

Diseño de un agente tutor web basado en emociones: elección del mejor desencadenante según el perfil del estudiante

Elección del desencadenante óptimo basado en el reconocimiento de emociones faciales y el perfil del estudiante

¿Cuándo intervenir? Determinación de umbrales emocionales para activación de agentes tutores en educación en línea

Mariano León Nájera, ORCID: 0000-0002-3229-8329, Universidad Autónoma del Estado de México, México

Magally Martínez Reyes, ORCID: 0000-0002-2643-6748, Universidad Autónoma del Estado de México, México

Anabelem Soberanes Martín, ORCID: 0000-0002-1101-8279, Universidad Autónoma del Estado de México, México

Resumen

Palabras clave

OpenCV; rendimiento; reconocimiento facial; sistema experto; visión artificial.

Abstract

Keywords

OpenCV; performance; face recognition; expert system; computer vision; artificial vision; facial recognition.

Introducción

En los últimos años la investigación sobre el estudio de las emociones ha ido en aumento, cada vez son más los científicos que dedican sus esfuerzos para profundizar en este tema. Partiendo de la evidencia que da por sentado Charles Robert Darwin en el año 1872 propone que los animales como los seres humanos exteriorizan emociones discretas, de entre las cuales presenta las emociones de ira, miedo, asco, entre otras. Años más tarde otros científicos de entre los cuales destaca Paul Ekman utilizaron enfoques similares a los de Darwin para organizar las emociones. Muchos de estos científicos nuevos han centrado sus experimentos en examinar la fisiología de las emociones centrándose en la expresión facial (Ekman P. , 2016).

El autor Ekman (1981) propone una clasificación de emociones básicas, afirma que existen evidencia clara de universalidad en las expresiones de: alegría, ira, desagrado, tristeza y la mezcla entre el miedo/sorpresa. Esto lo encontró a través de diferentes métodos y técnicas, partiendo de la medición muscular del rostro, de la acción visible, pues el autor define que una expresión facial es la manifestación de indicadores musculares específicos y distintos para cada tipo de emoción básica. Formulada a partir de bases: en teoría, inductiva y anatómica, en su búsqueda por la exactitud de los datos obtenidos, observados y cuestiones sin formular,

busca en la retroalimentación facial encontrar el control neural y correlación autónoma de la expresión psicofisiológicos entre la cara y el cerebro; es decir, sin duda existe una correlación en la dirección del pensamiento (Vargas, 2016).



Expresiones de universalidad.

Fuente: Adaptación personal (Ekman, 1981)

Los cursos en línea tradicionales presentan el mismo contenido a todos los estudiantes, requiriendo que completen el temario estándar para finalizar su formación. Sin embargo, es ampliamente reconocido que cada alumno posee características únicas, derivadas de sus conocimientos previos, experiencias y estilos de aprendizaje. Estas diferencias individuales asociadas a la forma en que procesan información y emplean sus capacidades cognitivas han impulsado el desarrollo de métodos de enseñanza-aprendizaje flexibles y dinámicos. Tal evolución busca ofrecer interactividad adaptativa, ajustándose a las necesidades educativas, preferencias y perfiles específicos de cada estudiante (Castro, Martínez, Cruz, Juárez, & Soberanes, 2013).

Las plataformas de aprendizaje son entornos educativos de varios sistemas de información, elaborados con amplio rango de aplicaciones informáticas instaladas en un servidor cuya función es la de facilitar al profesorado la creación, administración, gestión y distribución de cursos a través de Internet para todo aquel alumno interesado en ellos (Rodríguez, 2005). Los elementos que componen una plataforma de aprendizaje tal y como lo refiere el sitio de internet Captera (2022), que presenta elementos novedosos indispensables para cumplir con su propósito son las siguientes herramientas: creación de cursos, gestión de cursos, pruebas y evaluaciones, gestión de alumnos, gestión de rutas de aprendizaje, aula virtual, aprendizaje social y análisis e informes. Así mismo refiere que en respuesta a las solicitudes de parte de los usuarios han mejorado la experiencia de aprendizaje al realizar la combinación de estos elementos, innovando con elementos novedosos para que los usuarios que interactúan en ellas puedan encontrar lo siguiente: entregar contenido de aprendizaje, seguimiento del progreso del alumno, fomentar la participación y el debate de los alumnos y agregar y analizar datos.

Las plataformas de aprendizaje modernas incorporan sistemáticamente estas características, denominadas colectivamente como perfil del estudiante. Este perfil constituye una base de conocimiento que integra supuestos explícitos sobre el usuario (Vélez, 2009), los cuales resultan determinantes para el comportamiento del sistema educativo. En el contexto de la educación virtual, el modelo del estudiante se estructura mediante dimensiones clave que

reflejan tanto los elementos fundamentales del proceso educativo como sus interrelaciones dinámicas (González, 2008).



Dimensiones del perfil del estudiante.

Fuente: Adaptación personal (González, 2008)

Cada dimensión contiene información específica sobre el estudiante dentro de la plataforma de aprendizaje, la cual responde a un propósito concreto en el marco del proceso educativo. En su mayoría, esta información es proporcionada directamente por el propio alumno, bajo un principio de honestidad, sinceridad y plena conciencia sobre el valor y la finalidad de sus respuestas. El estudiante es consciente de que forma parte de un proceso formativo trascendental y que falsear o manipular la información podría derivar en resultados que afectarían negativamente su propio desempeño académico (León-Nájera, 2021). De las once dimensiones que conforman el perfil del estudiante, en esta investigación se abordarán únicamente dos, seleccionadas por su correlación directa: la dimensión del estado anímico y la dimensión emocional.

La integración de actividades lúdicas vinculadas a los estilos de aprendizaje dentro de las plataformas virtuales debe tener un impacto significativo en las estrategias de enseñanza. Esto es especialmente relevante considerando que los estudiantes a distancia interactúan de forma asincrónica e intensiva con los materiales didácticos disponibles en los sistemas de gestión del aprendizaje (LMS). Conocer el perfil del estudiante a través de los datos que se encuentran en la plataforma de aprendizaje permite diversificar las actividades y los procedimientos didácticos, facilitando que los alumnos adquieran un mayor conocimiento de sí mismos. Este autoconocimiento se considera un indicador clave en el desarrollo de la conducta vocacional (Amaya & Cuéllar, 2016).

El reconocimiento de emociones en tiempo real puede implementarse mediante dos enfoques tecnológicos principales. Uno primero que combina Open CV con Python, mientras que el

segundo utiliza Tensor Flow con JavaScript. Es importante destacar que Open CV y Tensor Flow son herramientas conceptualmente distintas, diseñadas para propósitos diferentes lo que hace que su comparación directa no sea del todo equivalente.

Mientras que Open CV, es una popular biblioteca de visión artificial, sus principales algoritmos son Eigenfaces, Fisherfaces y Local Binary Patterns Histograms, es decir, se centra en el procesamiento de imágenes en tiempo real e incluye implementaciones sin patente de los algoritmos (OpenCV, 2025). Tensor Flow es simplemente una biblioteca para trabajar con modelos de aprendizaje automático para computadoras de escritorio, dispositivos móviles, la web y la nube. En pocas palabras, proporciona capacidades sólidas para implementar sus modelos en cualquier entorno, lo cual ayuda a revelar y eliminar sesgos en los datos para obtener resultados justos y éticos de los modelos (tensorflow, 2025).

La propuesta de realizar reconocimiento facial mediante la identificación de emociones en tiempo real desde un computador, presentada por Mariano León (2021), emplea las herramientas de OpenCV y Python, y constituye una alternativa prometedora. Las pruebas se llevaron a cabo utilizando el método *EigenFaceRecognizer*, el cual depende directamente de los recursos del dispositivo donde se ejecutan los análisis. Además, el entorno físico también influye significativamente en los resultados: factores como la calidad de la cámara, la iluminación y los ángulos de captura de imagen o video son determinantes en la precisión del reconocimiento facial.

Es cierto que, al desarrollar una investigación mediante este método, es necesario calibrar y ajustar diversos aspectos para alcanzar el objetivo propuesto. Existen factores que, aunque intervienen de manera indirecta, pueden afectar significativamente los resultados. Entre ellos se encuentran la iluminación que depende del entorno donde se realice el análisis y la calidad del video o imagen, la cual está condicionada por la cámara utilizada. Un ejemplo concreto fue el intento de analizar el interior de un autobús de pasajeros con el fin de desarrollar una aplicación capaz de identificar los asientos vacíos y reportarlos a los usuarios en tiempo real. Sin embargo, esta implementación enfrentó varios obstáculos que impidieron su funcionamiento óptimo, entre ellos: la baja resolución de las imágenes, condiciones inadecuadas de iluminación, ángulos de captura poco favorables y objetos que obstruían la visibilidad hacia el fondo del autobús (Romero-Torres, Martínez Reyes, & León-Nájera, 2024). El software era el apropiado, sin embargo, no todos los elementos que estaban involucrados en el proceso eran los adecuados.

Al utilizar la herramienta TensorFlow para el análisis de expresiones faciales, se identifican siete emociones, de las cuales cinco corresponden a las expresiones universales propuestas por Ekman. Además, se hace una distinción entre miedo y sorpresa, que Ekman agrupa como una sola categoría y se incluye la expresión de neutralidad como una categoría adicional. En el estudio *Reconocimiento de emociones faciales en un conjunto de datos mediante redes neuronales convolucionales* (Tümen, 2017), se presenta una propuesta para analizar un gran número de rostros mediante este tipo de redes. Los resultados obtenidos incluyeron el número de imágenes por emoción, la proporción de datos y la tasa de éxito, resumidos de la siguiente manera:

- **Ira:** 453 imágenes, 9.1% de proporción, 48.1% de éxito.

- **Disgusto:** 28 imágenes, 5.1% de proporción, 82.1% de éxito.
- **Miedo:** 423 imágenes, 8.2% de proporción, 47.5% de éxito.
- **Felicidad:** 951 imágenes, 10.5% de proporción, 71.1% de éxito.
- **Tristeza:** 701 imágenes, 11.5% de proporción, 46.2% de éxito.
- **Sorpresa:** 367 imágenes, 9.1% de proporción, 80.7% de éxito.
- **Neutral:** 666 imágenes, 10.7% de proporción, 46.7% de éxito.

En la etapa final del análisis, la red neuronal convolucional (CNN) alcanzó una precisión del 58.5% en los datos de verificación y del 57.1% en los datos de prueba. Considerando que la tasa de éxito promedio en humanos para este conjunto de datos es de aproximadamente 65% \pm 5%, los resultados obtenidos por la red son aceptables, especialmente dado que no se aplicó ningún preprocesamiento a las imágenes (Coronel Caján , 2021).

Por tal razón la propuesta que se implementará será bajo la herramienta de Tensor Flow ya que gracias a su versatilidad de ejecutar modelos de aprendizaje automático se ajusta al desarrollo y comprobación de las emociones a la hora del estudio en plataformas de aprendizaje, esto es gracias a su tecnología que le permite realizar este tipo de análisis en un sitio web sin tener grandes recursos ni software especializado. Además, promete su aplicación en dispositivos móviles, no como una app sino, a través del uso del explorador web.

Por esta razón, la propuesta que se implementará utilizará la herramienta TensorFlow, ya que su versatilidad para ejecutar modelos de aprendizaje automático la hace adecuada para el desarrollo y la validación del reconocimiento de emociones durante el estudio en plataformas de aprendizaje. Esta elección se basa en su capacidad tecnológica para realizar análisis directamente en un sitio web, sin requerir recursos computacionales avanzados ni software especializado. Además, permite su aplicación en dispositivos móviles a través del navegador web, sin necesidad de desarrollar una aplicación nativa.

Objetivo e hipótesis

Establecer umbrales emocionales precisos que permitan definir el momento oportuno para la intervención de agentes tutores automatizados en entornos de educación en línea, a partir del análisis del perfil del estudiante en su estado emocional.

Al establecer umbrales de las emociones precisas mediante el análisis del perfil del estudiante, entonces será posible determinar de manera efectiva el momento oportuno para la intervención de agentes tutores automatizados en entornos de educación en línea.

Desarrollar un modelo que integre el reconocimiento de emociones faciales en tiempo real y el análisis del perfil del estudiante, con el propósito de determinar umbrales y establecer el desencadenante óptimo para la intervención de un agente tutor automatizado en entornos de educación en línea.

La integración de un modelo que combine el reconocimiento de emociones faciales en tiempo real y el análisis del perfil del estudiante permitirá identificar umbrales adecuados, lo que

favorecerá intervenciones oportunas por parte de agentes tutores automatizados en la educación en línea.

Metodología

Herramientas utilizadas (modelo facial, dataset, tipo de estudiantes, etc.).

Para el desarrollo de la aplicación encargada de seleccionar el desencadenante adecuado que active al agente tutor, se empleará *Tensor Flow* como herramienta principal. Esta elección se debe a su capacidad para identificar las expresiones universales propuestas por Ekman: ira, disgusto, miedo, felicidad, tristeza y sorpresa, además de la expresión neutral, la cual no había sido mencionada previamente. Cabe destacar que, a diferencia de la clasificación de Ekman quien agrupa las expresiones de miedo y sorpresa como una sola, *Tensor Flow* las distingue por separado (Coronel Caján , 2021).

Como complemento a la herramienta mencionada anteriormente, se empleará *JavaScript (JS)*, un lenguaje de programación ligero ampliamente utilizado en el desarrollo de sitios web interactivos. Su principal ventaja radica en su capacidad para añadir funcionalidades dinámicas en tiempo real, gracias a su modelo de ejecución (*just-in-time*), que combina interpretación y compilación (MDN, 2025). JavaScript es una elección adecuada debido a sus características técnicas: es un lenguaje basado en prototipos, multiparadigma (con soporte para programación orientada a objetos, imperativa y declarativa), monohilo y dinámico. Estas propiedades permiten a los desarrolladores manipular tanto el contenido como la apariencia de una página web sin necesidad de recargarla, optimizando así la experiencia del usuario.

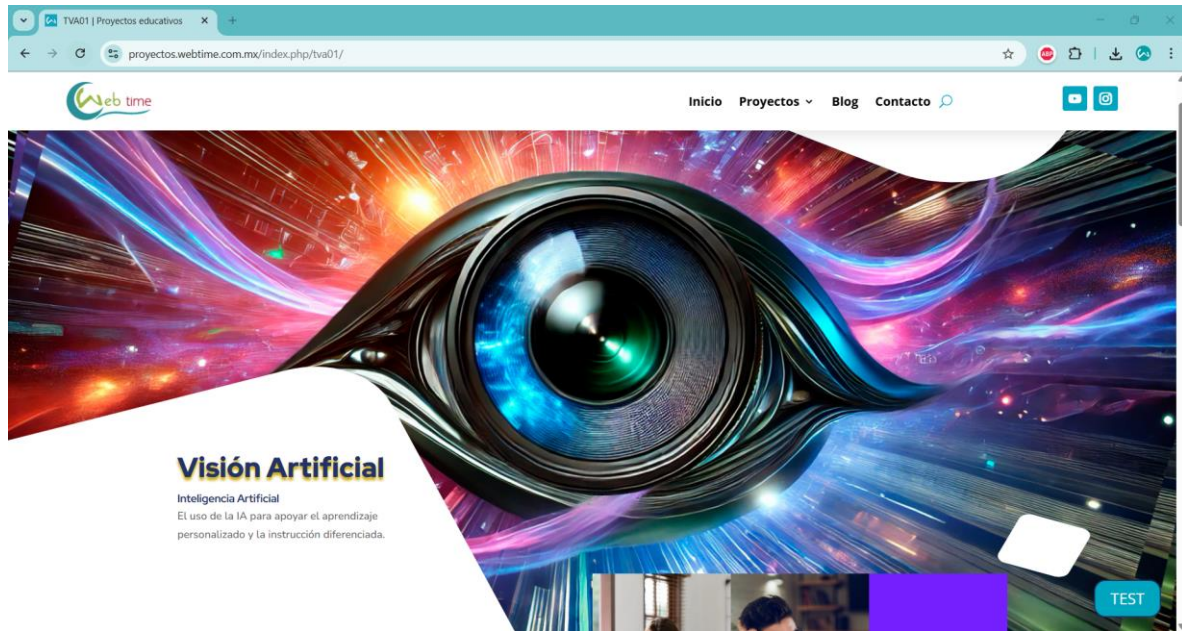
Para consolidar la implementación de *Tensor Flow* y *JavaScript* se desarrolló un simulador de actividades interactivo en un sitio web basado en WordPress con el tema Divi, utilizando PHP 7.4 como backend y MySQL para la gestión de bases de datos. Para el desarrollo se implementaron scripts personalizados en JavaScript, los cuales se acoplaron a TensorFlow para indicar las instrucciones que se realizarían para conseguir el objetivo deseado combinando así la flexibilidad del CMS con tecnologías avanzadas para crear una plataforma dinámica y escalable.

Para integrar TensorFlow y JavaScript en un entorno funcional, se desarrolló un simulador de actividades interactivo alojado en un sitio web basado en WordPress. Como infraestructura clave se utilizó:

- WordPress con el Tema Divi con scripts personalizados en JavaScript.
- PHP 7.4 para la lógica en el servidor.
- MySQL para la gestión y almacenamiento de datos.

Los scripts personalizados en JS garantizaron la integración perfecta con TensorFlow, traduciendo las salidas en instrucciones ejecutables para alcanzar los objetivos para sistema. Esta combinación aprovecha la flexibilidad de WordPress como CMS y la potencia de tecnologías avanzadas, resultando en una plataforma dinámica, escalable y de alto

rendimiento, es decir, facilitando futuras actualizaciones o la incorporación de nuevos modelos con IA.



Las actividades que se contemplaron fueron en lo general las que suelen implementar en las plataformas de aprendizaje. Es cierto como lo refiere la autora Reyes (2018) en la que refiere que los objetos de aprendizaje que deben de implementarse deben ser sometidos a una metodología, esto con el propósito que cumplan dentro de la actividad. Es por ello que se aclara que no se utilizaron objetos de aprendizaje estructurados a ese nivel, ya que la intención no es la construcción de alguno de ellos, sino, simplemente rescatar las aplicaciones más comunes que se pueden encontrar dentro de estas plataformas. Por tal razón se seleccionaron las acciones de lectura, reproducción de un video, un juego educativo, y la reproducción de audio.

Cómo se diseñó la evaluación de las emociones.

El diseño pedagógico busca inducir en el estudiante un estado cognitivo que active su necesidad de solicitar ayuda. Para lograr este objetivo, se implementaron cuatro estrategias basadas en la emoción compuesta miedo/sorpresa de Ekman:

1. **Estimulación textual:** Un texto multilingüe seleccionado para desafiar la competencia lingüística del alumno.
2. **Inducción emocional:** Un cortometraje animado de tono sensible, cuidadosamente elegido para evocar la respuesta emocional objetivo.
3. **Interactividad:** Un juego educativo de sopa de letras personalizable, donde los estudiantes: seleccionan la temática, proponen sus propias palabras objetivo e interactúan con la solución generada dinámicamente.
4. **Desafío auditivo:** Un audio en lengua no nativa con complejidad controlada, diseñado para superar la zona de confort lingüístico.

Esta secuencia didáctica multicomponente crea una progresión intencional desde el estímulo emocional hasta la necesidad demostrable de asistencia.

Criterios para seleccionar la mejor emoción desencadenante.

El sitio organiza las actividades seleccionadas en secciones diferenciadas dentro del sitio web, donde cada una ocupa un apartado específico. Previo al inicio de la prueba, el sistema solicita al estudiante información demográfica relevante (nivel educativo, edad, sexo) para permitir la posterior clasificación de los datos en la base de datos. Al activar el botón 'Test' ubicado en la parte inferior de la interfaz, se inicia un proceso automatizado que abre una ventana emergente, activa la cámara del dispositivo y comienza el análisis en tiempo real de las expresiones faciales del alumno durante la resolución de los desafíos.

Para cuantificar las reacciones, se ha implementado un contador programado que registra la frecuencia con la que se detecta cada expresión facial. Esta codificación permite identificar si el estudiante requiere asistencia, lo cual constituye el objetivo central del diseño propuesto. Las expresiones faciales detectadas se almacenan en una base de datos, con el fin de determinar cuál es el desencadenante más efectivo. Es importante señalar que este mismo flujo de captura, análisis y registro se aplica de forma uniforme a las cuatro estrategias didácticas implementadas en el sitio, lo que garantiza consistencia en el proceso de evaluación.

El sitio web incorporó una sección especializada de análisis que presenta los resultados mediante un gráfico de pastel, el cual sintetiza visualmente el desempeño de las cuatro expresiones faciales más relevantes detectadas durante cada una de las actividades implementadas. Esta representación gráfica permite identificar de manera inmediata cuál expresión ha demostrado mayor predominancia estadística, así como las posibles combinaciones significativas entre ellas. El gráfico cumple una función decisiva en el proceso de selección, ya que transforma los datos crudos en información visualmente accesible que facilita la determinación del desencadenante más efectivo -ya sea una expresión individual o una combinación de ellas- basándose en su frecuencia e intensidad de aparición durante las interacciones de los estudiantes con las diferentes pruebas.

Diagrama de flujo: Cómo el agente decide intervenir (emociones + perfil → reglas → acción).

Organización de contenidos:

Cada actividad se aloja en secciones independientes dentro del sitio web.

Recolección inicial de datos:

Previo a la prueba, el estudiante proporciona información demográfica (nivel educativo, edad, sexo) para permitir clasificación analítica posterior.

Mecánica de evaluación:

Al activar el botón 'Test' (ubicado en la zona inferior de la interfaz), se:

Inicia una ventana emergente

Activa la cámara del dispositivo

Comienza el análisis en tiempo real de expresiones faciales

Sistema de registro:

Un contador programado registra la frecuencia de cada expresión facial detectada

Cuando se identifica la necesidad de asistencia (objetivo pedagógico), este conjunto de expresiones sirve como base para determinar el desencadenante óptimo

Diseño del sistema

Arquitectura técnica: TensorFlow.js + pipeline de datos (emociones + perfil).

Dataset: ¿Usarás FER2013 + datos propios de estudiantes?

Criterios para el "desencadenante óptimo"

Variables de entrada:

Emociones (Ekman) + métricas de perfil (ej. "tiempo en pregunta difícil").

Reglas de activación:

Ejemplo: "Si tristeza > umbral X Y tiempo en tarea > Y segundos → Activar tutor."

Resultados

Análisis de qué expresión fue más efectiva.

Relación con los perfiles del estudiante.

Discusión

Interpretación de los hallazgos.

Implicaciones pedagógicas.

Ventajas del enfoque web y del agente tutor activado por emociones.

Conclusiones

Recapitulación de los hallazgos.

Aplicaciones futuras.

Limitaciones y futuras líneas de investigación.

Referencias

- Amaya, A., & Cuéllar, A. (2016). Estilos de aprendizaje de los alumnos de posgrado a distancia de la Universidad Autónoma de Tamaulipas. *Apertura: Revista de Innovación Educativa*, 8-21. Obtenido de <https://hdl.handle.net/20.500.12802/9278>
- Captera. (27 de 04 de 2022). *Captera.com*. Recuperado el 25 de 04 de 2022, de Learning management system software: <https://www.capterra.com/learning-management-system-software/#shortlist>
- Castro, R., Martínez, M., Cruz, R., Juárez, C., & Soberanes, A. (2013). Sistema Adaptativo del Perfil del Estudiante. *4º Coloquio Internacional de Cómputo e Informática*, 91-102.
- Coronel Caján , E. (2021). *Universidad Señor de Sipán*. Obtenido de Reconocimiento de expresiones faciales de tristeza utilizando aprendizaje profundo: <https://repositorio.uss.edu.pe/bitstream/handle/20.500.12802/9278/Coronel%20Caj%C3%A1n%20Erick%20Arturo.pdf?sequence=1&isAllowed=y>
- Ekman, P. &. (1981). Expresiones faciales de la emoción. *Estudios de psicología*, 115-144. Obtenido de Recuperado de sitio web: http://www.academia.edu/download/49320332/Paul_Ekman_-_Estudio_Expresiones_Faciales_De_La_Emocion.pdf
- Ekman, P. (2016). What Scientists Who Study Emotion Agree About. *Perspectives on Psychological Science*, 31–34. doi:10.1177/1745691615596992

- González, H. &. (2008). Modelo dinámico del estudiante en cursos virtuales adaptativos utilizando técnicas de inteligencia artificial. *Avances en Sistemas e Informática*, 5(1), 199-206. Obtenido de <https://repositorio.unal.edu.co/handle/unal/70148>
- León-Nájera, M. (20 de 10 de 2021). Diseño del algoritmo del perfil del estudiante con inteligencia artificial, integrando el reconocimiento facial. Valle de Chalco, México, México. Obtenido de <http://ri.uaemex.mx/handle/20.500.11799/112765>
- OpenCV. (22 de Febrero de 2025). *Reconocimiento facial con OpenCV*. Obtenido de OpenCV: https://docs.opencv.org/4.10.0/da/d60/tutorial_face_main.html
- Reyes Fuentes, M. (2018). Objetos de Aprendizaje para Implementar en Moodle. *Revista Iberoamericana de Producción Académica y Gestión Educativa*, 46 - 67. Obtenido de <https://www.pag.org.mx/index.php/PAG/article/view/742>
- Rodríguez, J. S. (2005). Plataformas tecnológicas para el entorno educativo. *Acción pedagógica*, 18-24. Obtenido de <https://dialnet.unirioja.es/descarga/articulo/2968980.pdf>
- Romero-Torres, J., Martínez Reyes, M., & León-Nájera, M. (2024). EL CONTROL DE USUARIOS A BORDO DEL AUTOBUS EN EL TRANSPORTE. *Fortaleciendo la Competitividad en las*, 37-55. Obtenido de https://www.rilco.org/wp-content/uploads/2025/03/LibroMexicoCentro_Fortalecimiento.pdf
- tensorflow. (22 de Febrero de 2025). *Introducción a TensorFlow*. Obtenido de tensorflow: <https://www.tensorflow.org/learn?hl=es-419>
- Tümen, V. S. (2017). Facial emotion recognition on a dataset using convolutional neural network. *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*. doi:10.1109/idap.2017.8090281
- Vargas, J. D. (2016). Inteligencia emocional y percepción de las emociones básicas como un probable factor contribuyente al mejoramiento del rendimiento en las ventas: Una investigación teórica. *Universitas Psychologica*, 73-86. Obtenido de <https://revistas.javeriana.edu.co/index.php/revPsycho/article/view/8693>
- Vélez, J. B. (11 de 11 de 2009). Entorno de aprendizaje virtual adaptativo soportado por un modelo de usuario integral. Girona, España. Obtenido de <http://hdl.handle.net/10803/7607>