

**ANÁLISIS DE LA EXPERIENCIA EMOCIONAL EN USUARIOS DE RECURSOS
EDUCATIVOS ABIERTOS MEDIANTE RECONOCIMIENTO FACIAL DE
EXPRESIONES EN UN DISPOSITIVO JETSON NANO**



**Anteproyecto de trabajo de grado
Modalidad: Trabajo de Investigación**

**Nombre1 Nombre2 Apellido1 Apellido2
Nombre3 Nombre4 Apellido3 Apellido4**

Director: PhD. Nestor Milciades Díaz Mariño

Codirectora: PhD. Carolina Gonzalez Serrano

**UNIVERSIDAD DEL CAUCA
FACULTAD DE INGENIERÍA ELECTRÓNICA Y TELECOMUNICACIONES
DEPARTAMENTO DE SISTEMAS
INGENIERÍA DE SISTEMAS
POPAYÁN CAUCA
OCTUBRE 2023**

TABLA DE CONTENIDO

1. PLANTEAMIENTO DEL PROBLEMA.....	2
1.1. Pregunta de investigación.....	3
2. ESTADO DEL ARTE.....	3
2.1. Revisión Sistemática de la Literatura (RSL).....	3
2.2. Resultados.....	8
2.3. Conclusión.....	8
4. OBJETIVOS.....	9
4.1. General.....	9
4.2. Específicos.....	9
5. ACTIVIDADES Y CRONOGRAMA.....	10
5.1. Actividades.....	10
5.2. Cronograma.....	13
6. RECURSOS, PRESUPUESTO Y FUENTES DE FINANCIACIÓN.....	14
7. CONDICIONES DE ENTREGA.....	14
8. REFERENCIA BIBLIOGRÁFICA.....	15



1. PLANTEAMIENTO DEL PROBLEMA

Los recursos educativos abiertos (REA) hacen referencia a un término empleado por primera vez en la conferencia de la UNESCO del año 2002 para indicar aquellos materiales (libros, videos, audios y, cualquier elemento diseñado para la enseñanza y aprendizaje) que tienen una licencia especial para ser utilizados por educadores y estudiantes, sin necesidad de pagar regalías o derechos de autor [1].

Desde el surgimiento de la web 2.0 a mediados del año 2004 hasta la actualidad, la cantidad de recursos académicos digitales ha aumentado significativamente, debido a que el rol del profesorado ha sido modificado, pasando de ser un consumidor de materiales educativos, a formar parte de la creación y publicación de los mismos [2]. Por lo anterior, en [1] han sugerido la necesidad de asegurar la calidad de dichos recursos, ya que este cambio de paradigma deja a un lado a terceros como las editoriales educativas quienes antes se encargaban principalmente de esta tarea. Actualmente el deber de evaluar los REA recae sobre las instituciones educativas, coordinadores de programa/curso y los educadores responsables de la actividad docente en base a su contexto de aplicación.

En [3] han propuesto diversas maneras de evaluar los recursos educativos digitales, entre las que se encuentran: la usabilidad, dimensión pedagógica, seguimiento y control del estudiante, pertinencia de los contenidos, y metadatos. Durante la evaluación se involucra tanto a evaluadores como estudiantes, docentes e investigadores. Se usan como instrumentos de evaluación planillas, lista de preguntas y formatos con criterios de puntajes establecidos, en su mayoría de forma manual. Por lo anterior, es importante no solo analizar las características inherentes de los materiales, sino también, la influencia que estos tienen en el aprendizaje de los estudiantes, involucrando en el proceso a distintos perfiles profesionales de acuerdo a su rol frente a dichos recursos.

A fin de evaluar la calidad de los REA, uno de los criterios es determinar la experiencia emocional de los usuarios mientras hacen uso de los mismos, permitiendo identificar el estado emocional y cognitivo del usuario. Para realizar esta tarea de forma automática se han empleado distintos algoritmos de reconocimiento facial de expresiones¹ (FER) con un índice de precisión de hasta el 99% en escenarios controlados [5], que se basan principalmente en modelos de aprendizaje profundo (DL) como las redes neuronales convolucionales (CNN).

Para obtener resultados aceptables en el menor tiempo posible con algoritmos FER, se requiere que sean ejecutados en máquinas con hardware dedicado. Por ese motivo, en este trabajo se propone hacer uso de dispositivos contruidos sobre una sola placa de circuito [8], los cuales son computadoras pequeñas con alta capacidad de cómputo y en similitud a sistemas embebidos pueden ser configurados para cubrir necesidades específicas en tiempo real. Estos dispositivos en contraste a las computadoras tradicionales, tienen la cualidad de ser: más veloces, menor tamaño físico, mayor cantidad de puertos, facilidad de integración con varios periféricos y bajo costo [6].

¹ FER es un proceso de caracterización de rostros humanos en base a imágenes y/o videos, que permite determinar emociones básicas como felicidad, disgusto, ira, tristeza, miedo y asco, planteadas por Friesen et al. [4].

Entre los dispositivos de este tipo se pueden mencionar al Raspberry PI 3 [7], Raspberry PI 4, Jetson Nano y Jetson TX2 (mencionados en orden de costo ascendente), la familia Jetson es la única en el mercado que incorpora nativamente una unidad de procesamiento gráfico (GPU). En las pruebas presentadas en [8], con un modelo de CNN para el reconocimiento de prendas de vestir y diversos tamaños de dataset, Jetson Nano obtiene una precisión similar a Jetson TX2 variando en apenas $\pm 0.1\%$, mientras que al compararlo con Raspberry PI 4, Jetson Nano obtuvo mejor desempeño entre 0.3% y 2.3%. En cuanto al tiempo de cálculo utilizando un dataset de 10k imágenes, Jetson Nano tarda 314 segundos menos que Raspberry PI 4, y 26 segundos más a comparación del Jetson TX2. De acuerdo a lo anterior, se toma la decisión de elegir al Jetson Nano para la implementación de un prototipo en base al costo/beneficio que éste ofrece al compararlo con otros sistemas.

Finalmente, en [9][10][11] han realizado exitosamente la implementación de algoritmos FER mediante modelos de CNN para la identificación de emociones durante el uso de recursos educativos, sin embargo, el despliegue de estos algoritmos se realizó en computadoras modestas de escritorio y laptops, limitando así su precisión en los resultados. Para complementar, en [12][13] hicieron uso de soluciones API para la captura de emociones, lo cual restringe el uso de esas herramientas en situaciones donde se requiere una conexión constante a internet y pago por el servicio. Concluyendo, el análisis de los datos recopilados en los estudios nombrados fueron hechos a través de técnicas estadísticas tradicionales y con datos obtenidos en pocas sesiones de evaluación de los materiales educativos digitales, evidenciando de esta manera la falta de un conjunto de datos robusto y un método de análisis automatizado.

1.1. Pregunta de investigación

¿Es posible determinar la experiencia emocional del usuario durante el uso de un recurso educativo abierto (REA) mediante una herramienta software con reconocimiento facial de expresiones (FER)?

2. ESTADO DEL ARTE

2.1. Revisión Sistemática de la Literatura (RSL)

Para este punto se realizó una RSL en la base de datos Scopus, ScienceDirect y Web of Science, que tuvo la siguiente pregunta de investigación:

PQ1: ¿Cuáles técnicas de aprendizaje de máquina se usan para determinar dentro de un marco educativo la experiencia emocional mediante el reconocimiento facial?

Además, se formulan las siguientes preguntas de mapeo:

- **MQ1:** ¿Cuál es la técnica de ML más usada para la detección de emociones usando FER?
- **MQ2:** ¿En qué tipo de dispositivos realizan el despliegue de los algoritmos FER?
- **MQ3:** ¿Qué tipo de métricas establecen para indicar la experiencia emocional?
- **MQ4:** ¿Cuáles técnicas de análisis de datos usan para estudiar las emociones?

A partir de la pregunta de investigación propuesta para la RSL, se generó la cadena de búsqueda:

(algorithm OR "machine learning") AND (technique OR method) AND ((detection OR recognition OR analysis) AND (emotion OR sentiment)) AND ("facial expression recognition" OR fer) AND (education OR student OR classroom)

Los siguientes resultados responden a la pregunta PQ1 y están acotados en los últimos 7 años (2015 - 2022) para un total de 21 artículos, se organizan por tipo de técnica de ML y cronológicamente. Se sintetizan principalmente de acuerdo al problema planteado, la técnica de ML propuesta y los resultados.

Redes Neuronales Convolucionales

En 2019, J. Pei y P. Shan [9], desarrollan un algoritmo basado en CNN FlowNetC que analiza las microexpresiones de estudiantes en un salón de clases, las emociones básicas se agrupan y clasifican en "agrado", "entusiasmo" e "interés", a partir del porcentaje de las emociones complejas se genera un puntaje con la media de todos los resultados de los estudiantes durante la clase y se separan en 4 niveles: muy desconcentrado, desconcentrado, concentrado y muy concentrado.

En 2020, T. Liu y J. Wang [10], plantean el problema de no adaptar dinámicamente los métodos pedagógicos pertinentes en función de las emociones y el comportamiento cambiante del alumno en el aprendizaje en línea. Como solución propusieron una innovadora técnica que toma una secuencia de 6 imágenes (frames) de un video, cada imagen identifica porcentualmente una de las emociones básicas y al final las categoriza en "confuso", "satisfecho", "insatisfecho" y "frustrado". Usan el API de Windows Azure para el reconocimiento facial y extracción de características faciales, después se procesa la data obtenida en la CNN diseñada.

En 2020, H. Lee y D. Lee [11], este estudio enfatiza la evaluación del estudiante centrada en el proceso (PFA) y no en resultados. Se diseña un marco de trabajo con algoritmos de AdaBoost y Haar feature-based para la extracción de puntos característicos de las regiones faciales y, una CNN para la clasificación de niveles de dificultad (fácil, difícil y neutral). Las pruebas de la herramienta se llevaron a cabo con imágenes capturadas desde la webcam de una laptop en tiempo real de estudiantes de pregrado y postgrado de carreras afines a la computación que resuelven ejercicios básicos de programación en una aplicación interactiva.

En 2020, W. Wang, K. Xu [4], diseñan un marco de trabajo para analizar las emociones de estudiantes en clases virtuales, emplean la herramienta IntraFace para la extracción de características de rostros faciales y posteriormente pasan la información a un modelo de CNN para la detección de emociones básicas, al final el software presenta al tutor un histograma con los resultados de la lección, a fin de que pueda ajustar sus estrategias de enseñanza.

En 2021, T. Liu y J. Wang [15], realizan la captura de imágenes en con una cámara infrarroja argumento que este tipo de dispositivos pueden aumentar la precisión de los algoritmos de aprendizaje profundo basados en FER (DLB) en el escenario donde la captura de los rostros es en un ambiente oscuro, proponen solucionar el problema de las expresiones faciales ambiguas etiquetando las imágenes en base

a emociones con una probabilidad de Churly, permitiendo una mejora en comparación a trabajos previos. Para la detección de emociones se usa un modelo CNN basado en GoogleNet y se entrena con los dataset de Oulu-CASIA y CK+, obteniendo un rendimiento del 75.71% y 83.14% respectivamente.

En 2021, Y. Bhatti y A. Jamil [16], trabajan en un framework para analizar las emociones de los instructores mientras dictan una sesión de enseñanza. Se afirma que este tipo de problemas no ha sido explorado previamente y por lo tanto, requieren diseñar una BD para dicho fin. Emplean Viola Jones para la detección de rostros y una CNN basada en DenseNet201 para la detección de emociones, que arroja una precisión del 62.50% para el dataset FER2013 y del 96.80% para el dataset JAFFE. Los resultados del estudio sólo indican el nivel de precisión del modelo para identificar las emociones básicas.

En 2021, K. Nithiyasree y A. Nisha [17] elaboran un estudio básico a fin de detectar el nivel de interés de estudiantes en una sección de lectura haciendo uso de la predicción de emociones faciales. Se implementa una arquitectura de CNN confiable como la VGG-16 para la detección de emociones, se entrena con el dataset FER2013 además con las imágenes de los participantes para identificarlos, y obtienen un rendimiento del modelo igual a 89%. Los resultados son presentados en un gráfico circular que indica el porcentaje de tiempo que cada emoción estuvo presente en la sesión de prueba.

En 2021, Y. Cui y S. Wang [18], plantean la importancia de caracterizar las emociones a partir de una secuencia de imágenes en una CNN 3D (tercera dimensión es el tiempo), sin embargo, las pruebas realizadas se hacen con imágenes estáticas. Explican a detalle el marco de trabajo, que indica el momento oportuno donde el tutor debe aumentar o disminuir la dificultad de la sesión. No se explica la arquitectura de la CNN propuesta. Plantean evaluar al usuario de la herramienta por la categorización ponderada de emociones básicas como “positiva” y “negativa”. Al final no evalúan el framework propuesto.

En 2021, M. Hasnine y H. Bui [19], se centran en detectar el nivel de compromiso de los estudiantes en un entorno virtual, detallan la arquitectura, usan OpenCV para la detección de rostros, dlib para identificar a la persona, CNN para la detección de emociones y para la detección de los ojos extraen la información de los puntos de referencia facial. Al final se obtiene un puntaje que resulta de la multiplicación del peso de la emoción detectada y la apertura de los ojos, que deriva en una clasificación de “altamente comprometida”, “comprometida” y “no comprometida”.

En 2021, T. Abdallah y I. Elleuch [20], identifican el problema de no contar con suficientes BD de datos públicos para entrenar modelos CNN para la detección de emociones. Para hacer frente a la problemática usan la técnica de transferencia de aprendizaje, primeramente toman como base la CNN de VGG-16, la modifican y entrenan para el propósito de identificar emociones con el dataset FER2013, finalmente la re-entrenan con un propio dataset el cual llamaron SBC2020. Por lo general la mayoría de modelos CNN entrenados y probados con FER2013 dan una precisión entre 60% y 75%, este modelo alcanzó una precisión del 79.4% demostrando que la solución propuesta es eficiente.

En 2022, C. Thomas y K. Punneeth Sarma [14], estudian el comportamiento emocional de estudiantes y determinan si el estilo de presentación de una sesión de aprendizaje afecta su nivel de compromiso. El proceso de detección de emociones se realiza con una CNN con precisión del 76% que analiza una secuencia de imágenes, las emociones se clasifican en "comprometido" o "distruido". En conclusión, el estilo de presentación del maestro no influye estadísticamente en el compromiso del estudiante.

En 2022, C. Hou y J. Ai [5], detectan el problema que los docentes tienen en las clases en línea al no entender el estado de escucha de los estudiantes en tiempo real, como solución se propone emplear el análisis de emociones mediante el uso de la CNN llamada VGG-16 combinada con ECANet, este modelo propuesto se entrena con el dataset CK+ dando una precisión del 99.18%, para el pre-procesamiento de imágenes usan OpenCV con otra CNN nombrada MTCNN. Como aporte del estudio, se especifica una manera de calcular una puntuación que refleja el entusiasmo y la concentración del estado de escucha del estudiante.

En 2022, S. Batra y H. Wang [21], estudian la combinación de técnicas de reconocimiento facial y detección de emociones para reconocer el nivel de compromiso de estudiantes durante el aprendizaje en línea. Las técnicas evaluadas son CNN y SVM, con HOG, SIFT Y SURF. Al final indican que la técnica con mejor precisión es ResNet-18 con un 80%, además los autores aclaran que los algoritmos empleados no pueden diferenciar entre un estudiante completamente comprometido con uno que lo esté parcialmente, ya que no existen límites claros para separar las dos clases.

En 2022, X. Li [22], propone un modelo de CNN basado en LeNet para mejorar la precisión en la detección de emociones básicas, el artículo detalla cada una de las partes del modelo. Para evaluar la CNN diseñada, emplean el dataset CK+ y obtienen una tasa de reconocimiento incrementada del 2.53% a comparación de LeNet, sin embargo, consume más tiempo.

En 2022, X. Qiao y X. Zheng [23], diseñan un sistema para el seguimiento en tiempo real de estados emocionales de los alumnos durante sesiones de aprendizaje. Implementan un marco de trabajo que incluye la detección de emociones, posición y movimiento de cabeza y ojos, mediante un modelo de CNN. Finalmente la información de cada sesión se almacena en una BD y, se genera un reporte con indicadores de puntaje para estados de "activo", "neutral", "negativo" y "sueño". El autor precisa que por motivos de limitaciones técnicas, el algoritmo evaluado no es muy preciso.

Redes Neuronales Recurrentes

En 2015, J. Khalfallah y J. Slama [24], crean una herramienta web hecha en Java EE para hacer uso remoto de laboratorios de electrónica, los estudiantes acceden desde sus laptops, el sistema captura las imágenes de los rostros mediante una librería de JavaScript que implementa un algoritmo de Constrained Local Models (CLM) que integra un modelo basado en aprendizaje automático para detectar las emociones, después las emociones básicas se agrupan y, clasifican en "entendió" o "no entendió", al final de la sesión los usuarios responden a un cuestionario para calificar la experiencia con el software.

En 2018, D. Yang, A. Alsaddon y P. Prasad [25], proponen un marco de trabajo para resolver problemas de reconocimiento de emociones basados en FER en un entorno de aprendizaje virtual, hacen uso de la técnica Haar cascaded para identificar los ojos y boca de un rostro, y clasifican las imágenes mediante una RNN compuesta de 1 capa de entrada, 20 ocultas y de 7 salida (emociones básicas). Los resultados indican la precisión del modelo para identificar cada emoción básica y el tiempo de procesamiento.

Soluciones API (modelos privados)

En 2022, I. Dubovi [12], analiza el desarrollo del nivel de compromiso y cómo este puede afectar los logros de aprendizaje en estudiantes de medicina que usan una aplicación simulada en VR. Capturan las emociones básicas y movimientos de los ojos, mediante el software Imotions, además complementan la detección de emociones (complejas) con auto encuestas a los participantes, posteriormente examinan la información con técnicas estadísticas como ANOVA. Concluyen afirmando la importancia de combinar técnicas automáticas objetivas y subjetivas para el análisis de emociones.

En 2020, G. Tonguc y B. Ozaydin Ozkara [13], realizan este estudio bajo la premisa de la importancia de las emociones en la educación. Utilizan la herramienta de Microsoft Emotion Recognition API para la detección de emociones durante la impartición de una conferencia a estudiantes universitarios, con los datos obtenidos de las emociones básicas se realizó un análisis mediante la técnica estadística Manova, también consideraron factores como la hora del día, el perfil de los estudiantes, género, entre otras. Concluyen con pautas a los tutores a fin de mantener la atención constante de los estudiantes.

En 2022, F. Otamendi [26], afirma que el post-procesamiento de la información obtenida tras la detección de emociones aún se encuentra inmadura, por eso propone usar el método de control de proceso estadístico (SPC) utilizado ampliamente en la industria, para hallar comportamientos anómalos en un proceso, en este caso de las emociones. La identificación de las emociones se realiza con el software Imotions. Al final da como resultado un conjunto de gráficas que indican la cantidad e intensidad de emociones de un individuo.

Otras técnicas

En 2016, Krithika L.B y Lakshmi Priya [27], diseñan un sistema para detectar el nivel de concentración de estudiantes durante una sesión de clases, el análisis para determinar cuán concentrado está el usuario se hace por medio de la detección de movimientos y posición, de la cabeza y cada ojo. El algoritmo implementado usa Viola Jones y Local Binary Patterns (LBP). Al finalizar la sesión, el sistema entrega un reporte general e individual de los estudiantes, e indica al tutor en cuáles momentos no prestaron atención.

2.2. Resultados

MQ1: ¿Cuál es la técnica de ML más usada para la detección de emociones usando FER?

Dentro del contexto educativo, se ha observado que la técnica de ML más ampliamente adoptada con cerca del 72% de los trabajos presentados, son las CNN que implementan modelos con arquitecturas como VGG16 [5], GoogleNet [21] y DenseNet201 [15]. Esta técnica ha demostrado ser altamente eficaz presentando índices de precisión que alcanzan hasta el 99% en ambientes controlados.

MQ2: ¿En qué tipo de dispositivos realizan el despliegue de los algoritmos FER?

En ninguna investigación dentro del ámbito educativo se halló implementación de técnicas de ML incorporadas en dispositivos de cómputo de bajo costo, únicamente en estaciones de trabajo (PCs de escritorio y laptops) [4][24][27] y servidores en la nube [26].

MQ3: ¿Qué tipo de métricas establecen para indicar la experiencia emocional?

La mayoría de los estudios utilizan la detección de emociones básicas como indicador del nivel de concentración del estudiante, ya sea en entornos virtuales o presenciales mediante cámaras incorporadas en computadoras. Algunos trabajos [11][24] van más allá al agrupar estas emociones básicas en emociones complejas, y solo en [5] se presenta una función ponderada que cuantifica el entusiasmo y la concentración del estudiante. En contraste, en [19][27] evalúan la concentración del estudiante considerando la posición y el movimiento de la cabeza y los ojos.

MQ4: ¿Cuáles técnicas de análisis de datos usan para estudiar las emociones?

Cerca del 48% de los estudios realiza un post-procesamiento de las emociones obtenidas por los modelos de ML. En [12][13][14] emplean técnicas estadísticas sofisticadas como MANOVA y SPC para evaluar la variación de las emociones entre grupos de estudiantes y las fluctuaciones emocionales en el tiempo, mientras que otros [4][17][19][27] optan por un análisis más simple que implica la representación gráfica de cambios emocionales a lo largo del tiempo y acumuladores porcentuales de estas emociones.

2.3. Conclusión

Es evidente que los modelos de ML destinados a FER y a la captura de gestos en tiempo real han alcanzado un nivel de precisión notable. Sin embargo, en contraste, se percibe una falta de madurez en el post-procesamiento de los resultados obtenidos por dichos modelos. Por lo tanto, se abre un campo de estudio para impulsar la creación de un conjunto de datos sólido que incluya además los tiempos de respuesta e interacción con el contenido, y sirva como base para el desarrollo de métricas objetivas capaces de cuantificar la experiencia emocional de un usuario durante la interacción con recursos educativos. En este contexto, la minería de datos se presenta como una técnica fundamental para la generación y análisis de las métricas propuestas. Mediante la explotación de grandes volúmenes de datos [29] recopilados durante el uso de recursos educativos, es posible identificar patrones, tendencias y relaciones ocultas, que podrían arrojar luz sobre la experiencia emocional.

3. APORTES

El aporte de la propuesta de trabajo consiste en elaborar un prototipo que incorpore un algoritmo de ML desplegado en un dispositivo de bajo costo Jetson Nano, que determine de manera innovadora la experiencia emocional del usuario mientras hace uso de los REA empleando técnicas de minería de datos, a partir de información obtenida de las características y gesticulaciones faciales.

El resultado de este trabajo permitirá la integración de las líneas en Inteligencia Artificial y Educación, con la de Ciencia de Datos del Grupo de Investigación en Inteligencia Computacional (GICO) para crear mayor sinergia al interior del mismo, abriendo la posibilidad de desarrollar una área de aplicación que facilite la incorporación de soluciones de Inteligencia Artificial al ámbito educativo.

4. OBJETIVOS

4.1. General

Diseñar una herramienta software que determine la experiencia emocional de usuarios de REA en un dispositivo de cómputo de bajo costo, mediante reconocimiento facial de expresiones y análisis de datos, para apoyar el ejercicio docente en la adopción de materiales educativos.

4.2. Específicos

- Implementar un modelo de ML para el reconocimiento facial y detección de emociones, seleccionado a partir de la caracterización de modelos de una revisión sistemática de la literatura.
- Desplegar el modelo de ML seleccionado en un dispositivo de bajo costo Jetson Nano, incluyendo una aplicación para la interacción del usuario con el REA.
- Implementar una estrategia de analítica de datos basada en técnicas de minería de datos que permita caracterizar la experiencia emocional de los usuarios durante el uso de REA, a partir de la información de reconocimiento facial y detección de emociones.
- Evaluar el desempeño del prototipo propuesto mediante la realización de experimentos con usuarios que permitan verificar el correcto reconocimiento facial y la identificación de emociones durante el uso de un REA.

5. ACTIVIDADES Y CRONOGRAMA

5.1. Actividades

Para el desarrollo de la herramienta software de reconocimiento facial y detección de emociones en usuarios de REA, se plantea seguir el método del marco de trabajo para la Investigación en Ciencias del Diseño [28].

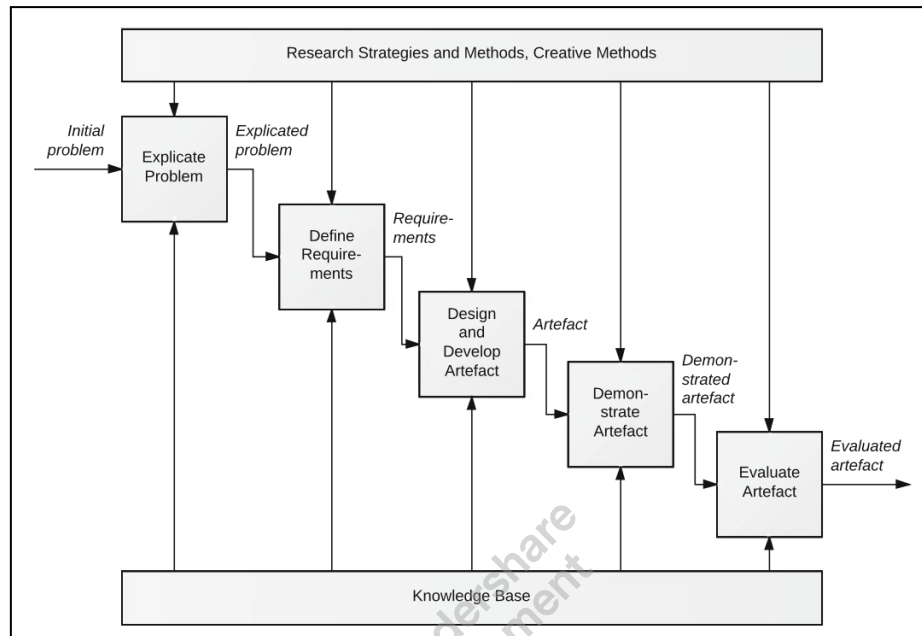


Figura 2. Actividades propuestas en Investigación en Ciencias del Diseño [26]

I. Explicación del problema

Gracias a la adopción paulatina del concepto REA en todo el mundo, se ha acelerado la frecuencia de generación y publicación de recursos académicos digitales. Este crecimiento ha planteado un nuevo reto tanto para creadores, como consumidores de este tipo de materiales, debido a que no ha sido sencillo asegurar la calidad de los mismos no solo por la cantidad excesiva, sino también, porque depende de cómo se decide adaptarlos para que sean contextualmente relevantes en su aplicación [1].

II. Definición de requerimientos

Este trabajo propone diseñar y construir un prototipo con el dispositivo de bajo costo Jetson Nano, que incluya un algoritmo FER basado en un modelo de ML y un modelo de minería de datos capaz de indicar por medio de reportes la experiencia emocional de los usuarios de REA.

Se identifican dos tipos de usuarios para el uso del prototipo: el evaluador (docente, investigador) y un sujeto (estudiante) a quien se pondrá a prueba los materiales académicos.

Para el evaluador, el prototipo deberá permitir la realización de configuraciones y despliegue de los REA. A su vez para el sujeto, el prototipo deberá brindar una visualización interactiva con el REA mediante periféricos como teclado y, un dispositivo apuntador (ratón) o directamente en una pantalla táctil.

III. Diseño y desarrollo del artefacto

Se diferencian dos etapas principales para el desarrollo del prototipo:

La primera etapa incluirá la exploración e implementación de modelos de ML para el análisis de imágenes mediante la técnica FER. Incluyendo la configuración necesaria para el despliegue en el dispositivo Jetson Nano.

En la segunda etapa se propone realizar un modelo de minería de datos para hallar patrones en los datos recopilados por el modelo de ML y diseñar un reporte con los resultados del proceso de minería. Esta actividad se llevará a cabo con la metodología CRISP-DM [29].

A. Entendimiento del negocio

El reconocimiento facial de expresiones (FER) en [19][27] se ha empleado para identificar emociones en estudiantes durante el uso de recursos educativos. Los resultados de esta técnica se usan para realizar análisis estadísticos manuales, que informan al evaluador de fluctuaciones en el comportamiento emocional de los individuos.

Objetivo del negocio

- Diseñar un reporte que indique la experiencia emocional de usuarios de REA.

Objetivos de la minería de datos

- Generar un conjunto de datos etiquetados y no etiquetados de emociones que indiquen la experiencia positiva o negativa de un usuario de REA.
- Validar el desempeño de los modelos generados.

Criterios de éxito de la minería de datos

- Los evaluadores y docentes puedan interpretar coherentemente los resultados plasmados en el reporte diseñado.

B. Entendimiento de los datos

Se plantea recopilar un conjunto de datos a partir de los resultados del modelo de ML que se pondrá a prueba en usuarios de distintos niveles educativos. En colaboración con el personal docente, se establecerá cuáles variables adyacentes pueden aportar más información a fin de entender los resultados de cada sesión. Adicionalmente, se etiquetarán los datos obtenidos mediante evaluaciones que indicarán el grado de comprensión de los usuarios frente al REA.

C. Preparación de datos

Con ayuda de herramientas de visualización y gestión de datos, se propone realizar un primer acercamiento para hallar patrones evidentes entre las variables, eliminar y/o completar valores vacíos, y desechar variables que no aporten mayor información.

D. Desarrollo de modelos

En este punto, se llevará a cabo la selección de diversas técnicas con el propósito de analizar y comprender en profundidad la experiencia emocional del usuario. Se aplicarán estrategias de análisis de tendencias temporales, segmentación de usuarios, análisis de correlación con otros factores relevantes y visualización de datos emocionales. Estas técnicas permitirán desentrañar patrones, tendencias y relaciones complejas en los datos emocionales, lo que enriquecerá la comprensión de la experiencia emocional de los usuarios y respaldará la toma de decisiones informadas para adaptar estrategias y mejorar sus necesidades.

E. Evaluación

Se evaluarán los distintos modelos que se implementaron en el punto anterior a fin de medir la precisión de cada uno al momento de calificar la experiencia emocional de un usuario REA en variedad de situaciones.

F. Despliegue

En este punto se definirá el trabajo necesario para hacer accesible el modelo seleccionado desde el dispositivo de cómputo de bajo costo. Se evaluarán las compatibilidades de software y hardware.

IV. Demostración del artefacto

Para esta actividad se espera diseñar un caso de uso donde los docentes modifiquen y adapten junto al investigador una variedad de recursos educativos tomados del portal Colombia Aprende². Después se seleccionarán algunos estudiantes quienes individualmente pasarán a resolver un conjunto de actividades frente al prototipo mediante los periféricos que se adapten para la interacción con el REA. Al finalizar cada sesión con el estudiante se realizarán una serie de preguntas cortas con el objetivo de conocer su experiencia en ciertos momentos de la prueba, esto permitirá comparar las emociones reales percibidas con el resultado obtenido del modelo de minería de datos.

V. Evaluación del artefacto

De acuerdo a las actividades de demostración se pretende realizar al menos dos sesiones por cada estudiante, esto con la finalidad de ir mejorando en cada iteración el modelo de minería de datos y el diseño del reporte. Cuando se finalicen todas las sesiones, se consultará con los docentes si el prototipo logra evaluar correctamente la calidad de los REA, si es el caso, entonces indagar sobre la viabilidad de aplicarlo en un entorno real.

² "Portal CREA - Centro de Recursos Educativos Abiertos | Colombia Aprende." <https://www.colombiaprende.edu.co> (accessed Oct. 07, 2023).

[illegible]

6. RECURSOS, PRESUPUESTO Y FUENTES DE FINANCIACIÓN

En la *Tabla 2* se describen los gastos requeridos para el desarrollo del proyecto.

Según el Decreto 1279 de 2002 define el valor del punto salarial en dieciocho mil ochocientos cuarenta y cinco pesos colombianos (\$18.845). Considerando un plazo de 26 semanas para el desarrollo del trabajo de grado, se tiene en cuenta el tiempo de trabajo para el director y el codirector de 2 horas por semana con un valor de referencia de 2.5 puntos y, el estudiante de 26 horas por semana con un valor de referencia de 1.5 puntos.

Tabla 2. Presupuesto

RUBROS	FUENTES		TOTAL
	ESTUDIANTES	DEPARTAMENTO	
Personal	\$19.108.830,00	\$4.899.700,00	\$24.008.350,00
Equipo	\$350.000,00	\$950.000,00	\$1.300.000,00
Software	\$0,00	\$0,00	\$0,00
Viajes y salidas de campo	\$0,00	\$0,00	\$0,00
Bibliografía	\$0,00	\$0,00	\$0,00
Materiales	\$0,00	\$0,00	\$0,00
Servicios técnicos	\$0,00	\$0,00	\$0,00
Publicaciones	\$0,00	\$0,00	\$0,00
Administración	\$0,00	\$0,00	\$0,00
Comunicaciones	\$0,00	\$0,00	\$0,00
Otros	\$1.770.000,00	\$0,00	\$1.770.000,00
TOTAL	\$21.228.830,00	\$5.849.700,00	\$27.078.530,00

Los rubros de “otros” incluyen derecho de grado y matrícula del estudiante.

7. CONDICIONES DE ENTREGA

Una vez terminado el trabajo de grado se obtendrán los siguientes productos:

- **Monografía del trabajo de grado**, contendrá especificaciones sobre la realización y despliegue de los modelos de FER y minería de datos en el dispositivo Jetson Nano, se planifica estructurar con los siguientes puntos: 1. Introducción (Planteamiento del problema, justificación, objetivos y resultados), 2. Marco teórico, 3. Estado del arte, 4. Metodología, 5. Análisis de resultados, 6. Conclusiones y trabajo futuro y 7. Referencias y anexos.
- **Artículo**, un artículo de investigación con el desarrollo del prototipo y los resultados obtenidos, considerando el formato de la IEEE.
- **Prototipo**, una herramienta software desplegada en un dispositivo de cómputo de bajo costo con reconocimiento facial de expresiones capaz de indicar la experiencia emocional de usuarios de REA.

8. REFERENCIA BIBLIOGRÁFICA

- [1] N. Butcher, A. Kanwar, S. UvalicTrumbic, Commonwealth of Learning., and UNESCO., "Guía básica de recursos educativos abiertos (REA)," 2015, Accessed: Nov. 03, 2022. [Online]. Available: <https://unesdoc.unesco.org/ark:/48223/pf0000232986>
- [2] C. J. González Ruiz, S. Martín Gómez, and A. Vega Navarro, "Portales educativos: la producción de materiales didácticos digitales," @tic. Rev. d'innovació Educ. ISSN-e 1989-3477, N°. 20, 2018 (Ejemplar Dedic. a Spring (January-June)), págs. 89-97, no. 20, pp. 89–97, 2018, doi: 10.7203/attic.20.12139.
- [3] I. Aguilar Juárez, J. A. De La Vega, O. Lugo Espinosa, and A. Z. Hidalgo, "Análisis de criterios de evaluación para la calidad de los materiales didácticos digitales Analysis of evaluation criteria for the quality of digital learning materials," Rev. CTS, n°, vol. 25, pp. 73–89, 2014.
- [4] W. Wang, K. Xu, H. Niu, and X. Miao, "Emotion Recognition of Students Based on Facial Expressions in Online Education Based on the Perspective of Computer Simulation," Complexity, vol. 2020, 2020, doi: 10.1155/2020/4065207.
- [5] C. Hou, J. Ai, Y. Lin, C. Guan, J. Li, and W. Zhu, "Evaluation of Online Teaching Quality Based on Facial Expression Recognition," Futur. Internet, vol. 14, no. 6, 2022, doi: 10.3390/fi14060177.
- [6] Á. Guerrero Aguirre, P. Jimena, and R. Giraldo, "Sistema embebido de bajo costo para visión artificial Low cost embedded system for machine vision," Sci. Tech. Año XIX, vol. 19, no. 2, 2014.
- [7] T. A. Salih and M. Basman Gh, "A novel Face Recognition System based on Jetson Nano developer kit," IOP Conf. Ser. Mater. Sci. Eng., vol. 928, no. 3, 2020, doi: 10.1088/1757-899X/928/3/032051.
- [8] A. A. Suzen, B. Duman, and B. Sen, "Benchmark Analysis of Jetson TX2, Jetson Nano and Raspberry PI using Deep-CNN," HORA 2020 - 2nd Int. Congr. Human-Computer Interact. Optim. Robot. Appl. Proc., pp. 3–7, 2020, doi: 10.1109/HORA49412.2020.9152915.
- [9] J. Pei and P. Shan, "A Micro-expression Recognition Algorithm for Students in Classroom Learning Based on Convolutional Neural Network," Trait. du Signal, vol. 36, no. 6, pp. 557–563, Dec. 2019, doi: 10.18280/ts.360611.
- [10] M. Mukhopadhyay, S. Pal, A. Nayyar, P. K. D. Pramanik, N. Dasgupta, and P. Choudhury, "Facial Emotion Detection to Assess Learner's State of Mind in an Online Learning System," ACM Int. Conf. Proceeding Ser., pp. 107–115, Feb. 2020, doi: 10.1145/3385209.3385231.
- [11] H. J. Lee and D. Lee, "Study of Process-Focused Assessment Using an Algorithm for Facial Expression Recognition Based on a Deep Neural Network Model," Electron. 2021, Vol. 10, Page 54, vol. 10, no. 1, p. 54, Dec. 2020, doi: 10.3390/ELECTRONICS10010054.
- [12] I. Dubovi, "Cognitive and emotional engagement while learning with VR: The perspective of multimodal methodology," Comput. Educ., vol. 183, p. 104495, 2022, doi: <https://doi.org/10.1016/j.compedu.2022.104495>.
- [13] G. Tonguç and B. Ozaydın Ozkara, "Automatic recognition of student emotions from facial expressions during a lecture," Comput. Educ., vol. 148, p. 103797, Apr. 2020, doi: 10.1016/J.COMPEDU.2019.103797.
- [14] C. Thomas, K. A. V Puneeth Sarma, S. Swaroop Gajula, and D. B. Jayagopi, "Automatic prediction of presentation style and student engagement from videos," Comput. Educ. Artif. Intell., vol. 3, p. 100079, 2022, doi: <https://doi.org/10.1016/j.caeai.2022.100079>.

- [15] T. Liu, J. Wang, B. Yang, and X. Wang, "Facial expression recognition method with multi-label distribution learning for non-verbal behavior understanding in the classroom," *Infrared Phys. Technol.*, vol. 112, p. 103594, Jan. 2021, doi: 10.1016/J.INFRARED.2020.103594.
- [16] Y. K. Bhatti, A. Jamil, N. Nida, M. H. Yousaf, S. Viriri, and S. A. Velastin, "Facial Expression Recognition of Instructor Using Deep Features and Extreme Learning Machine," *Comput. Intell. Neurosci.*, vol. 2021, 2021, doi: 10.1155/2021/5570870.
- [17] K. Nithiyasree, A. Nisha, S. Shankar, N. Akshaykumar, and T. Kavitha, "Facial Emotion Recognition of Students using Deep Convolutional Neural Network," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 10, pp. 1430–1434, Apr. 2021, doi: 10.17762/TURCOMAT.V12I10.4462.
- [18] Y. Cui, Y. Cui, S. Wang, and R. Zhao, "Machine Learning-Based Student Emotion Recognition for Business English Class," *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 12, pp. 94–107, 2021.
- [19] M. N. Hasnine, H. T. T. Bui, T. T. Thu Tran, H. T. Nguyen, G. Akçapınar, and H. Ueda, "Students' emotion extraction and visualization for engagement detection in online learning," *Procedia Comput. Sci.*, vol. 192, pp. 3423–3431, 2021, doi: <https://doi.org/10.1016/j.procs.2021.09.115>.
- [20] T. Ben Abdallah, I. Elleuch, and R. Guermazi, "Student Behavior Recognition in Classroom using Deep Transfer Learning with VGG-16," *Procedia Comput. Sci.*, vol. 192, pp. 951–960, 2021, doi: <https://doi.org/10.1016/j.procs.2021.08.098>.
- [21] S. Batra et al., "DMCNet: Diversified model combination network for understanding engagement from video screengrabs," *Syst. Soft Comput.*, vol. 4, p. 200039, 2022, doi: <https://doi.org/10.1016/j.sasc.2022.200039>.
- [22] X. Li, "Expression Recognition of Classroom Children's Game Video Based on Improved Convolutional Neural Network," *Sci. Program.*, vol. 2022, 2022, doi: 10.1155/2022/5203022.
- [23] X. Qiao, X. Zheng, X. Sun, S. Li, and Y. Zhang, "Learners' States Monitoring Method Based on Face Recognition Technology," *Procedia Comput. Sci.*, vol. 202, pp. 172–177, 2022, doi: <https://doi.org/10.1016/j.procs.2022.04.024>.
- [24] J. Khalfallah and J. B. H. Slama, "Facial Expression Recognition for Intelligent Tutoring Systems in Remote Laboratories Platform," *Procedia Comput. Sci.*, vol. 73, pp. 274–281, 2015, doi: <https://doi.org/10.1016/j.procs.2015.12.030>.
- [25] D. Yang, A. Alsadoon, P. W. C. Prasad, A. K. Singh, and A. Elchouemi, "An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment," *Procedia Comput. Sci.*, vol. 125, pp. 2–10, 2018, doi: <https://doi.org/10.1016/j.procs.2017.12.003>.
- [26] F. J. Otamendi, "Statistical emotion control: Comparing intensity and duration of emotional reactions based on facial expressions," *Expert Syst. Appl.*, vol. 200, p. 117074, 2022, doi: <https://doi.org/10.1016/j.eswa.2022.117074>.
- [27] Krithika L.B and Lakshmi Priya GG, "Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric," *Procedia Comput. Sci.*, vol. 85, pp. 767–776, 2016, doi: <https://doi.org/10.1016/j.procs.2016.05.264>.
- [28] P. Johannesson and E. Perjons, *An introduction to design science*, vol. 9783319106328. 2014.
- [29] R. Wirth and J. Hipp, "CRISP-DM: towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, 29-39," *Proc. Fourth Int. Conf. Pract. Appl. Knowl. Discov. Data Min.*, no. 24959, pp. 29–39, 2000.