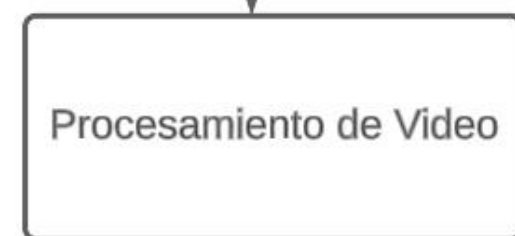


Streamlit

IMPLEMENTACIÓN DE MODELO DE RECONOCIMIENTO FACIAL



Video .mp4 / Entrada Webcam



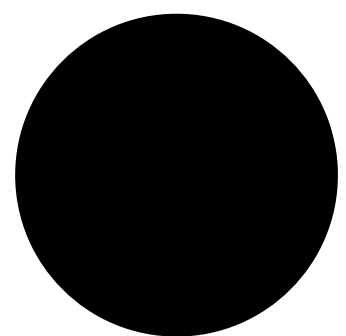
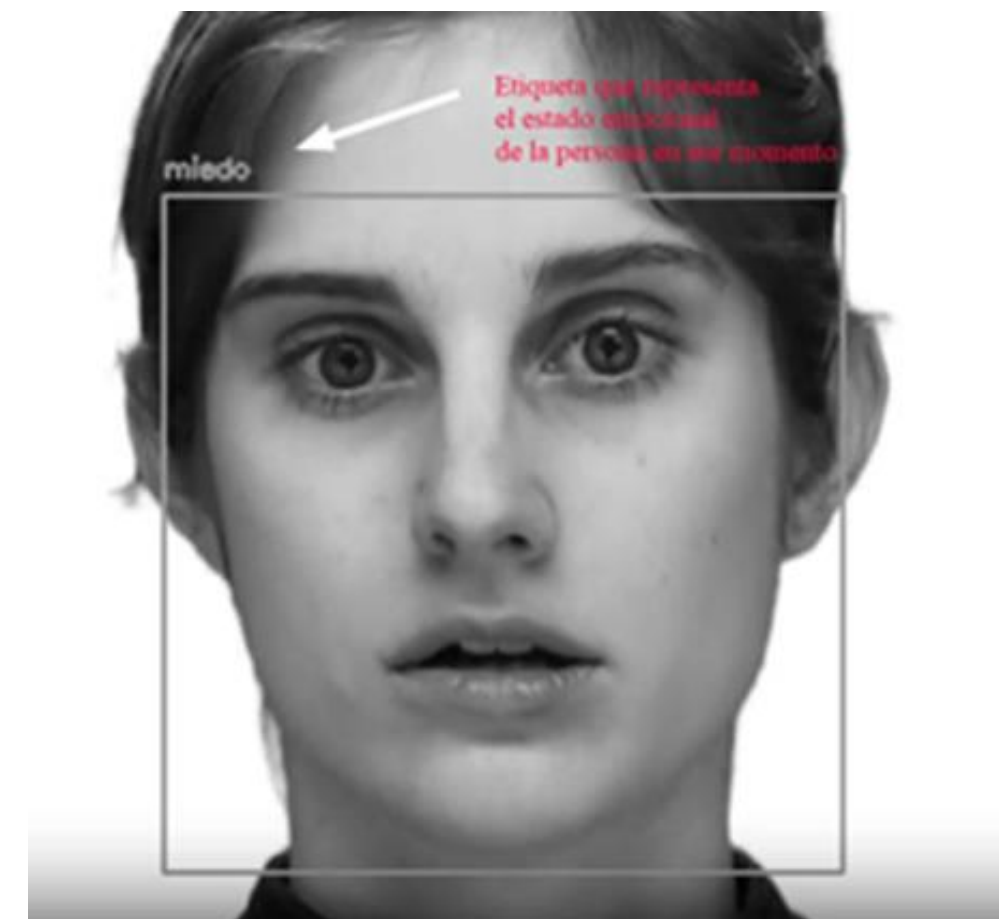
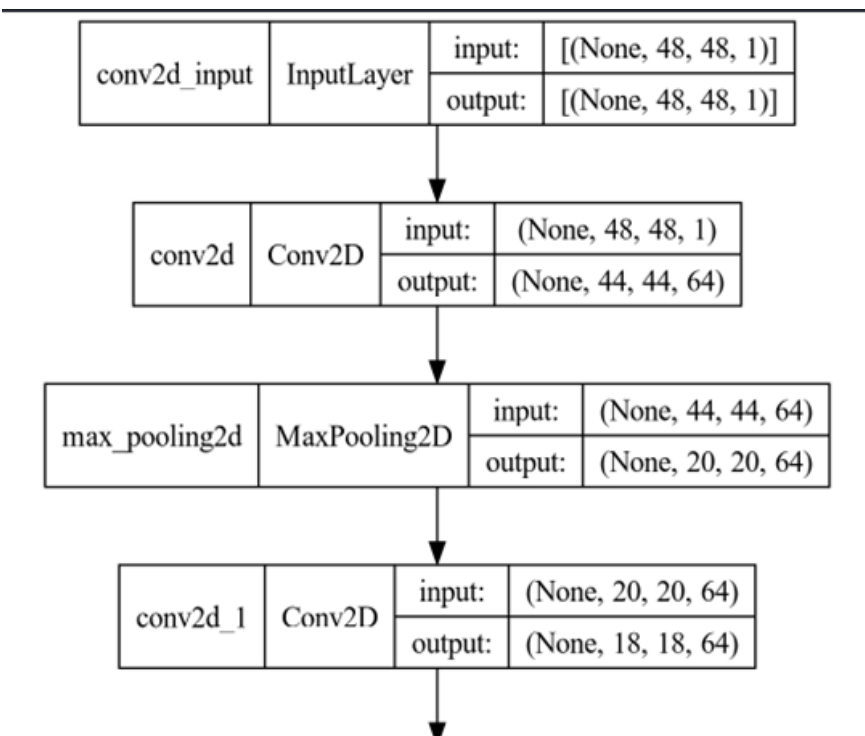
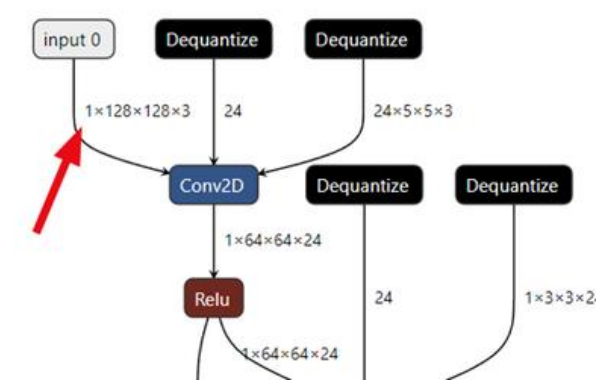
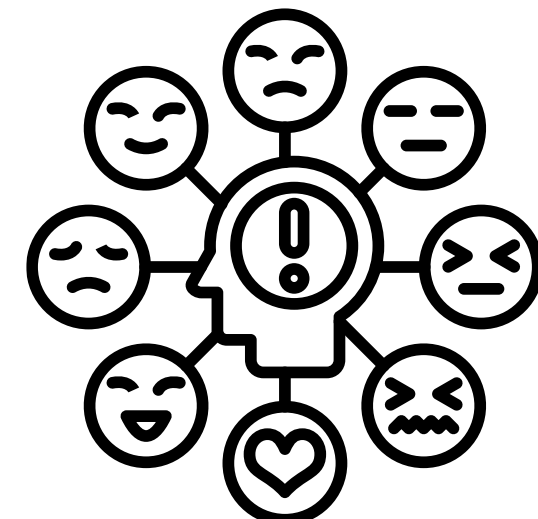
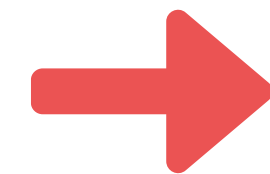
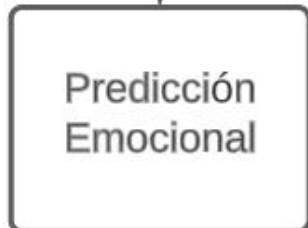
Fotogramas



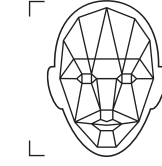
Fotogramas




Respuesta



DEMOSTRACIÓN (EMOCIONES FACIALES)





Menú

Menú

Reconocimiento de Emociones

Modelo

¿Cómo le gustaría hacer la predicción?

mfccs

Configuraciones

- 3 emociones
- 6 emociones
- 7 emociones
- Transcripción Whisper
- género

Cargar el archivo


Upload file

Drag and drop file here
Limit 200MB per file • MP4, MOV, AVI, MPEG4



Browse files

Video P12 (9).mp4 8.7MB

Video original




Video con emociones procesadas



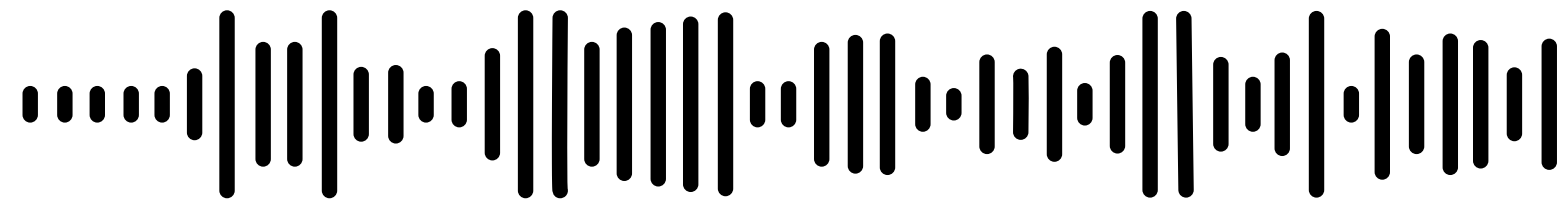
Línea Temporal de Emociones

sorpresa

enojado



DATOS UTILIZADOS EN EL ENTRENAMIENTO



01

MESD

Mexican Emotional Speech Database : 864 archivos de audio en WAV, 3 voces femeninas, 2 masculinas y 6 infantiles



02

TESS

Toronto emotional speech set: 2800 archivos, dos actrices femeninas (26 y 64 años)



03

SAVEE

Surrey Audio-Visual Expressed Emotion: 480 expresiones en inglés interpretadas por 4 actores masculinos



Angry

Happy

Disgusted

Surprised

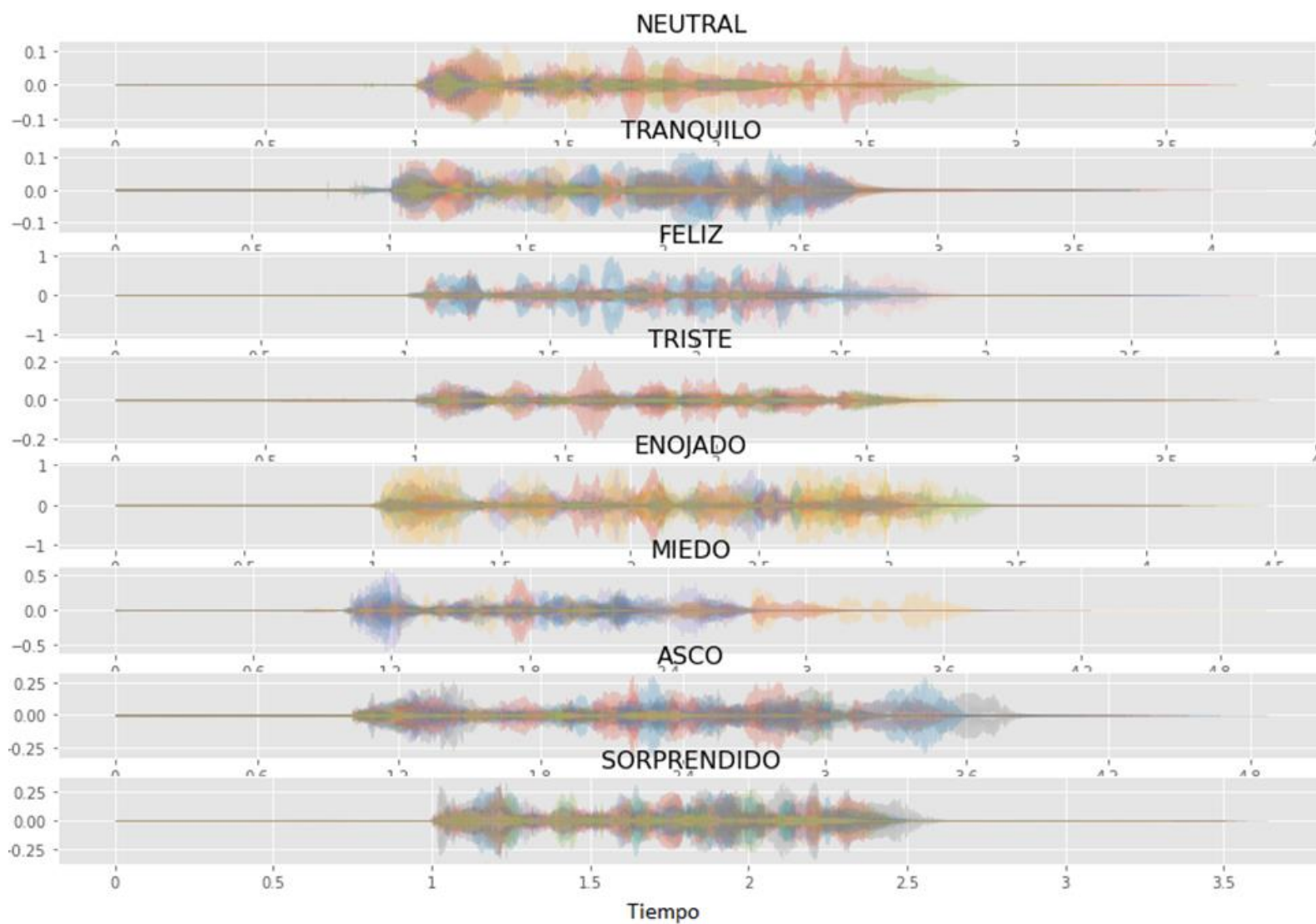
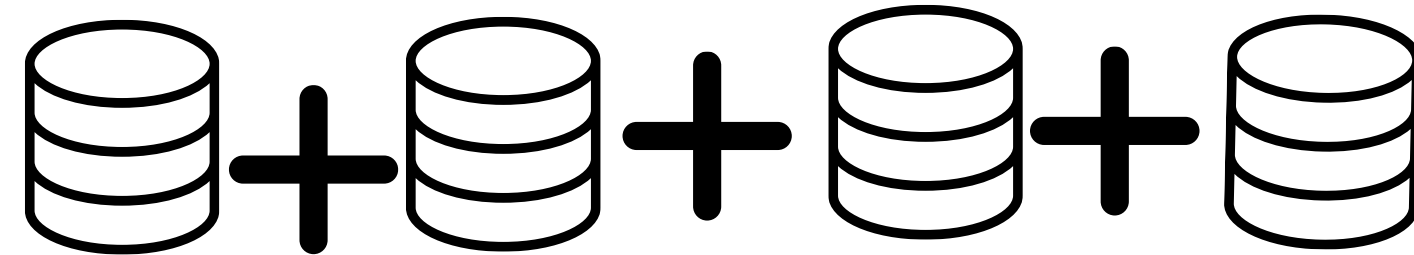
Calm

04

RAVDESS

The Ryerson Audio-Visual Database of Emotional Speech and Song : 7356 grabaciones de 24 actores

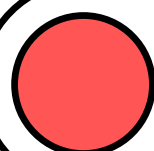
PRE-PROCESAMIENTO DE DATOS



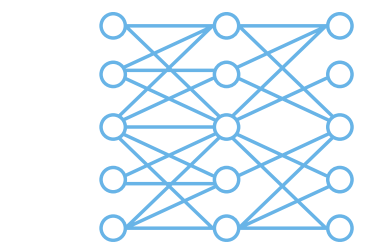
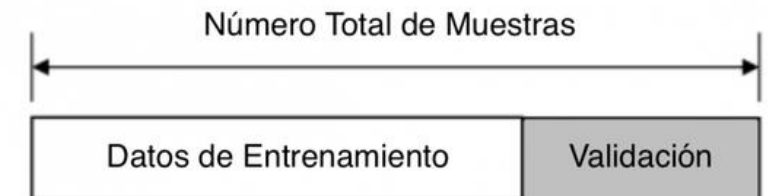
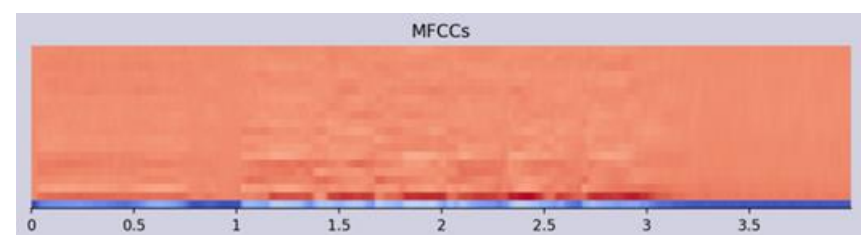
```
emotion_label,source,actors,path,emotion2,emotion3
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_bog_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_youth_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_phone_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_rush_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_anna_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_gym_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_deep_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_nice_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_judge_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_burn_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_weak_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_raise_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_voice_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_lid_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_hole_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_laugh_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_juice_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_goal_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_hate_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_bog_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_bean_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_wife_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_mob_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_fall_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_get_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_hurl_fear.wav,negative_female,fear_female
fear_female,TESS,female,./data/TESS/DMF_Fear/DMF_such_fear.wav,negative_female,fear_female
```

6737 registros



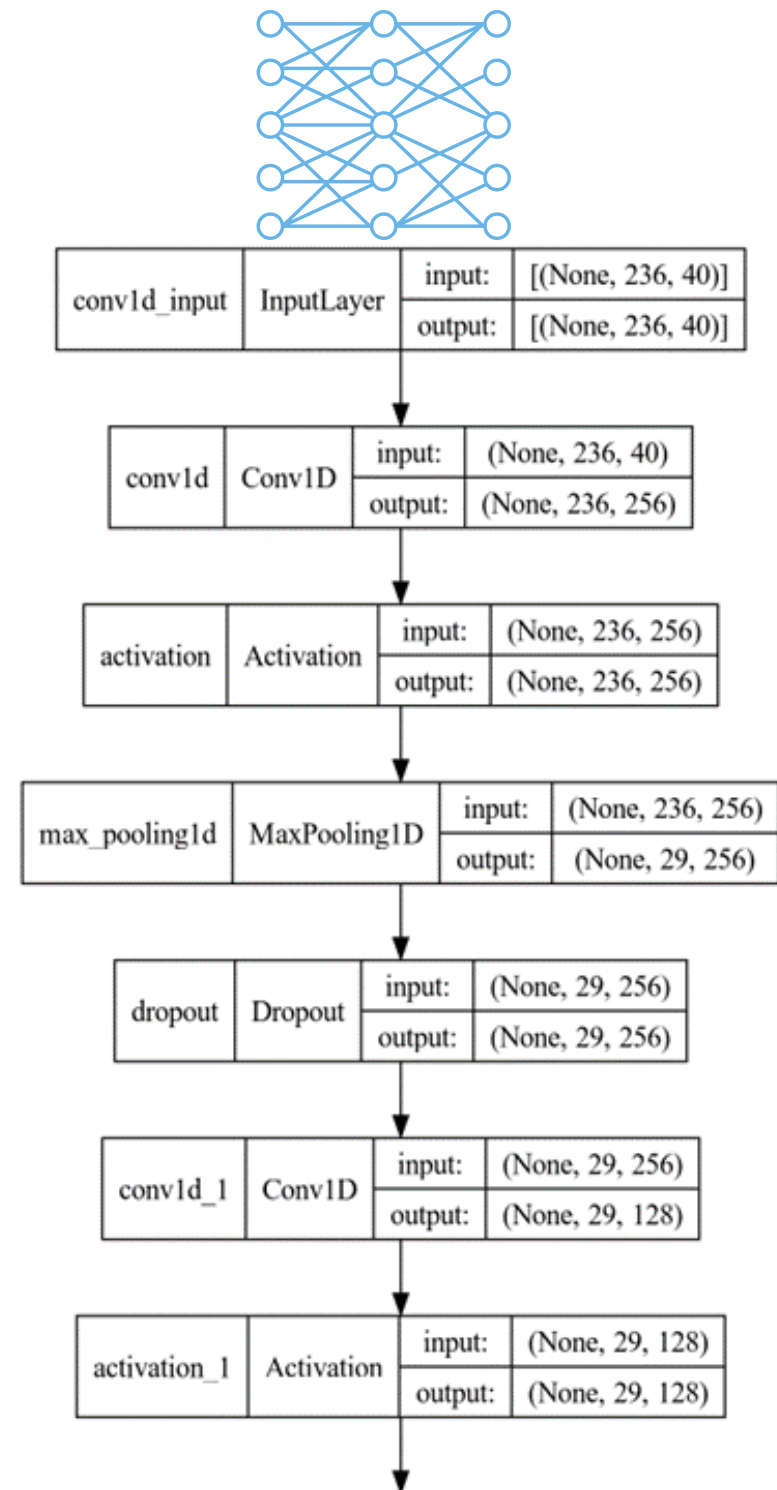
 DATOS COMBINADOS

EXTRACCIÓN MFCC



Entrenamiento

ENTRENAMIENTO DEL MODELO



Keras.callbacks

model_checkpoints

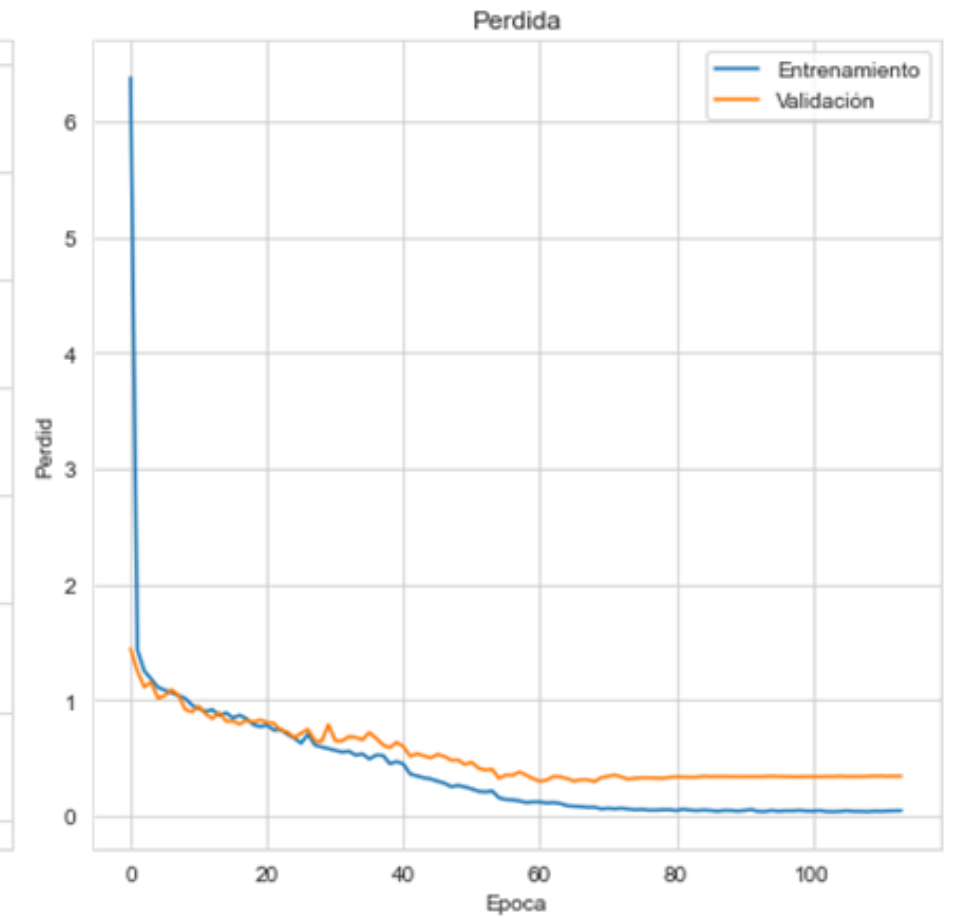
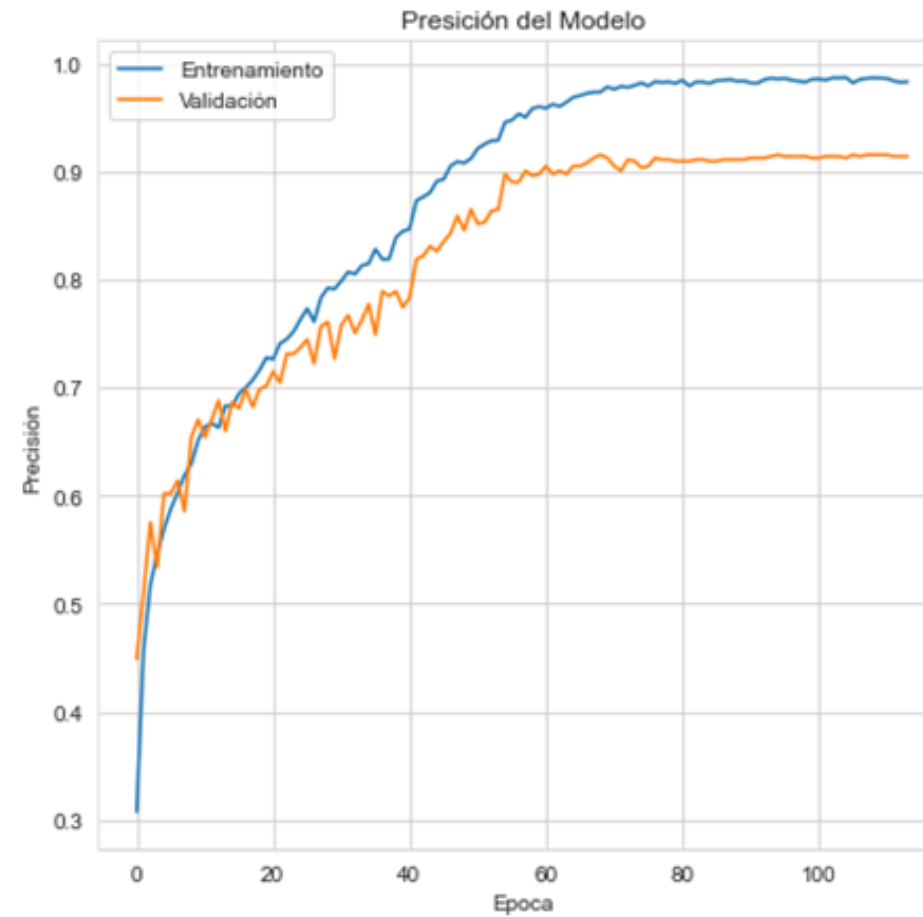
History = model.fit(X_train,Y_train,epochs = 500)

```

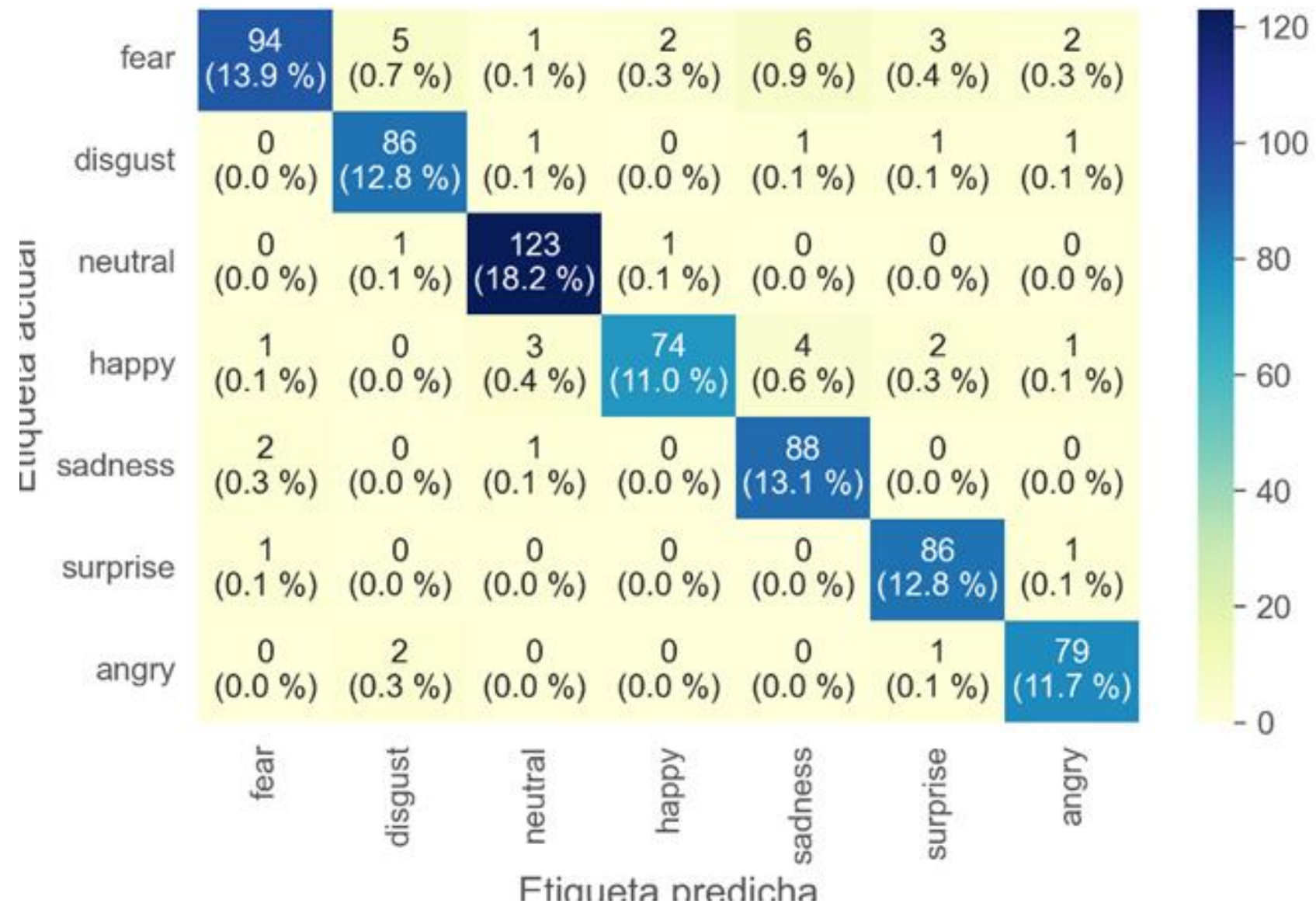
Epoch 1/500
2023-08-17 10:56:31.941644: I tensorflow/stream_executor/cuda/cuda_dnn.cc:366] Loaded cuDNN
337/337 [=====] - 2s 3ms/step - loss: 6.3769 - accuracy: 0.3073 - v
Epoch 2/500
337/337 [=====] - 1s 3ms/step - loss: 1.4351 - accuracy: 0.4575 - v
Epoch 3/500
337/337 [=====] - 1s 3ms/step - loss: 1.2533 - accuracy: 0.5176 - v
Epoch 4/500
337/337 [=====] - 1s 3ms/step - loss: 1.1864 - accuracy: 0.5457 - v
Epoch 5/500
337/337 [=====] - 1s 3ms/step - loss: 1.1119 - accuracy: 0.5689 - v
Epoch 6/500
337/337 [=====] - 1s 3ms/step - loss: 1.0886 - accuracy: 0.5876 - v
Epoch 7/500
337/337 [=====] - 1s 3ms/step - loss: 1.0658 - accuracy: 0.6027 - v
Epoch 8/500
337/337 [=====] - 1s 3ms/step - loss: 1.0410 - accuracy: 0.6188 - v
Epoch 9/500
337/337 [=====] - 1s 3ms/step - loss: 1.0180 - accuracy: 0.6292 - v
Epoch 10/500
337/337 [=====] - 1s 3ms/step - loss: 0.9582 - accuracy: 0.6509 - v
Epoch 11/500
337/337 [=====] - 1s 3ms/step - loss: 0.9274 - accuracy: 0.6636 - v
Epoch 12/500
337/337 [=====] - 1s 3ms/step - loss: 0.9049 - accuracy: 0.6667 - v
Epoch 13/500
337/337 [=====] - 1s 3ms/step - loss: 0.9197 - accuracy: 0.6628 - v
...
337/337 [=====] - 1s 3ms/step - loss: 0.0471 - accuracy: 0.9824 - v
Epoch 114/500
337/337 [=====] - 1s 3ms/step - loss: 0.0484 - accuracy: 0.9829 - v
Epoch 00114 early stopping

```

Output is truncated. View as a [scrollable element](#) or open in a [text editor](#). Adjust cell output [settings](#)...

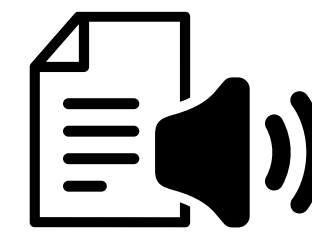


VALIDACIÓN DEL MODELO



Emoción	Precisión	Sensibilidad	Puntuación F1	Soporte
Miedo	0.96	0.83	0.89	113
Asco	0.91	0.96	0.93	90
Neutral	0.95	0.98	0.97	125
Feliz	0.96	0.87	0.91	85
Tristeza	0.89	0.97	0.93	91
Sorpresa	0.92	0.98	0.95	88
Enfado	0.94	0.96	0.95	82
Exactitud	0.93			674
Promedio ponderado	0.94	0.93	0.93	674

DEMOSTRACIÓN (EMOCIONES AUDIO)



Menú

Menú

Reconocimiento de Emociones

Modelo

¿Cómo le gustaría hacer la predicción?

mfccs

Configuraciones

- 3 emociones
- 6 emociones
- 7 emociones
- Transcripción Whisper
- género

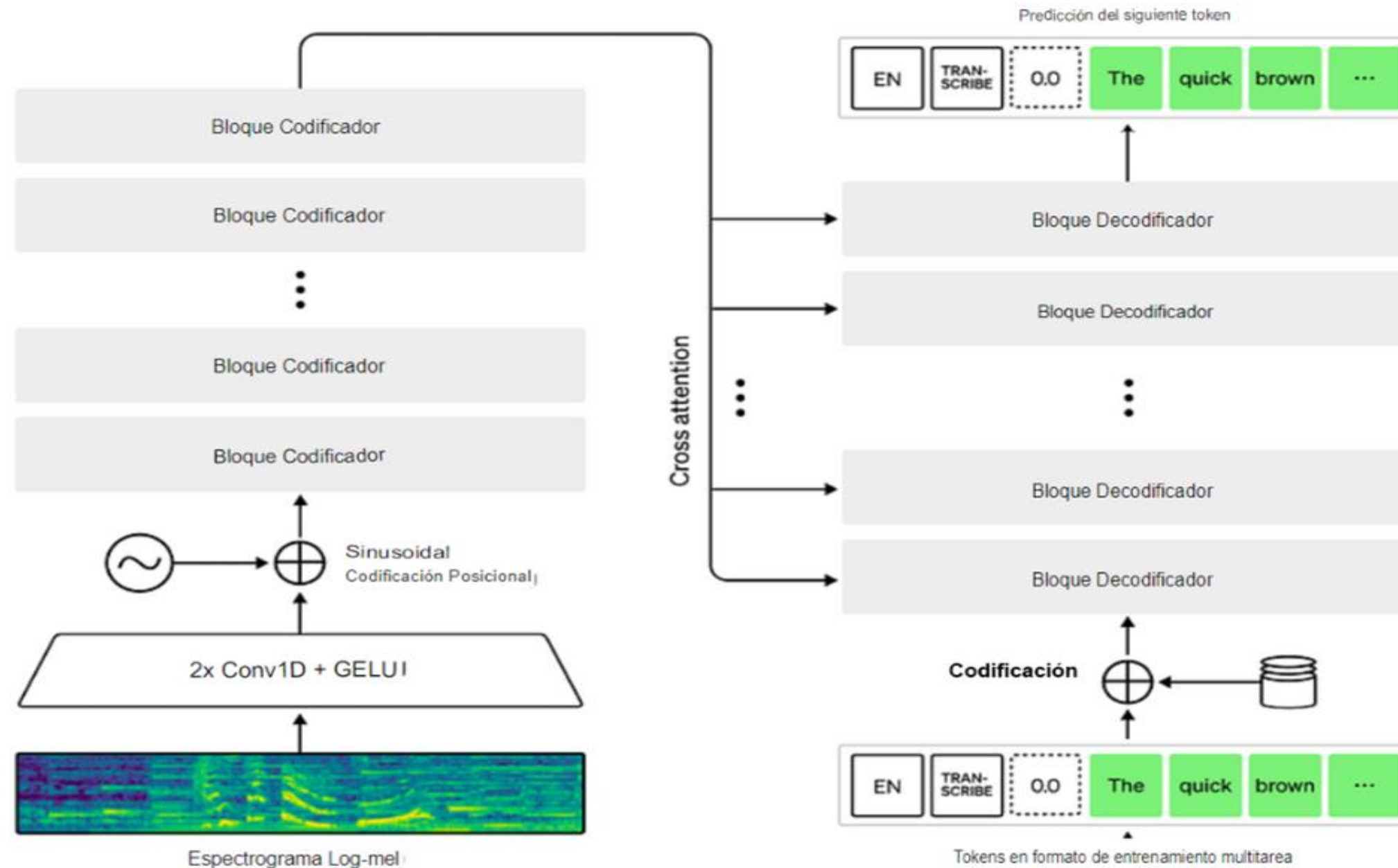
Linea Temporal de Emociones

Resumen de emociones

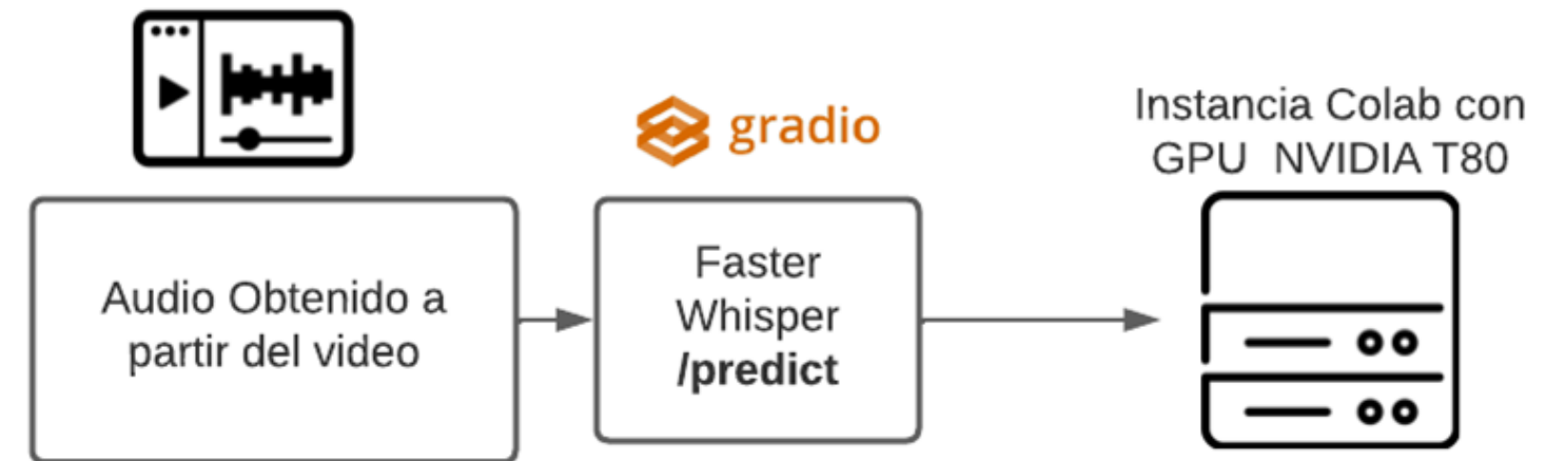
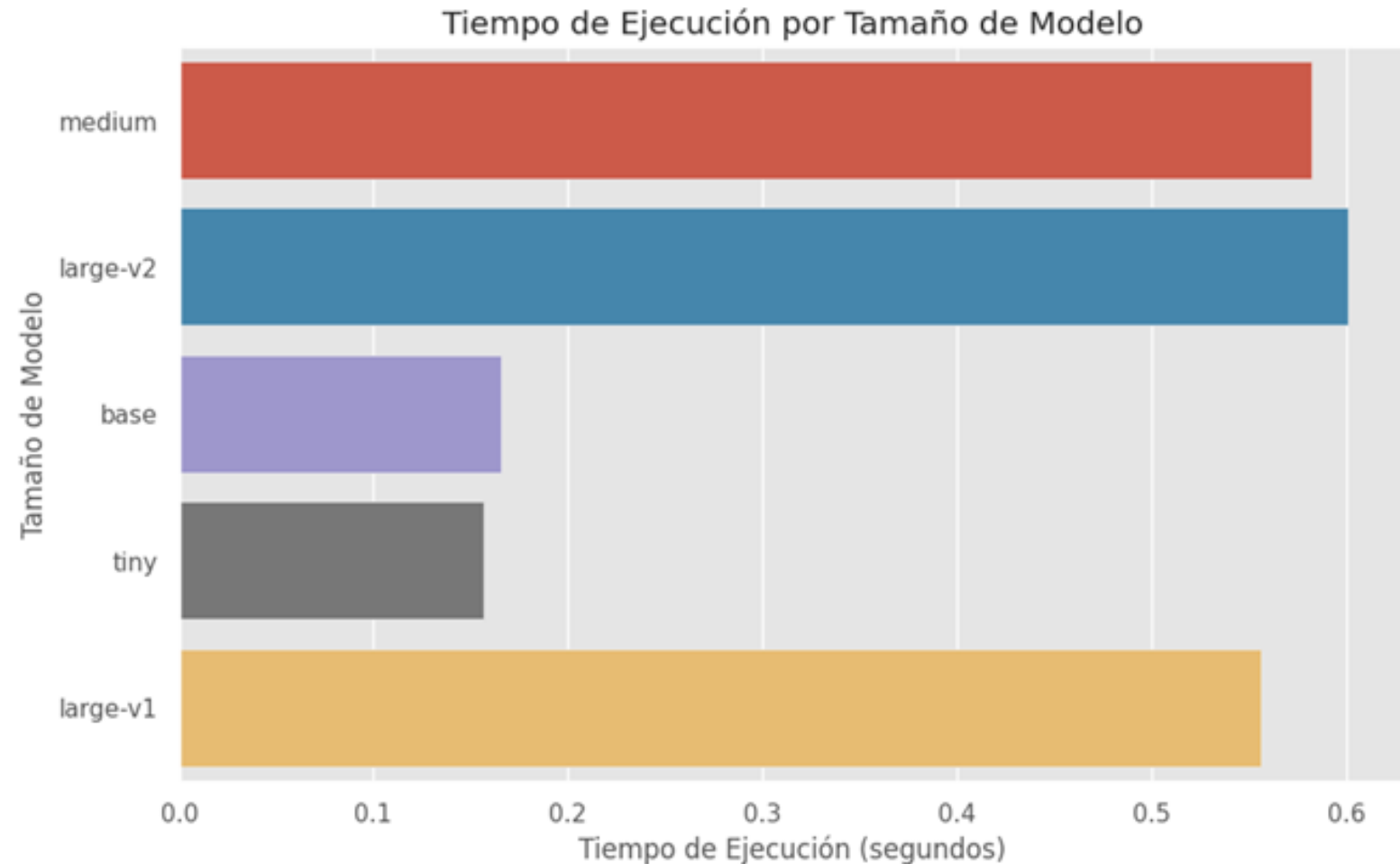
Emoción	Porcentaje
triste	50.8%
enojado	24.8%
asco	18.4%
neutral	1.0%
miedo	4.8%



Arquitectura Transformers:



IMPLEMENTANDO WHISPER



Implementación	Precisión	Tiempo	Uso de memoria máxima del GPU	Uso de memoria máxima del CPU
openai/whisper	fp16	90s	11325MB	9439MB
faster-whisper	fp16	54s	4755MB	3244MB
faster-whisper	int8	59s	3091MB	3117MB

```

simple_transcribe = gr.Interface(fn=ui.transcribe_webui_simple_progress
    if is_queue_mode else ui.transcribe_webui_simple,
    description=ui_description, article=ui_article, inputs=simple_inputs(),
    outputs=[
        gr.File(label="Descargar"),
        gr.Text(label="Transcripción"),
        gr.Text(label="Segmentos")
    ])

```

Detalle de las diferencias entre los modelos Whisper.

Modelo	Lenguaje	Tiempo (en segundos)	Frase	Memoria de GPU utilizada
tiny	Español	0.156766 segundos	Bueno, y yo me siento bien, creo que es una herramienta superfundamentada de V, que nos ayudó de un montón para de alguna forma facilitar la comunicación	16 mb
base	Español	0.165180 segundos	Bueno, yo me siento bien, creo que es una herramienta super fundamental y que nos ayuda un montón.	16 mb
medium	Español	0.581997 segundos	Bueno, yo me siento bien, creo que es una herramienta súper fundamental y que nos ayuda un montón	2912 mb
large-v1	Español	0.556571 segundos	Yo me siento bien, creo que es una herramienta super fundamental y que nos ayuda un montón	416 mb
large-v2	Español	0.600834 segundos	Bueno, yo me siento bien. Creo que es una herramienta súper fundamental y que nos ayudó un montón	1536 mb



hablando
puerta niños
LOS están
por la

Emoción obtenida a partir de la transcripción de texto:

00:00.000 --> 00:04.000

Los niños están hablando por la puerta.

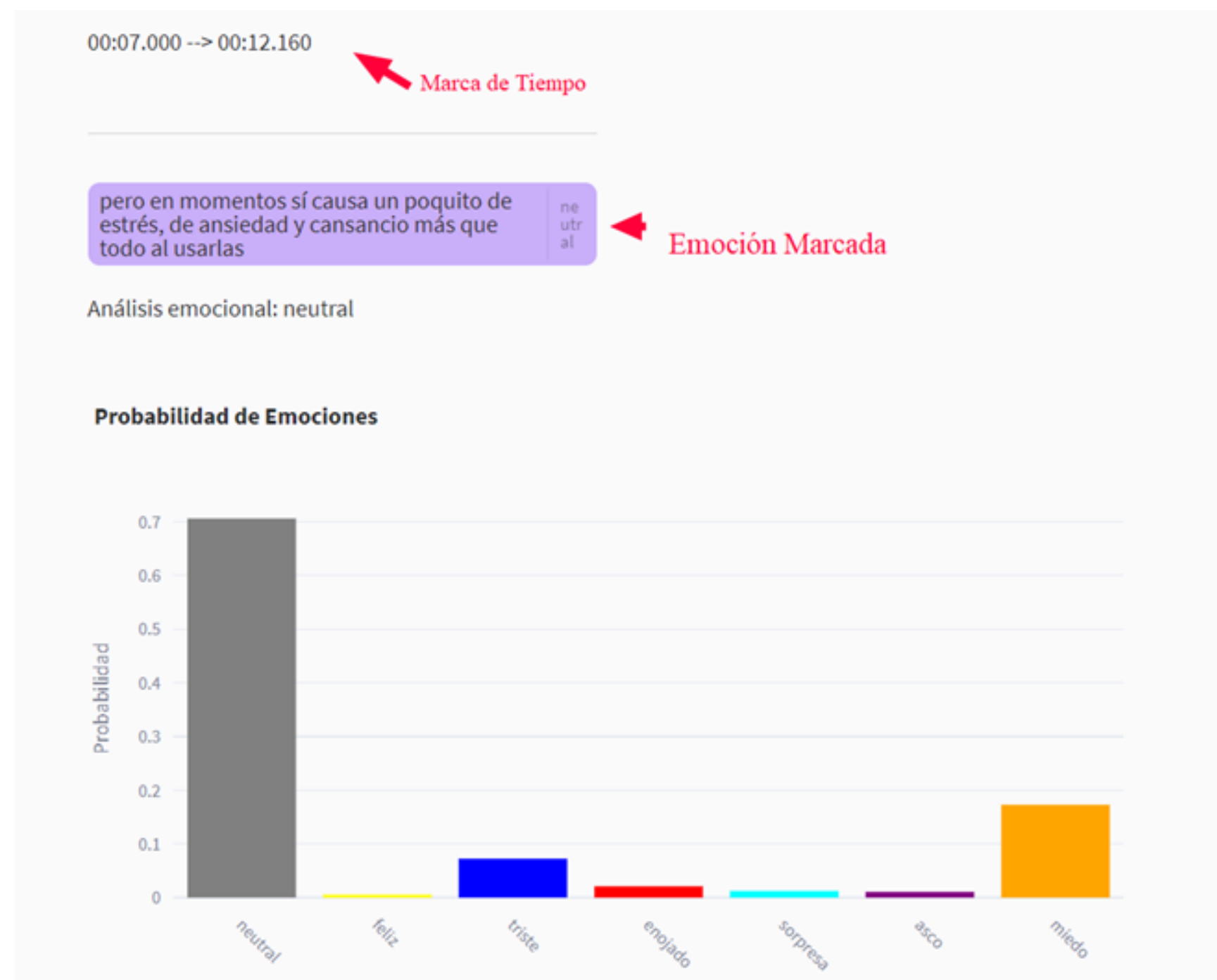
Análisis emocional: others

IMPLEMENTANDO EL MODELO NLP

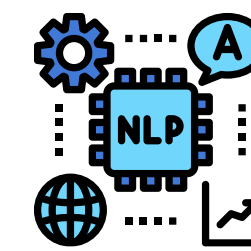
Puntajes F1 para los modelos encontrados en Pysentimiento

Modelo	Emoción
BERTin	50.2 +- 2.9
BETO	52.2 +- 1.4
Electricidad	46.3 +- 2.3
RoBERTa-es	53.1 +- 2.2
RoBERTuito	55.3 +- 0.8

Función: `analyze_emotion(self)`



DEMONSTRACIÓN (EMOCIONES TEXTO)



Menú

Menú

Reconocimiento de Emociones

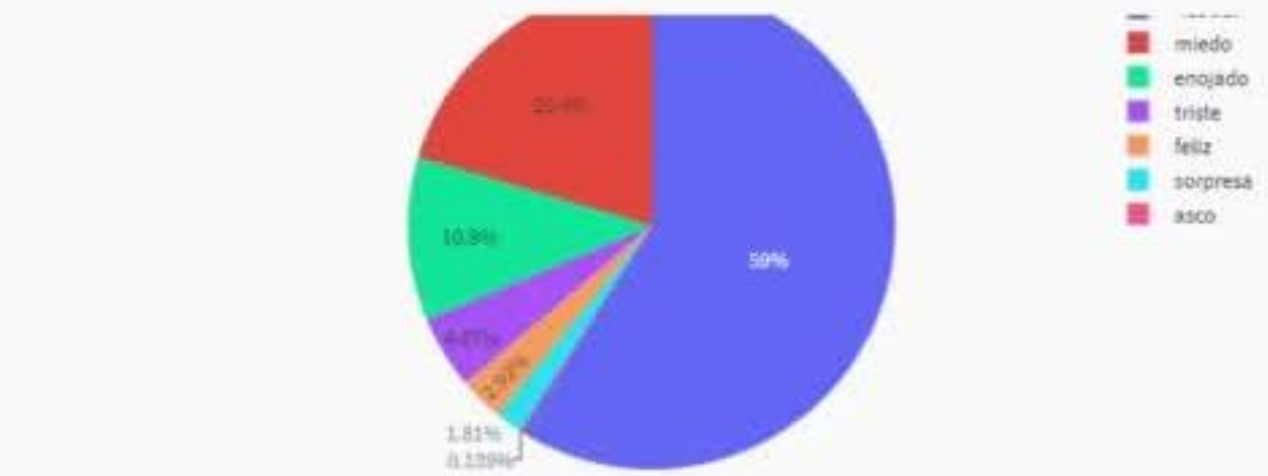
Modelo

¿Cómo le gustaría hacer la predicción?

mfccs

Configuraciones

- 3 emociones
- 6 emociones
- 7 emociones
- Transcripción Whisper
- género



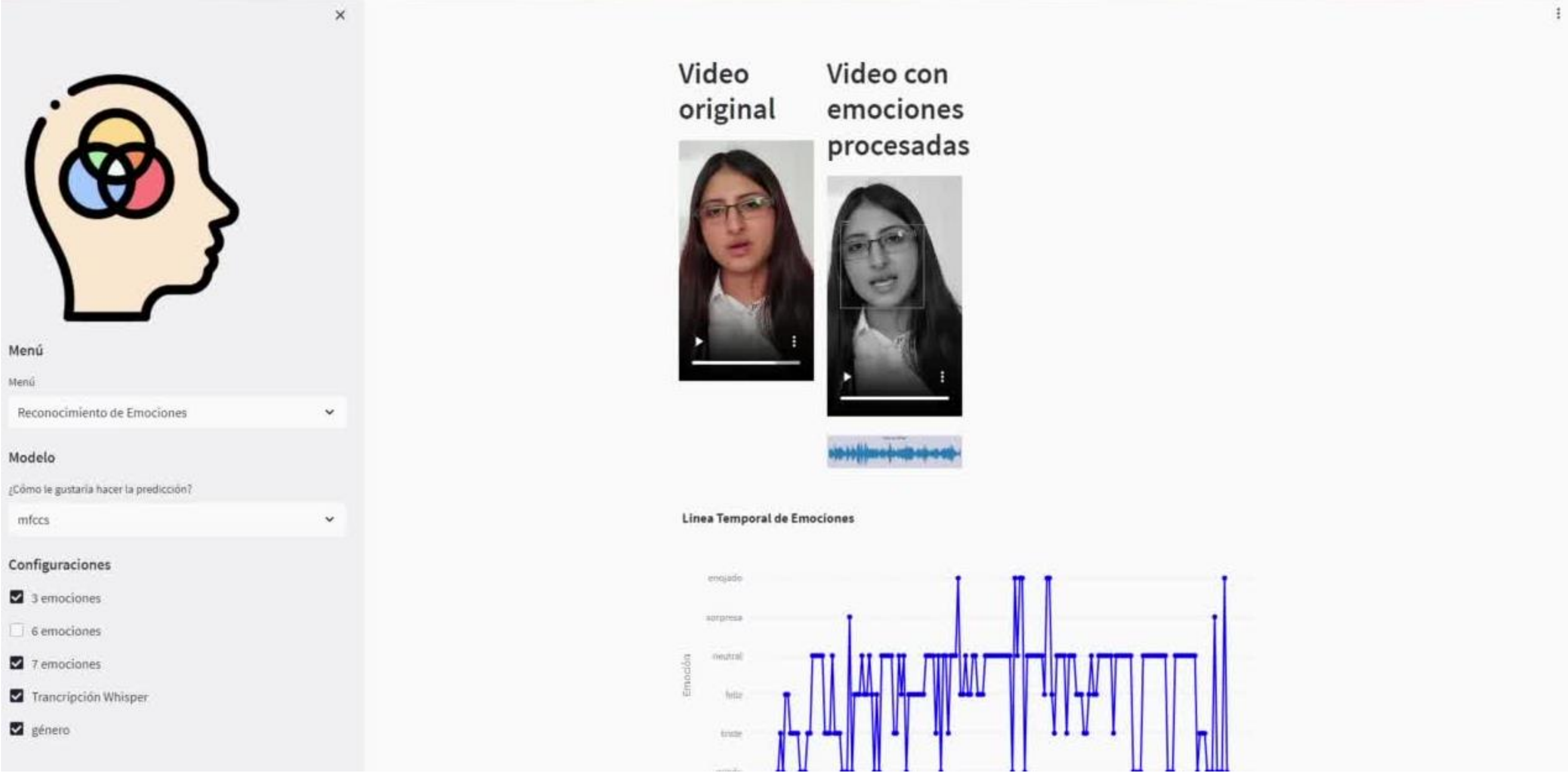
Analizando...



Predicciones



DEMONSTRACIÓN COMPLETA



The image shows a web application interface for emotion recognition. On the left is a settings sidebar, and on the right is the main analysis area.

Menú
Menú
Reconocimiento de Emociones

Modelo
¿Cómo le gustaría hacer la predicción?
mfccs

Configuraciones

- 3 emociones
- 6 emociones
- 7 emociones
- Transcripción Whisper
- género

Video original

Video con emociones procesadas

Línea Temporal de Emociones

Emoción

enojado
sorpresa
neutral
feliz
triste

The interface displays two video thumbnails side-by-side. The left one is labeled 'Video original' and the right one 'Video con emociones procesadas'. Below the right video is a small waveform visualization. The 'Línea Temporal de Emociones' is a line graph with a vertical axis labeled 'Emoción' and horizontal axis representing time. The y-axis categories are 'enojado', 'sorpresa', 'neutral', 'feliz', and 'triste'. The graph shows a blue line fluctuating between these levels, indicating detected emotion changes over time.

ESTADO DEL ARTE

05



ESTADO DEL ARTE



Paul Ekman ★★★★★

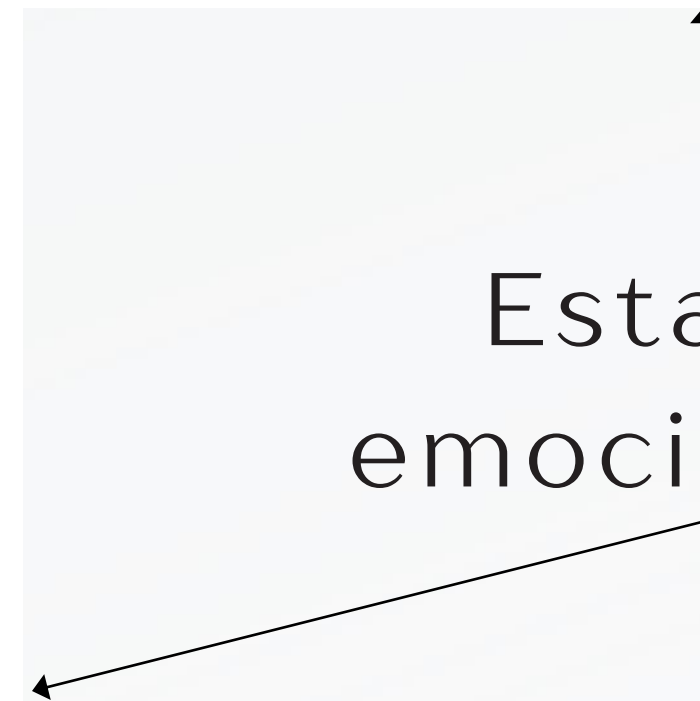
Paul Ekman es un psicólogo estadounidense conocido por su investigación pionera sobre las emociones y las expresiones faciales.



Codificación de emociones

Etiqueta	Traducción	Código
Angry	Enajo	1
Contemptuous	Desprecio	2
Disgusted	Asco	3
Fearful	Miedo	4
Happy	Feliz	5
Neutral	Neutral	6
Sad	Tristeza	7
Surprise	Sorpresa	8

Estados
emocionales



PLANIFICACIÓN DE LA REVISIÓN

Artículos Publicados	Título	Cita
AP1	Framework multimodal emocional en el contexto de ambientes dinámicos.	Ierache, J., Sattolo, I., & Chapperón, G. (2020). Framework multimodal emocional en el contexto de ambientes dinámicos. <i>Revista Ibérica de Sistemas e Tecnologías de Informação</i> , 40, 45–59.
AP2	Reconocimiento de emociones en la voz empleando redes neuronales y su integración en frameworks multimodales de educación emocional.	Ierache, J. S., & Elkfury, F. (2021). Reconocimiento de emociones en la voz empleando redes neuronales y su integración en frameworks multimodales de educación emocional. In <i>XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja)</i> .
AP3	Sistema integrado de reconocimiento de emociones para interacción hombre – máquina.	Martínez, F. H., Hernández, C. A., & Giral, D. A. (2023). Sistema integrado de reconocimiento de emociones para interacción hombre-máquina. <i>Información tecnológica</i> , 34(1), 117-128.
AP5	Use of the Student Engagement as a Strategy to Optimize Online Education, Applying a Supervised Machine Learning Model Using Facial Recognition	Noboa, A., Gonzalez, O., & Tapia, F. (2022). Use of the Student Engagement as a Strategy to Optimize Online Education, Applying a Supervised Machine Learning Model Using Facial Recognition. 283–295.

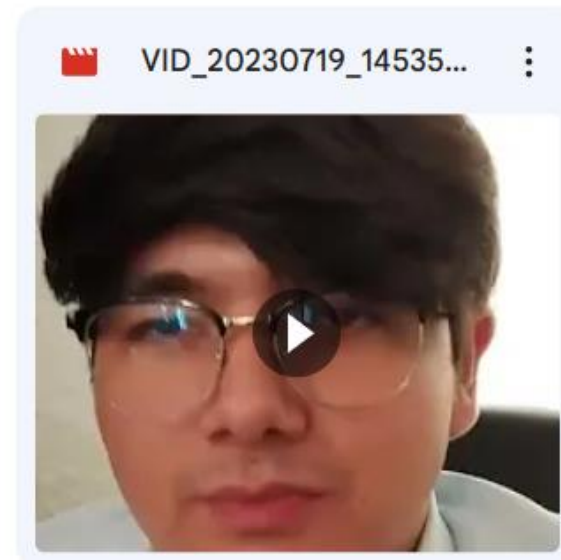
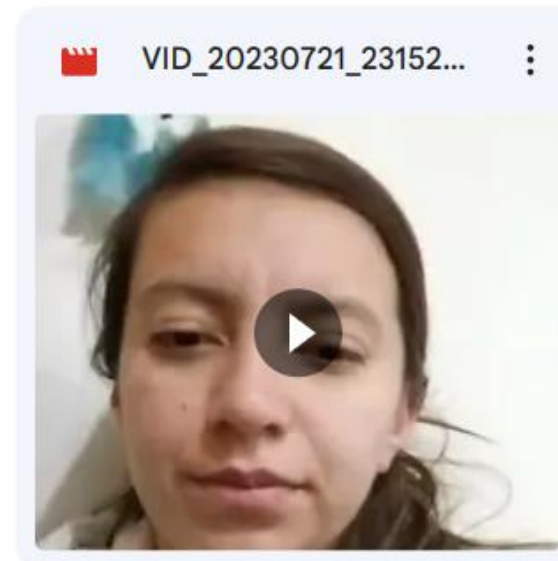
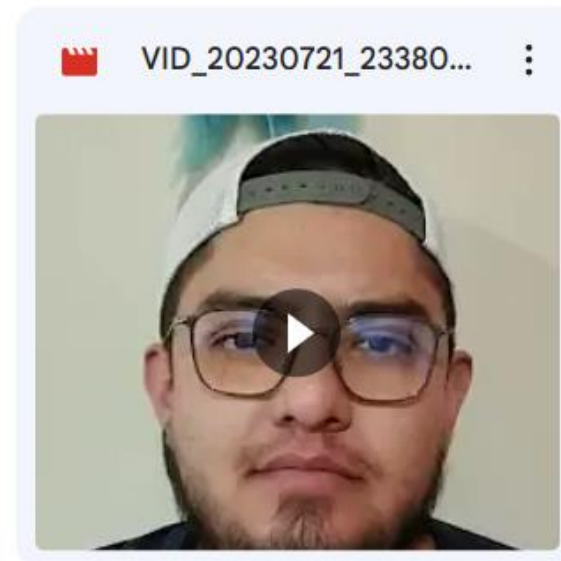
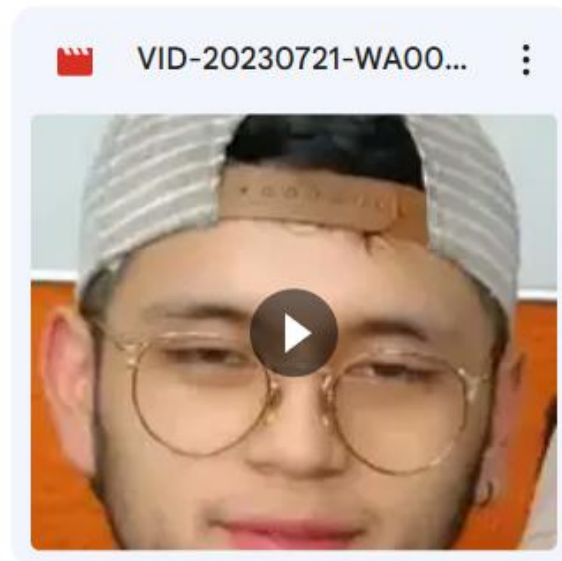


- Los investigadores en sus artículos dirigen sus estudios a diversos ambientes y consideran técnicas multimodales que permitan determinar resultados óptimos a través de reconocimiento facial, gestual o vocal.

RESULTADOS

06

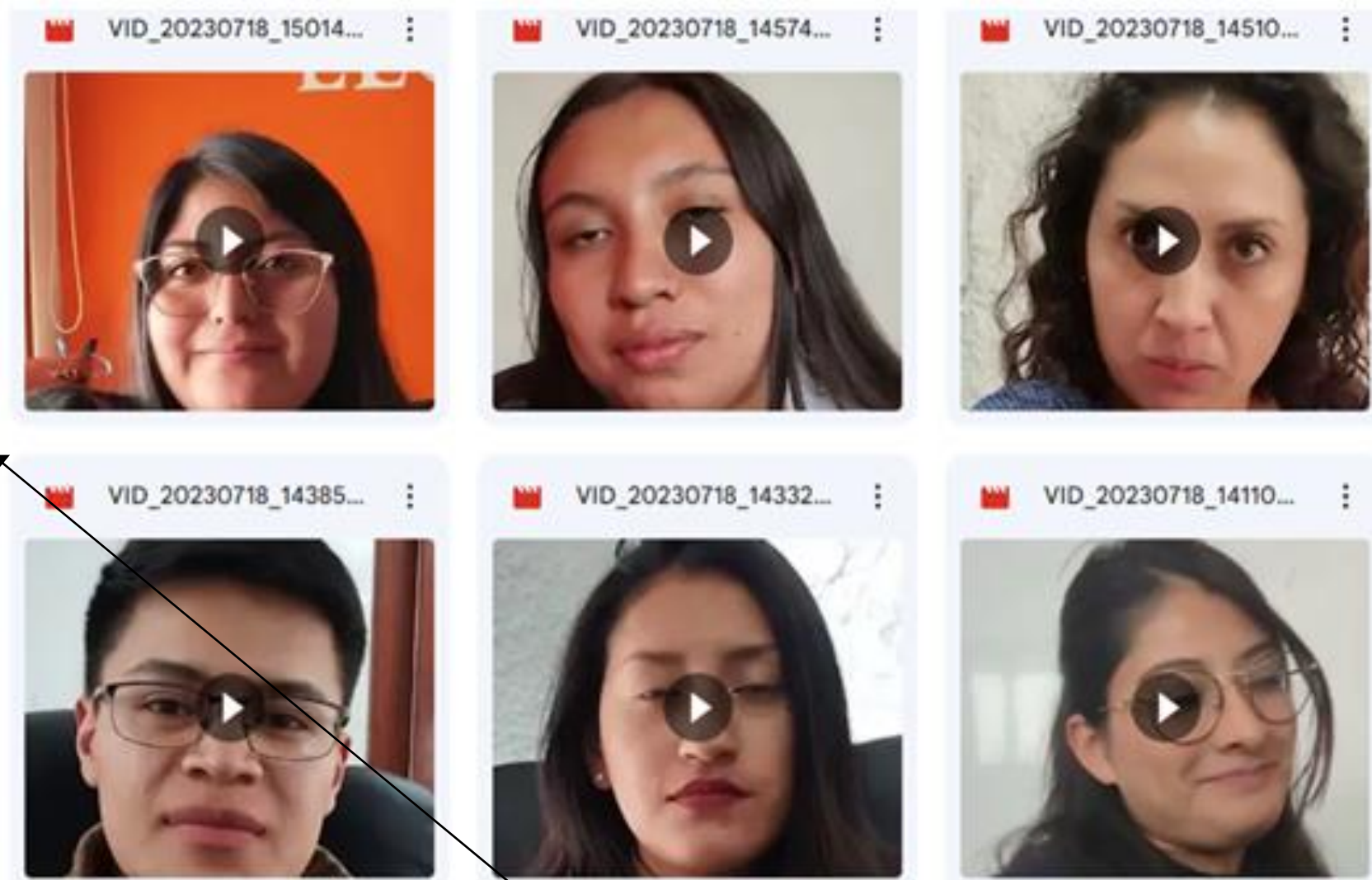
DESCRIPCIÓN DE LOS DATOS



El conjunto de datos manejados durante el proceso de reconocimiento emocional implica la realización de una entrevista en un entorno laboral no controlado, bajo la aplicación de un test que permite identificar sintomatología de tecnoestrés en los usuarios

Test de Evaluación de Tecnoestrés aplicado en la empresa GRUPO TAO Advisors

RECOLECCIÓN DE LOS DATOS



Dirigido a entornos no controlados del mundo real, capturando las expresiones faciales naturales y las emociones de las personas durante sus actividades cotidianas.

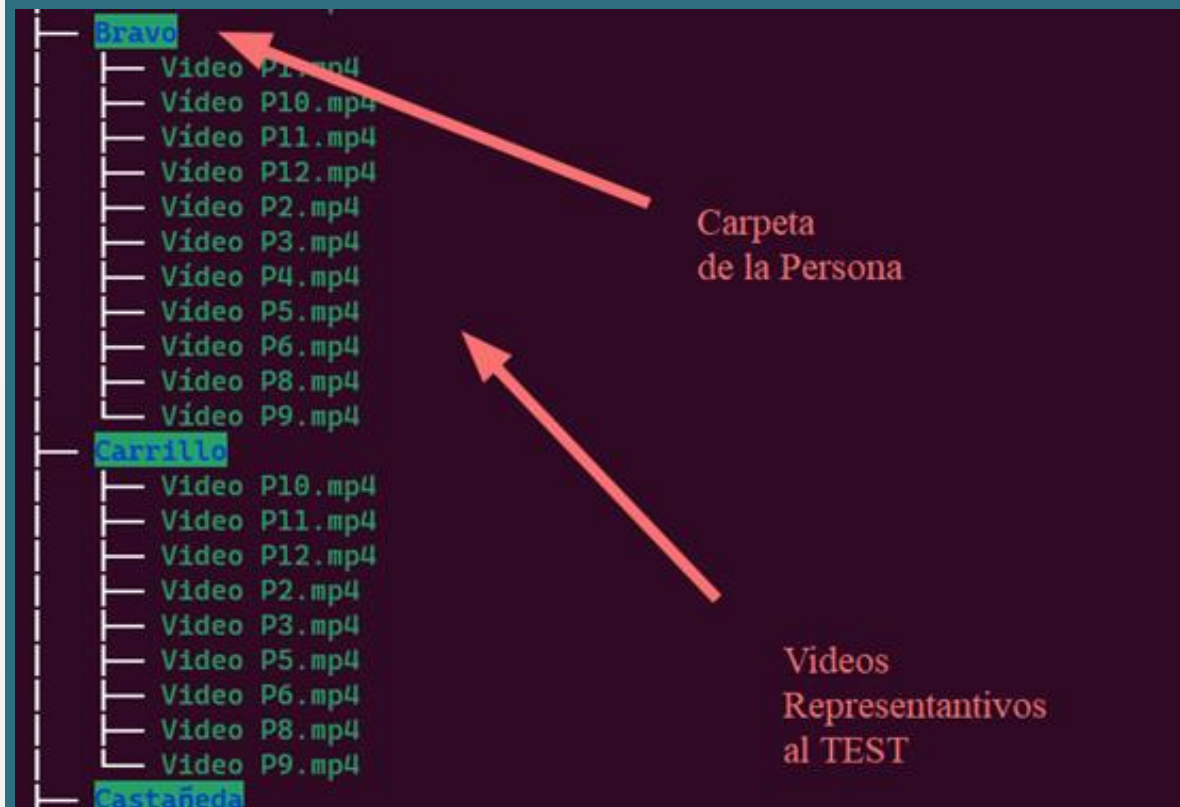
El framework permite cargar los videos recolectados, de esta manera se interpreta el estado emocional de la persona analizada.

Nombre Video	Genero	Encuestado/a	Área
VID_20230718_134816.mp4	Masculino	Jean Pierre Toapanta	Sistemas
VID_20230718_140611.mp4	Femenino	Ivone Ushiña	Contabilidad
VID_20230718_141109.mp4	Femenino	Karen Guerra	Contabilidad
VID_20230718_143327.mp4	Femenino	Melisa Peralta	Contabilidad
VID_20230718_143858.mp4	Masculino	Alex Carrillo	Sistemas
VID_20230718_145108.mp4	Femenino	Sara Chiriboga	Contabilidad
VID_20230718_145745.mp4	Femenino	Shyrley Castañeda	Contabilidad
VID_20230718_150147.mp4	Femenino	Karina Ochoa	Contabilidad
VID_20230718_150523.mp5	Femenino	Vanesa Gualoto	Contabilidad
VID_20230718_151118.mp6	Femenino	Joselyn Mantilla	Contabilidad
VID_20230718_151723.mp7	Femenino	Abigail Guerra	Contabilidad
VID_20230718_142409.mp8	Femenino	Karla Ciro	Contabilidad
VID_20230718_143632.mp9	Femenino	Fernanda Gualacata	Contabilidad
VID_20230718_144236.mp10	Femenino	Samanta García	Contabilidad
VID_20230718_144813.mp11	Femenino	Lizeth Maza	Contabilidad
VID_20230718_145352.mp12	Masculino	Nicolás Villavicencio	Sistemas
VID_20230718_145945.mp13	Masculino	Edwin Bravo	Sistemas
VID_20230718_231529.mp14	Femenino	Elizabeth Morejón	Contabilidad
VID_20230718_233802.mp15	Masculino	Carlos Quiroz	Sistemas
VID-20230721-WA0001.mp4	Masculino	Steeven Vargas	Sistemas

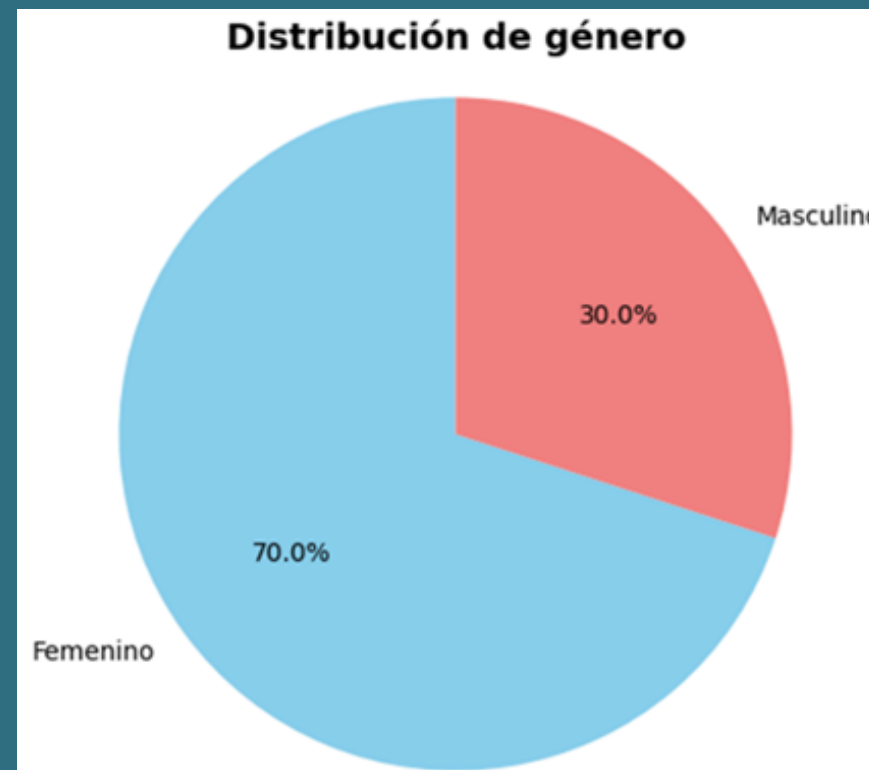
Análisis de los Datos

La muestra recolectada fue analizada previamente por el especialista en el campo psicológico, antes de ser procesada por el sistema se identificaron patrones de comportamiento relacionado con el tema propuesto

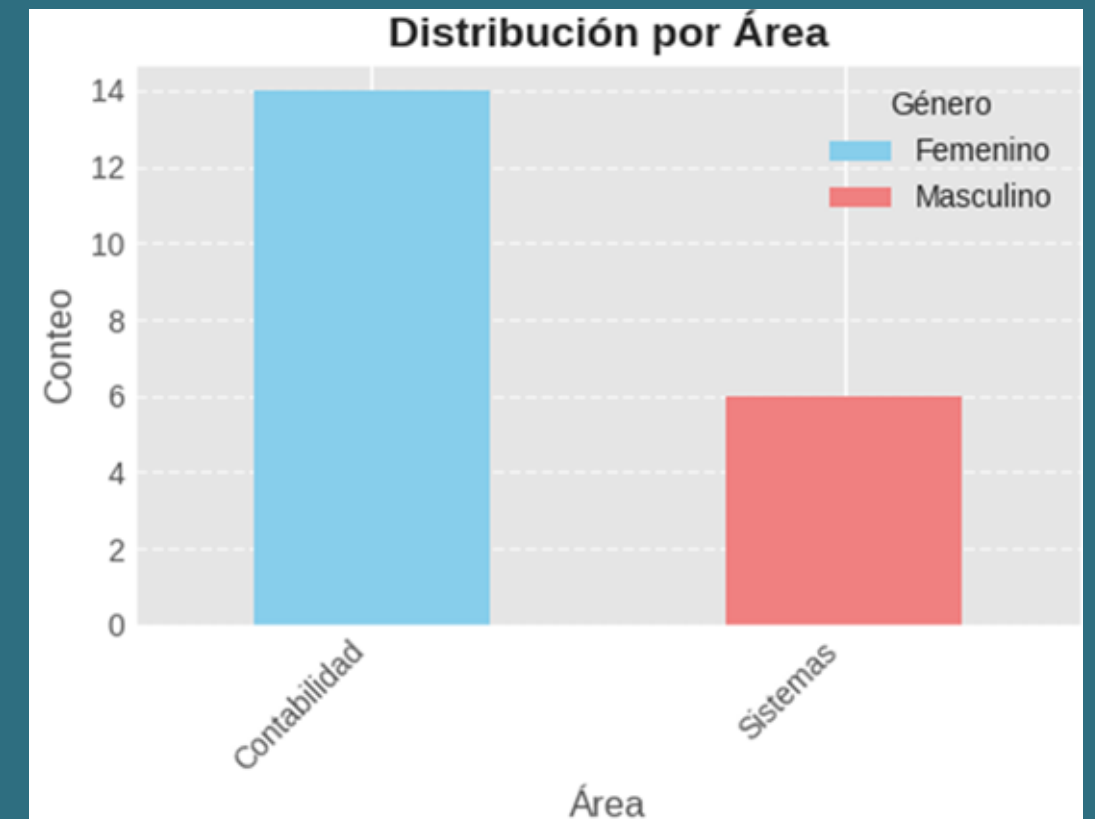
Extracción de la muestra .MP4



Detalle de los videos muestrales



Detalle de las personas participantes



RESULTADOS GENERALES

Resultados de Datos Generales

Sujeto de Prueba 1

- Género: Masculino
- Área Laboral: Sistemas

Figura 113

Video 1



Sujeto de Prueba 3

- Género: Femenino
- Área Laboral: Contabilidad

Figura 115

Video 3



Análisis Psicológico: Gesticulaciones neutrales. Reconoce las TICs como una herramienta que ayuda en varias áreas y tener el máximo provecho. Denota gran satisfacción con el uso y la presencia de las TICs. No ha experimentado estrés ansiedad o malestar al usar por tiempo prolongado las TICs.

Análisis Psicológico: Debido al ángulo y toma de la cámara es difícil reconocer adecuadamente las expresiones faciales. Mantiene buenas estrategias de afrontamiento, entiende que el aprendizaje puede ser frustrante en inicio.

RESULTADOS ESPECÍFICOS

Video original

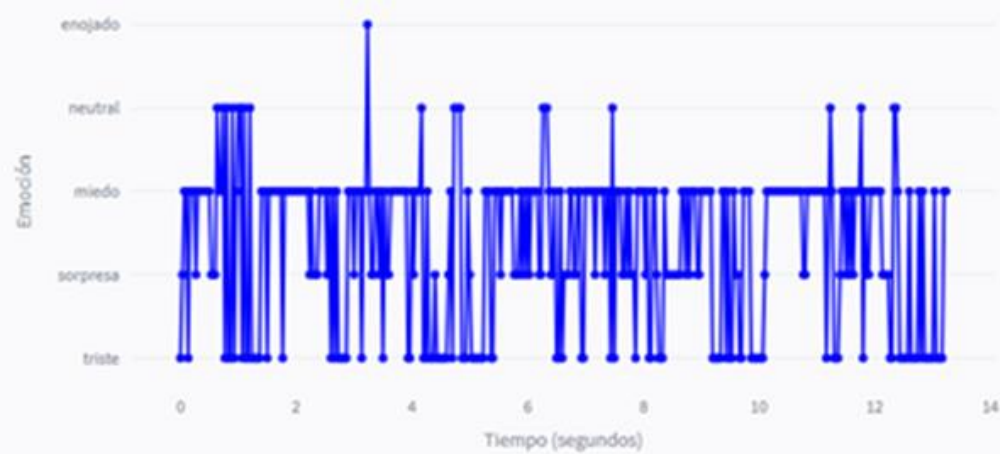


Video con emociones procesadas

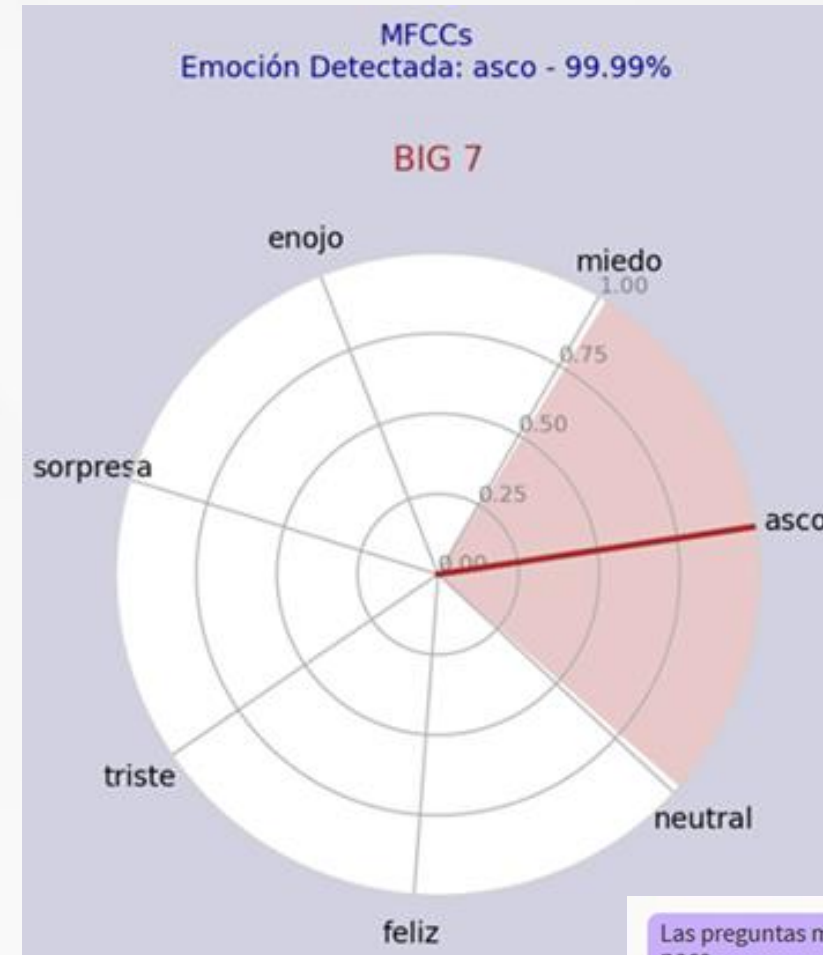
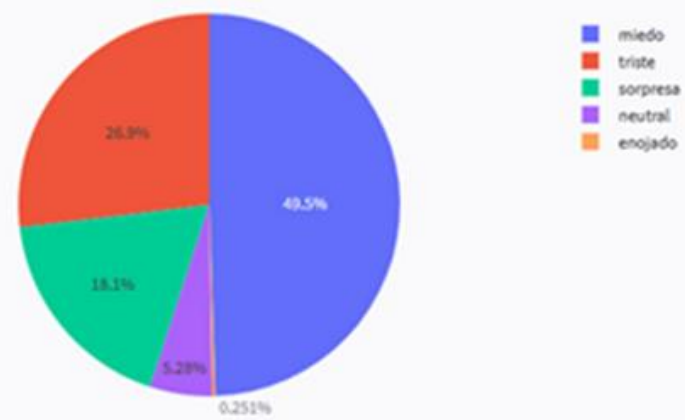


Se puede observar a través de una línea de tiempo los picos de las emociones que el sujeto a experimentando según el modelo predictivo facial

Línea Temporal de Emociones



Resumen de emociones

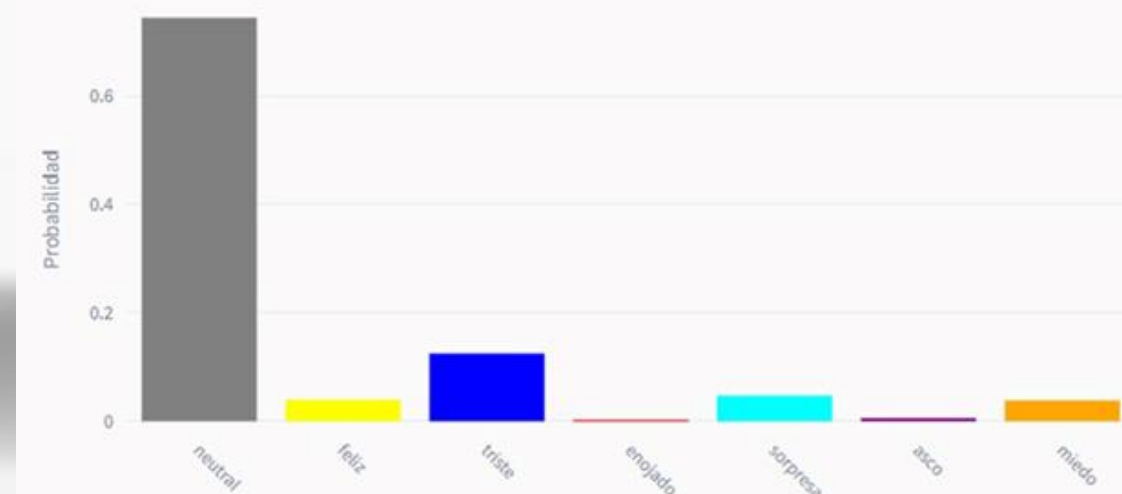


Por la voz su respuesta varía esto se debe a nerviosismo por ser una entrevista

Las preguntas me hicieron reflexionar un poco. neutral

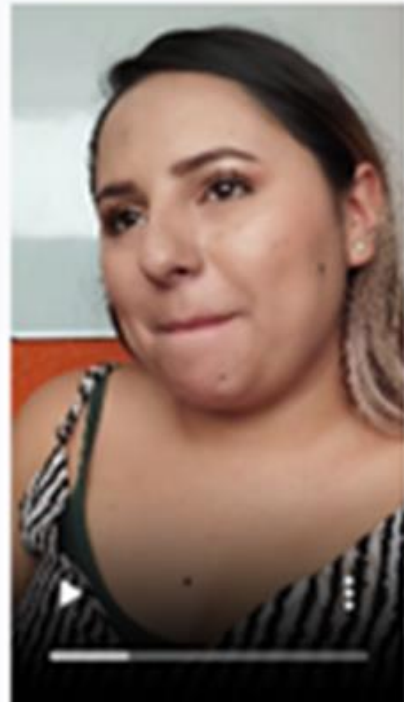
Análisis emocional: neutral

Probabilidad de Emociones

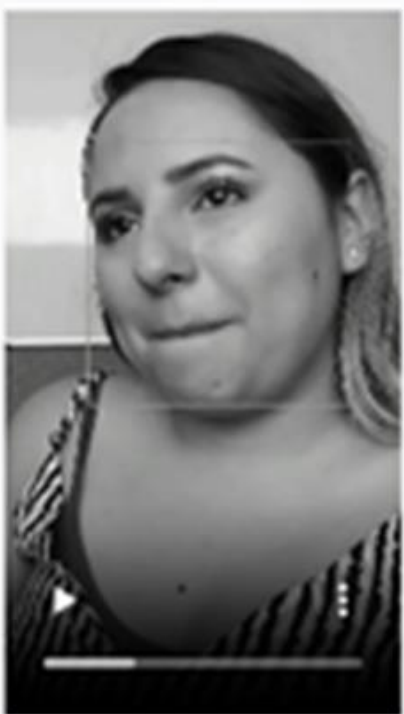


RESULTADOS ESPECÍFICOS

Video original

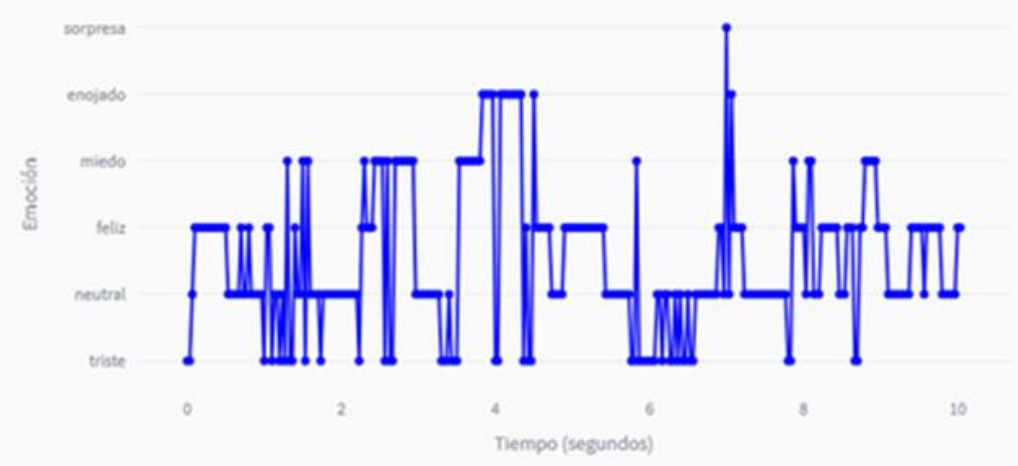


Video con emociones procesadas

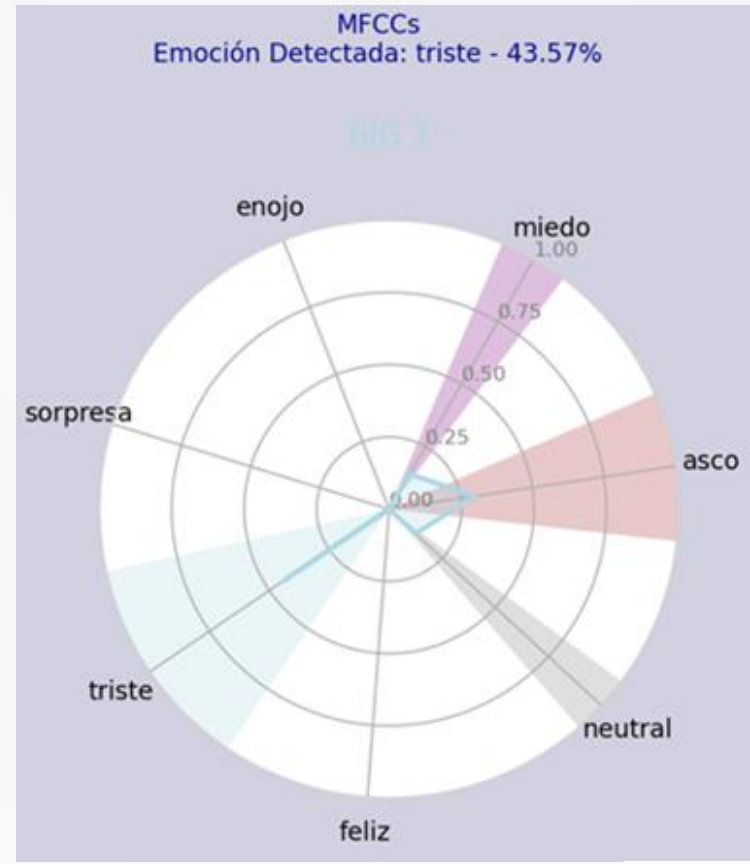
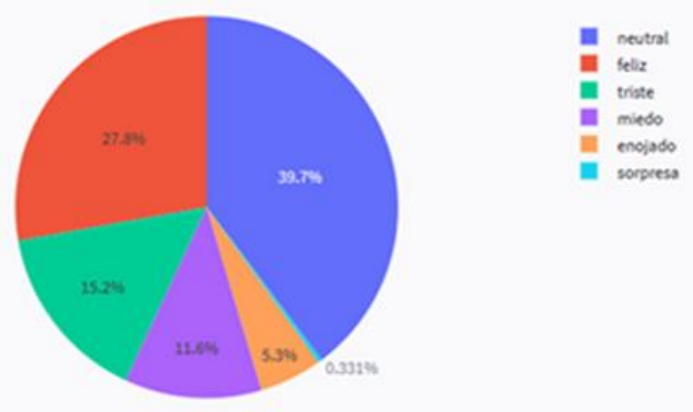


El sujeto muestra neutralidad, felicidad y tristeza con valores porcentuales más altos a diferencia del resto de emociones.

Línea Temporal de Emociones



Resumen de emociones

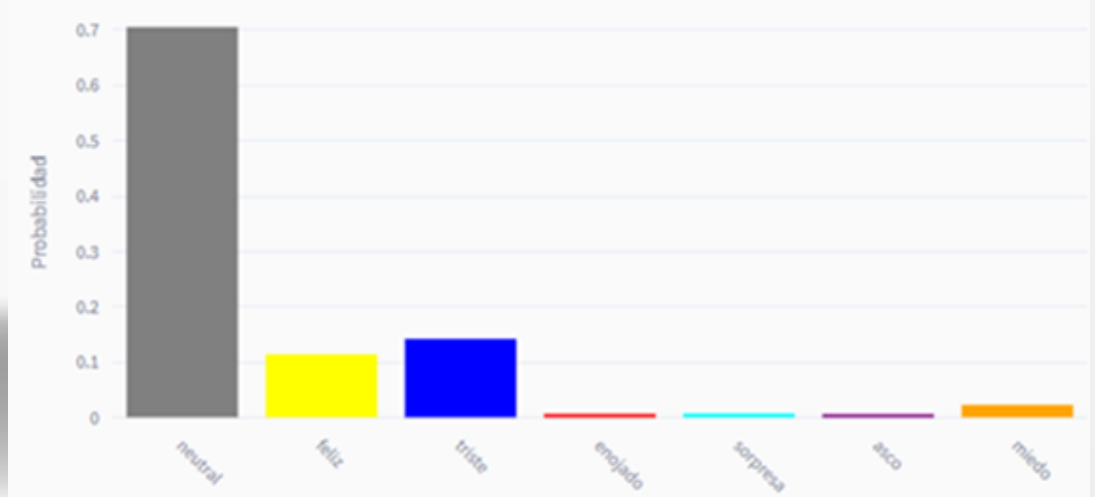


Se puede apreciar que el sujeto analizado expresa neutralidad, eso se debe a sintomatología de incertidumbre e inseguridad en su respuesta.

Actualmente bien, un poco estresada aprendiendo del sistema, los procesos, pero es un gobierno neutral

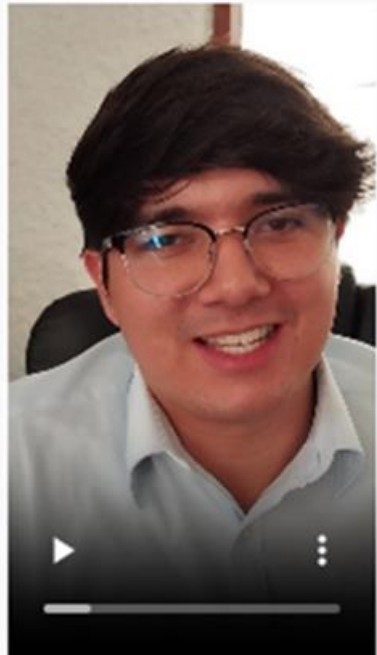
Análisis emocional: neutral

Probabilidad de Emociones

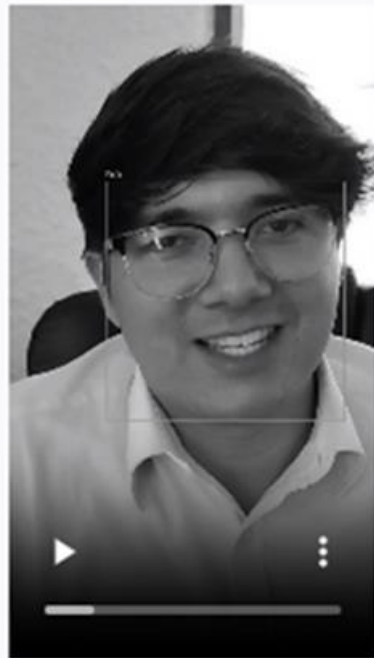


RESULTADOS ESPECÍFICOS

Video original

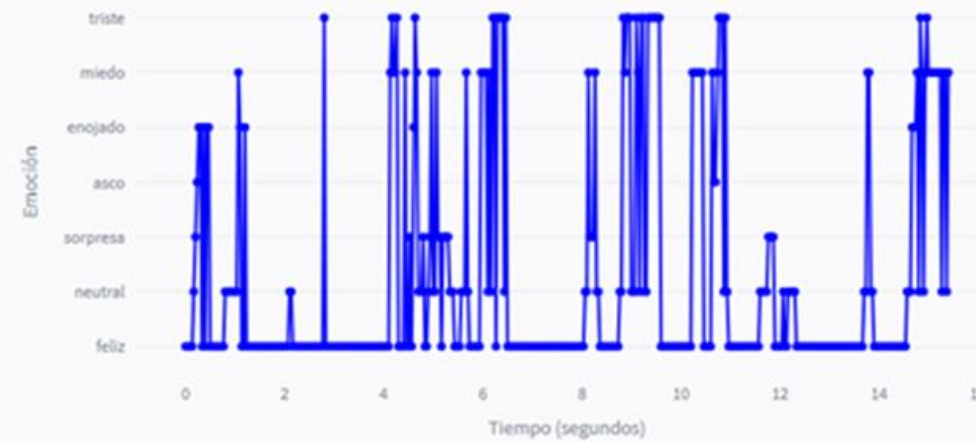


Video con emociones procesadas

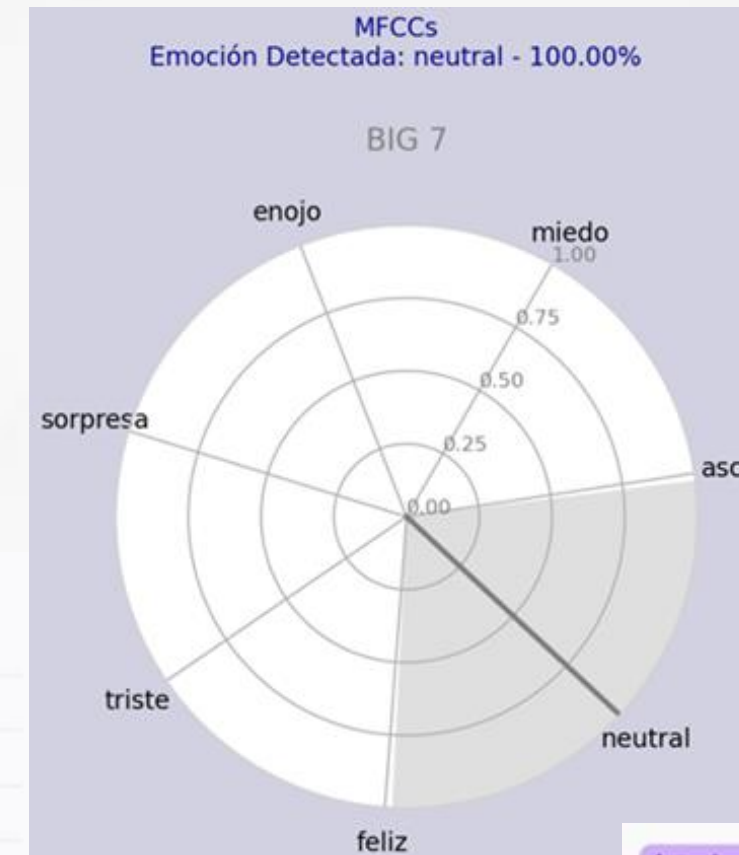
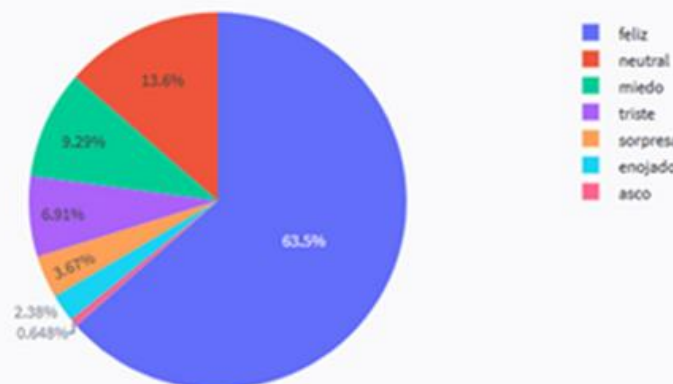


El sujeto muestra felicidad en mayor porcentaje seguido de neutralidad y miedo con valores porcentuales menores.

Linea Temporal de Emociones



Resumen de emociones

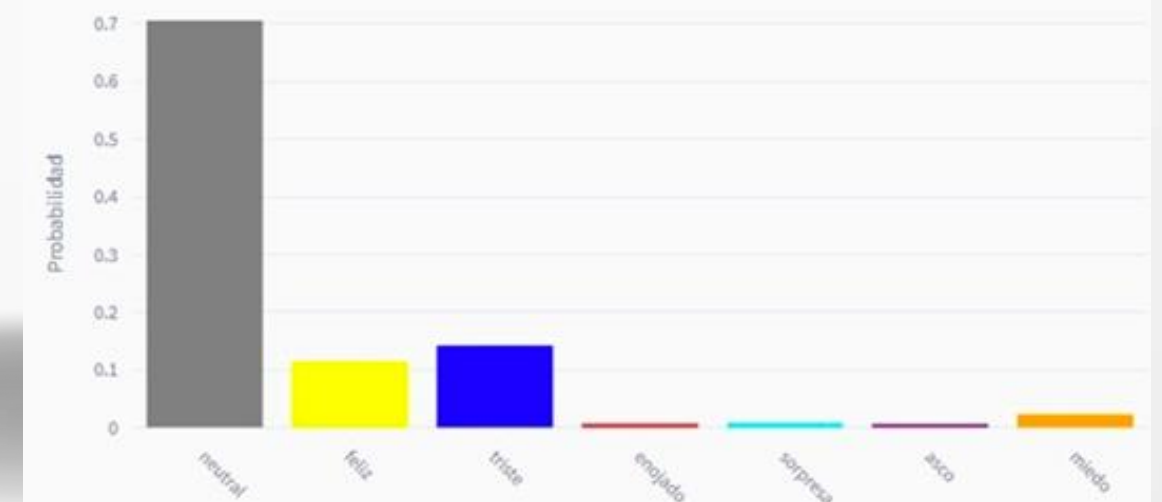


El sujeto analizado expresa neutralidad a través del texto, siendo su principal emoción durante el proceso de entrenamiento.

Actualmente bien, un poco estresada aprendiendo del sistema, los procesos, pero es un gobierno

Análisis emocional: neutral

Probabilidad de Emociones



CONCLUSIONES Y RECOMENDACIONES

07



CONCLUSIONES

- **Se logró la creación de un prototipo con la capacidad de reconocer y evaluar las emociones de un individuo en el contexto de la salud ocupacional. Esto se alcanzó mediante la combinación de un enfoque multimodal, que utiliza señales provenientes de expresiones faciales, gestos, comunicación no verbal y tono de voz.**
- Se realizó la revisión literaria sobre el reconocimiento multimodal de emociones, mediante esto se identificó estudios correspondientes a las arquitecturas y conjuntos de datos más utilizados dentro del campo del Deep Learning.
- Para el desarrollo y entrenamiento del sistema se realizó la implementación de un modelo de red neuronal convolucional de dos capas para el reconocimiento de emociones, la aplicación de técnicas de Hold-out en un conjunto de datos unificado, que abarca los corpus RAVDESS, TESS, MESD y SAVEE.
- La arquitectura propuesta de red neuronal convolucional de dos capas logró una precisión general del 93% aplicando técnicas de "Hold-out", el promedio ponderado sobre las siete clases de emociones en el conjunto de datos de pruebas, logrando específicamente: precisión del 97% para la emoción de miedo, 97% para disgusto, 93% para neutral, 96% para felicidad, 86% para tristeza, 89% para sorpresa y 92% para enojo.

CONCLUSIONES

- **Se logró determinar correctamente varias arquitecturas de reconocimiento de aprendizaje profundo y certeras en el reconocimiento de emociones: una red pre-entrenada BlazeFace para el reconocimiento de caras, una red neuronal convolucional de dos capas para el reconocimiento vocal y una red neuronal convolucional de tres capas pre-entrenada para el reconocimiento facial.**
- La aplicación de SCRUM como metodología ágil permitió que el proceso se desarrolle bajo comunicación de los desarrolladores y sus colaboradores al momento de recibir los videos, considerando a individuos enfocados en áreas administrativas y de sistemas quienes manejan y conocen sobre herramientas tecnológicas.
- El proyecto fue desarrollado en un lenguaje de programación Python, siendo uno de los lenguajes de categoría alta para desarrolladores expertos, de esta manera se supo aprovechar la amplia gama de librerías al momento de realizar la programación enfocada en reconocimiento emocional.
- Los niveles de tecnoestrés son evidentes en la muestra tomada, se puede identificar la presencia de sintomatología, generando un diagnóstico bajo niveles de estrés esperables acorde al trabajo que desarrollan diariamente.

RECOMENDACIONES

- **Al momento de realizar el entrenamiento de los modelos predictivos de reconocimiento emocional, se recomienda validar continuamente su rendimiento con métricas de evaluación estandarizadas y validadas, mediante esto podemos lograr un porcentaje de precisión más elevado.**
- Se recomienda que para procesos de reconocimiento emocional se empleen otro tipo de emociones basadas en la rueda de emociones de Plutchik de esta manera se podrá contemplar patrones más precisos entorno al comportamiento humano durante sus actividades diarias.
- Se recomienda que durante el proceso de extracción de datos el moderador no sólo enfoque la parte fisiológica de un individuo ya que para determinar el comportamiento y lenguaje no verbal requiere de mayor parte del cuerpo, esto se denomina inquietud psicomotora.
- Es recomendable que el sistema permita realizar el reconocimiento emocional bajo la combinación de audio y video con una prueba validada, de esta manera se busca garantizar la fiabilidad de los resultados arrojados en torno a criterios de tecnoestrés.

TRABAJOS FUTUROS

08



TRABAJOS FUTUROS

● Para futuras investigaciones se recomienda realizar este tipo de reconocimiento emocional enfocadas en un entorno laboral controlado, empleando buena iluminación y enfoque de manera frontal, de esta manera la captura de gesticulación sea favorable para el receptor de la muestra y para futuros diagnósticos en el área psicológica.

● Realizar el reconocimiento emocional bajo la combinación de audio y video con una prueba validada, de esta manera se busca garantizar la fiabilidad de los resultados arrojados en torno a criterios de tecnoestrés.

● Es recomendable que este tipo de proyectos sea dirigido a una población más extensa, ya que al estar dirigida en una muestra pequeña la generalidad de los datos influye en los resultados obtenidos bajando el nivel de confiabilidad de estos.