

**AVISO IMPORTANTE** - La memoria no podrá exceder de 20 páginas. Para rellenar correctamente esta memoria, lea detenidamente las instrucciones disponibles en la web de la convocatoria. Es obligatorio rellenarla en inglés si se solicita 100.000 € o más (en costes directos).

**IMPORTANT** – The research proposal cannot exceed 20 pages. Instructions to fill this document are available in the website. If the project cost is equal or greater than 100.000 €, this document must be filled in English.

## 0. DATOS DEL PROYECTO

**IP 1** (Nombre y apellidos): Álvarez Arranz, Álvaro José  
**IP 2** (Nombre y apellidos): Calderón Pazmiño, Miguel Angel  
**IP 3** (Nombre y apellidos): Gallardo Villanueva, Marcos  
**IP 4** (Nombre y apellidos): Martínez del Valle, Alba  
**IP 5** (Nombre y apellidos): Martínez Gijón, Rubén  
**IP 6** (Nombre y apellidos): Wang, Runtian

**TÍTULO DEL PROYECTO (ACRÓNIMO):** *Diagnosti-CAT*: predicción de enfermedades de salud mental mediante aprendizaje profundo (DPESMAP)

**TITLE OF THE PROJECT (ACRONYM):** *Diagnosti-CAT*: mental illness prediction through machine learning (DMIPML)

## 1. ANTECEDENTES Y ESTADO DEL ARTE

En estos últimos años, la salud mental ha emergido como un tema de gran importancia en la sociedad, reconociendo su gran impacto en el bienestar general de las personas. El aumento de la visibilidad de la salud mental en medios tradicionales e internet, el aumento de estos trastornos en la población, incluso las consecuencias de la pandemia por COVID-19 han puesto en primer plano la necesidad de abordar esta temática de manera urgente.

Asimismo, el *machine learning* está revolucionando diversos campos, incluido el de detección de enfermedades de salud mental, al proporcionar herramientas avanzadas basadas en computación. Por lo que, en tareas como el análisis de texto de notas médicas, hasta algunas como el procesamiento de imágenes cerebrales, el uso de aprendizaje automático ha demostrado ser de gran utilidad.

Para analizar el estado del arte de este proyecto se ha seguido la metodología propuesta por las pautas PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), lo que proporciona un marco estructurado y con gran detalle para la realización de revisiones sistemáticas y la presentación de esta información de una manera transparente. Este enfoque es capaz de asegurar rigor y reproducibilidad en la identificación, selección y análisis de la literatura relevante con relación con la detección de enfermedades mentales mediante el uso de *machine learning*.

El proceso de revisión se dividió en varias etapas clave. Estas incluyen la definición de la cuestión que nos ha llevado a esta investigación, la identificación de fuentes de información fiables, la selección de los estudios según una serie de criterios predefinidos en las búsquedas, la extracción de datos e información más relevante y la evaluación de la calidad de la metodología de los estudios incluidos. Además, se hace uso de bases de datos



especializadas y recursos bibliográficos reconocidos para garantizar la rigurosidad de la búsqueda bibliográfica.

### 1.1. Identificación y descripción del dominio

El uso de técnicas avanzadas de *machine learning* para el análisis y diagnóstico de trastornos mentales se ha convertido en un campo de investigación intensivo y prometedor. Este enfoque tecnológico responde a la necesidad creciente de herramientas de diagnóstico precisas y no invasivas que puedan identificar y evaluar trastornos como la depresión, la ansiedad, la esquizofrenia y el trastorno bipolar. Estudios recientes en este campo han explorado diversas metodologías, como la evaluación multimodal que integra audio, vídeo y datos textuales para captar una gama más amplia de indicadores de trastornos mentales, como los estudios que utilizan los conjuntos de datos DAIC-WOZ y AVEC, que han demostrado la viabilidad de aplicar modelos de aprendizaje profundo para el análisis comportamental y emocional.

### 1.2. Identificación de problemas similares

En una revisión más meticulosa de los resúmenes de artículos de relevancia para el proyecto *Diagnosti-CAT*, se ha averiguado que no todos se alinean con los objetivos específicos de nuestro proyecto. Por ejemplo, el estudio *Unsupervised Deep Learning to Detect Agitation From Videos in People With Dementia* introduce técnicas de aprendizaje profundo no supervisado para la detección de agitación en personas con demencia. Aunque es una contribución valiosa, su enfoque en un trastorno específico contrasta con nuestra necesidad de abordar un espectro más amplio de trastornos.

De manera similar, otros estudios como *The Reproducibility of Bio-Acoustic Features is Associated With Sample Duration, Speech Task, and Gender* y *Analyzing Acoustic and Prosodic Fluctuations in Free Speech to Predict Psychosis Onset in High-risk Youths* se exploran meticulosamente las características bioacústicas y prosódicas del habla. Aunque ofrecen detalles significativos sobre las condiciones específicas de los sujetos estudiados, ambos artículos centran sus hallazgos en aspectos que no integran plenamente las dimensiones lingüística, visual y de transcripción de audio que nuestro proyecto persigue combinar en un modelo de *machine learning* unificado.

Tras el análisis de muchos otros artículos, se ha prestado especial atención a una serie de artículos, debido a su relevancia directa y a la profundidad con que abordan problemas similares a los de nuestro proyecto, como estudios que integran evaluaciones multimodales y utilizan aprendizaje profundo para analizar señales lingüísticas, acústicas, visuales e, incluso, de transcripción de audio en el diagnóstico de las patologías objetivo. Es más, también demuestran metodologías avanzadas y resultados prometedores que pueden potenciar el diagnóstico y seguimiento en el contexto de *Diagnosti-CAT*, ofreciendo un marco sólido para la detección temprana y la intervención en salud mental.

### 1.3. Oportunidad científica

El estudio bibliográfico acaba llegando a la conclusión de que existe una oportunidad científica de crear un modelo multimodal de texto y vídeo para la predicción de la patología mental. En concreto, se determina la idoneidad de BERT para el análisis de las características textuales, y *Swin transformers* para las características de vídeo. Estos modelos han demostrado su eficacia en artículos como *Multimodal assessment of schizophrenia symptom severity from linguistic, acoustic and visual cues* y *Vision transformer equipped with neural resizer on facial expression recognition task*. Se construirá a su vez una nueva base de datos para cumplir los requisitos de un modelo semejante.



## 2. JUSTIFICACIÓN Y NOVEDAD DE LA PROPUESTA

### 2.1. Adecuación de la propuesta a las características y finalidad de la modalidad seleccionada.

La propuesta consiste en el desarrollo de un proyecto de investigación a cuatro años. Este proyecto se basa en la aplicación de una herramienta basada en el aprendizaje automático, diseñada para el pre-diagnóstico de cuatro patologías mentales tipo: depresión, ansiedad, trastorno de bipolaridad y esquizofrenia. Se ha establecido *Diagnosti-CAT* como nombre provisional para la herramienta.

Este proyecto incorpora una aproximación interdisciplinar al campo de la salud mental, combinando la psiquiatría clínica con técnicas de aprendizaje automático. Por este mismo motivo, el desarrollo de la herramienta aspira a combinar los conocimientos de profesionales de la salud mental como psicólogos y psiquiatras con los de científicos de datos e investigadores de *deep learning*. A su vez, abre la puerta a multitud de avances en el campo de la ciencia de datos aplicada a la salud mental, abriendo un marco de investigación inicialmente dedicado a las cuatro patologías anteriormente mencionadas, pero ampliable a otras muchas enfermedades no incluidas y menos comunes. En definitiva, el proyecto se propone también como un marco funcional sobre el que trabajar para ayudar a una mejor comprensión de este problema cada vez más visible en la sociedad.

La idea anterior tiene estrecha relación con la modalidad seleccionada, bajo la que se quiere acoger el proyecto. Esta herramienta se trata de un *proyecto de modalidad de investigación orientada*, ya que su objetivo final es resolver ese desafío social oculto en el pasado que es la salud mental. Últimamente está recibiendo la atención que merece, y por ese motivo, cada vez aparecen más personas con esa dolencia de la cual anteriormente eran incapaces de obtener explicación. Los centros sanitarios en la actualidad reciben un volumen creciente de pacientes que buscan consulta por patologías relacionadas con la salud mental. Por ese motivo, muchas veces se ven sobrepasados, y son incapaces de tratar a personas que realmente necesitan un tratamiento psiquiátrico con urgencia.

El fin último de la herramienta *Diagnosti-CAT* es ayudar a la eficacia del sistema de salud, realizando un pre-diagnóstico de los pacientes para averiguar qué clase de enfermedad padecen y la gravedad de la misma.

### 2.2. Justificación y contribución esperada del proyecto a la generación de conocimiento en la temática de la propuesta. Hipótesis de partida.

La hipótesis de partida del proyecto es muy sencilla. Existen numerosos estudios que avalan la capacidad de los sistemas de aprendizaje automático actuales para diagnosticar correctamente las enfermedades mentales de los pacientes siguiendo una serie de inputs. Estos inputs pueden ser marcadores biomédicos, cuestionarios, o incluso entrevistas. Además los datos vienen en multitud de modalidades como el audio, texto, vídeo, entre otras. Sin embargo, no hay mucho trabajo realizado sobre análisis clínico de estas enfermedades. La idea principal en la que se basa el proyecto es que se puede crear una base de datos fundamentada en entrevistas psiquiátricas cerradas a partir de la cual se pueda crear un modelo multimodal de aprendizaje automático que sirva para determinar tanto la enfermedad padecida, como la gravedad. Dentro de este sistema entran las siguientes enfermedades: Depresión, ansiedad, trastorno de bipolaridad y esquizofrenia.

Tomando esa idea como hipótesis base, se espera innovar en dos frentes distintos. En primer lugar, se plantea crear una base de datos novedosa que sirva para detectar indicios de las cuatro patologías objetivo. Esta base de datos se tratará de una entrevista con 20 preguntas prefijadas por expertos de la salud mental. En segundo lugar, se construirá un modelo multimodal nuevo que combina texto y vídeo, gracias al cual se extraerán las características necesarias para ofrecer el diagnóstico final y la gravedad.



### 2.3. Justificación y contribución esperada del proyecto a solucionar problemas concretos vinculados a la prioridad temática seleccionada.

El punto central de este proyecto es su capacidad esperada de mejorar la calidad del sistema de salud nacional. La idea no es diagnosticar al paciente, sustituyendo de algún modo el trabajo de los psiquiatras, sino establecer un pre-diagnóstico de los pacientes de nuevo ingreso. La aplicación de la herramienta está proyectada de la siguiente forma:

1. El nuevo paciente ingresa al centro sanitario por una dolencia perteneciente al campo de la salud mental.
2. Tras unas breves comprobaciones, se le redirige a la herramienta *Diagnosti-CAT*, la cual realizará una serie de preguntas de respuesta libre que el paciente deberá contestar en voz alta. En este proceso el paciente estará solo con la plataforma.
3. Los resultados serán analizados por un psiquiatra encargado, el cual tomará una decisión al respecto, priorizando al paciente según su enfermedad y/o gravedad.

Se espera que este proceso alivie la carga existente sobre los centros de salud, a su vez que permita detectar casos graves, priorizándolos y permitiendo darles la urgencia que necesitan. En definitiva, este proyecto pretende dar prioridad a la salud mental, agilizando y mejorando el contacto del doctor con el paciente, lo cual siempre será el fin último para mejorar la calidad de vida del enfermo.

## 3. OBJETIVOS, METODOLOGÍA Y PLAN DE TRABAJO

### 3.1. Objetivos generales y específicos.

El objetivo primero del proyecto es el desarrollo y validación de un modelo de aprendizaje automático que sea capaz de predecir la probabilidad y gravedad de enfermedades mentales como la ansiedad, la depresión, el trastorno bipolar y la esquizofrenia. Para ello, se usará una serie de respuestas estructuradas que serán proporcionadas por los pacientes en un cuestionario prefijado y las imágenes de los mismos obtenidas mediante grabaciones. Con ello se busca ofrecer una herramienta no invasiva, accesible y precisa que pueda ser utilizada por profesionales de la salud mental con el fin de detectar estas dolencias junto a su gravedad antes de tener un primer acercamiento; para así, optimizar el proceso de detección y permitir tratar a tiempo los problemas más graves.

Los objetivos del proyecto son los siguientes:

- **Objetivo Específico 1 (OE1):** Creación de una base de datos multimodal (texto y vídeo) estructurada de pacientes adultos con depresión, ansiedad, trastorno bipolar, esquizofrenia y sujetos sanos.
- **Objetivo Específico 2 (OE2):** Desarrollo de la plataforma *Diagnosti-CAT* mediante la utilización de algoritmos de machine learning como BERT y Swin transformers para la identificación y cuantificación de la gravedad.
- **Objetivo Específico 3 (OE3):** Obtención de métricas de rendimiento, como la precisión, sensibilidad, especificidad, entre otras, para la evaluación y comparación de nuestro modelo con métodos actuales.

### 3.2. Descripción de la metodología.

Para garantizar el éxito del proyecto *Diagnosti-CAT*, se adoptará una metodología metódica y detallada que nos permitirá abordar cada aspecto del desarrollo de una herramienta de diagnóstico de trastornos mentales. El proyecto se estructura en cuatro fases críticas, cada una diseñada para construir sobre el progreso de la anterior y optimizar los resultados del proyecto en su conjunto.



En la fase 1, nos centraremos en el diseño del estudio y la recolección de datos, asegurando que todas las operaciones se realicen dentro de los marcos éticos apropiados y que los datos recopilados sean de alta calidad y relevantes para nuestros objetivos. Durante la fase 2, el enfoque se desplazará hacia el desarrollo y la integración de los modelos de aprendizaje automático, utilizando avanzadas técnicas de procesamiento de texto y análisis de video para crear un sistema robusto y eficiente capaz de realizar diagnósticos precisos y evaluaciones de la gravedad de los trastornos mentales. La fase 3 se dedicará al desarrollo de la aplicación web, implementando una interfaz intuitiva y pruebas exhaustivas para garantizar que la aplicación sea segura, eficaz y fácil de integrar en los flujos de trabajo clínicos existentes. Finalmente, la fase 4 se enfocará en la validación clínica y la adopción de la herramienta en hospitales, asegurando su efectividad en entornos reales y su capacidad para desahogar eficientemente los servicios de atención psicológica y psiquiátrica. Esta fase incluirá pruebas piloto, evaluación de impacto y ajustes basados en el *feedback* de los expertos, con el objetivo de integrar plenamente la herramienta en los procedimientos hospitalarios estándar.

### 3.2.1 Fase 1: Establecimiento de Fundamentos y Recolección de Datos

La primera fase del proyecto es crítica, ya que establece la base sobre la cual se construirán todas las actividades subsiguientes. Esta fase está dividida en varias actividades clave que aseguran unos fundamentos sólidos para el desarrollo y validación del modelo.

#### Diseño del Estudio y Aprobaciones Éticas

Antes de comenzar cualquier recolección de datos, es esencial obtener las aprobaciones éticas necesarias. Esta tarea implica colaborar con los comités de ética de investigación para garantizar que todos los procedimientos cumplen con las normativas internacionales y nacionales pertinentes. El diseño del estudio deberá incluir protocolos detallados para el manejo de información sensible, garantizando la protección y privacidad de los participantes.

#### Formación y Organización del Equipo de Campo

Un equipo de campo bien preparado es esencial para la recolección de datos de alta calidad. Este equipo incluirá investigadores y técnicos capacitados en la interacción con pacientes bajo condiciones éticas y empáticas. La formación cubrirá aspectos técnicos de la recolección de datos y habilidades interpersonales para manejar situaciones sensibles con los participantes.

#### Recolección y Procesamiento de Datos

La recolección de datos se realizará en centros médicos colaboradores, utilizando equipos de grabación de alta definición para capturar datos audiovisuales y equipos de grabación de audio para las transcripciones. Las sesiones de recolección serán diseñadas para captar tanto las respuestas verbales como no verbales de los pacientes, para el análisis multimodal posterior.

- *Datos Audiovisuales:* Incluyen video y audio que serán analizados para obtener indicadores visuales de condiciones de salud mental. El audio será transcrito para su posterior procesamiento.
- *Transcripciones Textuales:* Las sesiones grabadas serán transcritas cuidadosamente, preservando características lingüísticas y paralingüísticas que pueden ser indicativas de trastornos mentales además de las etiquetas de los profesionales

#### Revisión de la calidad de los datos

Después de la recolección, todos los datos serán revisados para asegurar su calidad antes de su uso en la modelación. Esto incluye la verificación de la claridad del audio y del video, la precisión de las transcripciones y la integridad de los datos recogidos. Se emplearán técnicas de limpieza de datos para corregir cualquier inconsistencia o error encontrado.



## **Establecimiento del Estado del Arte y Revisión Literaria**

Paralelamente a la recolección de datos, se llevará a cabo una revisión exhaustiva de la literatura existente. Este estudio ayudará a identificar brechas en la investigación actual y a confirmar la relevancia y la innovación del enfoque propuesto. Utilizando bases de datos académicas y recursos bibliográficos, se hará un análisis crítico de trabajos previos, especialmente aquellos que utilizan metodologías multimodales y técnicas de aprendizaje automático en el diagnóstico de trastornos mentales.

## **Desarrollo Preliminar del Modelo**

Aunque la fase intensiva de desarrollo del modelo se programará para fases posteriores, un esbozo preliminar y la configuración inicial del modelo comenzarán en esta fase. Esto incluye la selección de algoritmos adecuados, la planificación de la arquitectura del modelo y pruebas iniciales con conjuntos de datos de prueba para validar las técnicas seleccionadas.

## **Resultados esperados de la Fase 1**

Al final de esta fase, se espera tener un conjunto robusto y éticamente recogido de datos multimodales, un equipo bien preparado y protocolos establecidos, y un entendimiento claro del estado del arte. Esto proporcionará una base sólida sobre la cual se construirán las fases subsiguientes del proyecto, centradas en el desarrollo intensivo del modelo y su posterior implementación en una aplicación de diagnóstico.

### **3.2.2 Fase 2: Desarrollo y Ajuste del Modelo**

La segunda fase del proyecto se centra en el desarrollo técnico del modelo de aprendizaje automático multimodal. Esta etapa es crucial, ya que se diseñará, implementará y perfeccionará el modelo que se utilizará para analizar y predecir condiciones de salud mental basándose en los datos recogidos en la fase anterior.

## **Desarrollo de Modelos de Aprendizaje Automático**

Basándonos en el estado del arte y las revisiones sistemáticas realizadas, seleccionaremos y adaptaremos tecnologías de procesamiento de lenguaje natural y análisis de imagen para desarrollar nuestro modelo:

- *Procesamiento de Lenguaje Natural (PLN)*: Utilizaremos BERT (*Bidirectional Encoder Representations from Transformers*) por su capacidad de comprender contextos complejos dentro de textos largos, lo que es ideal para analizar transcripciones clínicas detalladas.
- *Análisis de Imagen y Video*: Adoptaremos Swin Transformers, que se han demostrado eficaces para analizar secuencias de video debido a su enfoque en ventanas deslizantes y atención jerárquica, permitiendo una interpretación más fina de las expresiones faciales y los gestos corporales.

## **Integración Multimodal**

Uno de los desafíos más significativos será la integración efectiva de datos de texto y video para crear un modelo que evalúe simultáneamente múltiples tipos de información. Se desarrollará un marco que pueda fusionar eficazmente estas dos fuentes de datos, manteniendo la coherencia y la complementariedad entre las modalidades. Esto implicará técnicas de fusión de características y alineación de tiempo entre los datos de audio y video.

## **Ajuste Fino y Optimización de Modelos**

Después de la integración inicial, el modelo pasará por varias iteraciones de ajuste fino y optimización. Este proceso será esencial para adaptar los modelos preentrenados a nuestro conjunto de datos específico y nuestras necesidades de diagnóstico:



- **Validación Cruzada:** Utilizaremos la técnica de validación cruzada k-fold para evaluar la generalización del modelo sobre los datos no vistos durante el entrenamiento. Esta técnica divide el conjunto de datos en 'k' partes iguales, entrenando el modelo en 'k-1' de estas partes y validándolo en la parte restante. Este proceso se repite 'k' veces, con cada subconjunto usado exactamente una vez como conjunto de validación. La validación cruzada es esencial para evitar el sobreajuste y para asegurar que nuestro modelo funcione bien en general, no solo en el conjunto de datos con el que fue entrenado.
- **Ajuste de Hiperparámetros:** El ajuste de hiperparámetros es un paso crítico en la optimización de cualquier modelo de aprendizaje automático. Utilizaremos técnicas como la búsqueda en cuadrícula (Grid Search) y la búsqueda aleatoria (*Random Search*) para explorar sistemáticamente una amplia gama de valores para los hiperparámetros del modelo. Los hiperparámetros que ajustaremos incluirán:
  - **Tasa de Aprendizaje:** Fundamental para controlar la rapidez con la que el modelo aprende; tasas de aprendizaje muy altas pueden hacer que el aprendizaje salte sobre mínimos globales, mientras que tasas muy bajas pueden hacer que el proceso sea demasiado lento y susceptible al sobreajuste.
  - **Número de Capas en la Red Neuronal:** Ajustar la profundidad de la red puede tener un impacto significativo en la capacidad del modelo para aprender representaciones complejas de los datos.
  - **Configuración de la Atención:** Específicamente para modelos como BERT y los transformadores, ajustar cómo las capas de atención ponderan diferentes partes de los datos de entrada puede mejorar significativamente la relevancia de las características aprendidas para las tareas de diagnóstico.

### Evaluación Continua:

Durante el proceso de ajuste fino y optimización, evaluaremos continuamente el rendimiento del modelo utilizando un conjunto de validación independiente. Esto nos permitirá monitorear el progreso y hacer ajustes iterativos. Las métricas clave que utilizaremos incluyen la precisión, la sensibilidad, la especificidad y la puntuación F1, que nos proporcionarán una comprensión holística de la eficacia del modelo en la clasificación correcta de las condiciones diagnósticas.

- **Implementación de Herramientas y Automatización:** Para facilitar el ajuste fino y la optimización, implementaremos herramientas de automatización que nos permitirán ejecutar múltiples configuraciones de forma eficiente. Utilizaremos plataformas como *TensorFlow Extended (TFX)* y *PyTorch Lightning*, que proporcionan soporte para tuberías de ML escalables y robustas, gestionando desde la preparación de datos hasta la evaluación del modelo y el despliegue.

### Pruebas Preliminares del Modelo

Para evaluar la precisión, sensibilidad y especificidad del modelo, implementaremos un conjunto de pruebas estadísticas y técnicas de validación reconocidas:

- **Validación Cruzada:** Utilizaremos la técnica de validación cruzada k-fold para evaluar cómo el modelo generaliza a un conjunto de datos independiente que no se utilizó durante el entrenamiento. Esta técnica divide los datos en k subconjuntos y realiza el entrenamiento en k-1 de estos, utilizando el subconjunto restante para la prueba. Este proceso se repite k veces con cada subconjunto utilizado una vez como datos de prueba.
- **Curva ROC (Receiver Operating Characteristic):** Generaremos curvas ROC para cada clase diagnóstica para evaluar la capacidad discriminativa del modelo en diferentes puntos de corte de decisión. El área bajo la curva ROC (AUC) será una métrica clave para medir la calidad del modelo.



- **Matriz de Confusión:** Para cada prueba, construiremos matrices de confusión que nos permitirán visualizar la precisión de clasificación del modelo, incluyendo verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos. Esto nos ayudará a calcular métricas específicas como la precisión, el recall y el F1-score.

### Evaluación de la Robustez del Modelo

Dada la variabilidad en cómo se pueden manifestar los trastornos mentales entre diferentes individuos, es crucial que nuestro modelo maneje esta diversidad de manera efectiva:

- **Aumento de Datos:** Implementaremos técnicas de aumento de datos para enriquecer nuestro conjunto de entrenamiento y aumentar la diversidad de los datos sin necesidad de recopilación adicional. Esto incluirá técnicas como la modificación de las características del audio (cambios en el tono, velocidad, etc.), alteraciones en las imágenes (rotaciones, cambio de escala, etc.) y transformaciones textuales.
- **Pruebas de Sensibilidad:** Realizaremos pruebas de sensibilidad para evaluar cómo pequeñas variaciones en los datos de entrada afectan los resultados del modelo. Esto es crucial para entender la estabilidad del modelo frente a variaciones menores que podrían ocurrir en un entorno clínico real.
- **Validación en Subconjuntos de Datos Diversos:** Además de la validación cruzada general, realizaremos pruebas específicas en diversos subconjuntos de datos que reflejen diferentes demografías, tipos de trastornos y severidades de síntomas. Esto nos ayudará a asegurar que el modelo es efectivo y justo a través de diferentes grupos de pacientes.

### Documentación y Análisis de Resultados

Se documentará exhaustivamente todos los procesos y resultados obtenidos durante esta fase. Esto incluirá la descripción de la arquitectura del modelo, los procesos de entrenamiento, las decisiones de diseño, y los resultados de las pruebas. Este documento servirá como un recurso vital para las fases posteriores del proyecto y para futuras investigaciones.

### Resultados esperados de la Fase 2

Al finalizar esta fase, esperamos tener un modelo de aprendizaje automático bien desarrollado y optimizado, capaz de analizar y predecir con precisión las condiciones de salud mental a partir de datos multimodales. Este modelo será la piedra angular de la herramienta de diagnóstico que desarrollaremos en las fases subsiguientes, y su efectividad será crucial para el éxito del proyecto en su conjunto.

### 3.3.3 Fase 3: Desarrollo Exhaustivo y Pruebas Detalladas de la Aplicación Web

#### Planificación y Arquitectura del Sistema

El diseño arquitectónico de la aplicación es fundamental para soportar no solo la carga operativa actual, sino también para facilitar futuras expansiones. Optamos por una arquitectura de microservicios, que ofrece numerosas ventajas en términos de escalabilidad y mantenimiento. Cada servicio funcionará de manera independiente, lo que permite actualizaciones y pruebas sin afectar al resto del sistema.

- Para el *backend*, seleccionamos Node.js con Express.js por su capacidad para manejar numerosas conexiones simultáneas sin bloqueo, gracias a su modelo de operaciones de entrada/salida no bloqueante. Esta elección es esencial en un contexto clínico donde los tiempos de respuesta rápidos y el manejo eficiente de las solicitudes son críticos.





- En el *frontend*, React.js es la elección preferida por su enfoque en componentes reactivos y su vasto ecosistema, que facilita la creación de interfaces de usuario ricas y responsivas que pueden adaptarse a los complejos flujos de trabajo de diagnóstico médico.

La integración de modelos de aprendizaje automático como *BERT* y *Swin Transformers* requiere una atención especializada. Utilizaremos contenedores Docker para encapsular cada componente del modelo, garantizando así la consistencia del entorno de ejecución y simplificando el despliegue y la escalabilidad en diferentes infraestructuras, desde servidores locales hasta la nube.

### **Desarrollo y Pruebas de Funcionalidades**

El desarrollo se organizará en sprints de dos semanas, siguiendo principios ágiles para adaptarnos rápidamente a cualquier cambio o nuevo requisito que surja durante el proceso de desarrollo. Cada sprint incluirá etapas de planificación, desarrollo, pruebas y revisión, culminando con una demostración del progreso a los stakeholders.

Las APIs RESTful serán diseñadas para garantizar una comunicación eficiente y segura entre el frontend y el backend. Implementaremos autenticación robusta y mecanismos de autorización para proteger la privacidad y la integridad de los datos médicos sensibles.

Las pruebas unitarias y de integración serán fundamentales para asegurar la calidad del código y la funcionalidad del sistema. Utilizaremos Jest para las pruebas del frontend y backend, asegurando que cada componente funcione correctamente antes de integrarse en el sistema mayor.

### **Pruebas de Carga y Estrés**

Las pruebas de carga son cruciales para garantizar que la aplicación pueda manejar un alto volumen de usuarios simultáneos, especialmente durante picos de uso, sin degradación del rendimiento. Herramientas como JMeter simularán diferentes escenarios de carga para identificar cuellos de botella y puntos de falla en la aplicación.

Las pruebas de seguridad también recibirán alta prioridad. Además de las pruebas de penetración estándar, implementaremos pruebas específicas para evaluar la vulnerabilidad de la aplicación frente a ataques comunes, como inyección SQL, XSS y CSRF. Utilizaremos herramientas como OWASP ZAP para identificar y mitigar posibles vulnerabilidades.

### **Despliegue y Operaciones**

El despliegue continuo será facilitado por un pipeline de CI/CD configurado con Jenkins o GitHub Actions, permitiendo despliegues automáticos y seguros de nuevas versiones de la aplicación. Esto es esencial para mantener la aplicación actualizada con las últimas mejoras y correcciones de seguridad sin interrumpir el servicio a los usuarios.

El monitoreo en tiempo real del sistema se realizará mediante Prometheus, y los logs serán gestionados usando ELK Stack, lo cual permite un análisis detallado y en tiempo real del estado del sistema. Esto es crucial para la detección y resolución rápida de cualquier problema operativo o de rendimiento que pueda surgir.

### **Resultados esperados de la Fase 3**

Esta fase del proyecto es vital para transformar los modelos teóricos y los algoritmos de aprendizaje automático en una herramienta práctica y operativa que pueda ser utilizada de manera efectiva en entornos clínicos. Al final de esta fase, la aplicación no solo cumplirá con los requisitos técnicos y funcionales, sino que también será segura, eficiente y fácil de usar para los profesionales de la salud, lo que asegurará su adopción y eficacia en el diagnóstico y evaluación de trastornos mentales.



### 3.3.4 Fase 4: Validación Clínica y Adopción en Hospitales

#### Implementación en Hospitales y Evaluación Clínica

La cuarta fase del proyecto *Diagnosti-CAT* se centra en la validación clínica y la adopción en entornos hospitalarios para desahogar los servicios de atención psicológica y psiquiátrica. En esta fase, la aplicación desarrollada será sometida a rigurosas pruebas clínicas para validar su eficacia y seguridad, con el objetivo de proporcionar una herramienta confiable que priorice a los pacientes que requieran atención urgente basándose en la gravedad de sus síntomas psicológicos.

- **Pruebas Clínicas Rigurosas:** Inicialmente, la aplicación se implementará en un conjunto controlado de hospitales colaboradores. Estas pruebas involucrarán a pacientes que voluntariamente participen en sesiones donde se utilizarán tanto las evaluaciones estándar actuales como nuestra herramienta, para comparar y validar la precisión y confiabilidad de los diagnósticos y evaluaciones de gravedad proporcionados por *Diagnosti-CAT*.
- **Capacitación del Personal Médico:** Se organizarán sesiones de capacitación para el personal médico. Estas capacitaciones cubrirán el uso adecuado de la herramienta, interpretación de los resultados, y cómo integrar esta nueva tecnología en sus flujos de trabajo diarios sin interrumpir las operaciones existentes.
- **Integración con Sistemas Hospitalarios:** Trabajaremos en estrecha colaboración con los equipos de TI de los hospitales para integrar nuestra aplicación con los sistemas de registros médicos electrónicos (EMR) existentes, asegurando que el proceso sea fluido y que los datos generados por nuestra herramienta sean accesibles y útiles dentro del contexto más amplio del historial médico del paciente.

#### Monitoreo y Ajustes Continuos

- **Recopilación de Feedback:** El feedback de los médicos, enfermeras, y administradores hospitalarios será esencial. Este feedback nos permitirá ajustar la aplicación para mejor adaptarse a las necesidades reales del entorno clínico y mejorar la interfaz de usuario para facilitar su uso.
- **Evaluación Continua del Impacto:** Además de la funcionalidad y precisión, evaluaremos cómo la herramienta afecta la eficiencia del flujo de trabajo clínico y la rapidez con la que los pacientes reciben atención. Nuestro objetivo es asegurar que la herramienta ayude efectivamente a desahogar los servicios de atención psicológica y psiquiátrica en los hospitales.
- **Análisis de Datos a Gran Escala:** Utilizaremos los datos recopilados durante las pruebas clínicas para realizar un análisis a gran escala de la efectividad de la herramienta. Este análisis nos ayudará a identificar patrones que podrían no ser evidentes durante las evaluaciones individuales y a ajustar nuestros algoritmos para mejorar aún más la precisión y la utilidad de las evaluaciones.

#### Escalabilidad y Mantenimiento de la Aplicación

##### Mantenimiento Continuo

El mantenimiento continuo de la aplicación es muy importante para asegurar su eficacia y fiabilidad a largo plazo. Incluirá la actualización regular de software, la corrección de errores y la mejora de funcionalidades existentes. Esto garantiza que la aplicación permanezca al día con los últimos avances tecnológicos y cumpla con las cambiantes regulaciones de seguridad y privacidad de datos.

- **Actualizaciones de Software:** Implementaremos un calendario regular de actualizaciones para introducir mejoras en la funcionalidad y en la interfaz de usuario, así como para abordar cualquier vulnerabilidad de seguridad que pueda surgir. Esto



incluirá también la actualización de las librerías y frameworks subyacentes para aprovechar las mejoras en rendimiento y seguridad que ofrecen las nuevas versiones.

- **Corrección de Errores:** A pesar de los rigurosos procesos de prueba, es inevitable que aparezcan errores en el software. Mantendremos un sistema proactivo de seguimiento y corrección de errores para minimizar el impacto en los usuarios finales.
- **Soporte Técnico:** Proporcionaremos soporte técnico continuo para resolver cualquier problema que los usuarios puedan enfrentar. Esto incluirá soporte en línea, recursos de autoayuda y la posibilidad de contacto directo con nuestro equipo técnico para soluciones rápidas.

### Escalabilidad de la Plataforma

Asegurar la escalabilidad de la plataforma es esencial para manejar un aumento en la carga de trabajo sin comprometer el rendimiento. Prepararemos la aplicación para escalar tanto vertical como horizontalmente para responder a un incremento en la demanda.

- **Escalabilidad Vertical:** Incrementaremos la capacidad de nuestros servidores mediante la mejora de CPU, RAM y almacenamiento según sea necesario. Esto permitirá que la aplicación maneje un mayor volumen de procesamiento y almacenamiento de datos de manera más eficiente.
- **Escalabilidad Horizontal:** Implementaremos estrategias de escalabilidad horizontal, como el balanceo de carga y la replicación de bases de datos. Esto no solo mejora el rendimiento al distribuir la carga entre múltiples servidores, sino que también aumenta la redundancia y la disponibilidad de la aplicación.
- **Monitorización y Análisis de Rendimiento:** Utilizaremos herramientas avanzadas de monitorización para observar continuamente el rendimiento de la aplicación. Esto nos permitirá identificar y abordar proactivamente cualquier problema de rendimiento y escalar recursos antes de que afecten a los usuarios.

### Resultados esperados de la Fase 4

En esta fase final esperamos asegurar que *Diagnosti-CAT* no solo sea una herramienta tecnológicamente avanzada, sino también una solución práctica y efectiva en el ambiente clínico real. Al final de esta fase, esperamos que *Diagnosti-CAT* esté plenamente integrado en los procesos hospitalarios, proporcionando una herramienta valiosa para priorizar la atención médica y mejorar la eficiencia del tratamiento de trastornos mentales, además de seguir monitorizando esta herramienta para mantener su escalabilidad y funcionalidad.

#### 3.3. Plan de trabajo y cronograma.

Tabla 1. Tabla de plan de trabajo.

Plan de Trabajo					
Fase	Año(s)	Mes(es)	Actividad	Responsable	Descripción Detallada
1	1	1-6	Diseño del estudio y aprobaciones éticas	Equipo de gestión	Obtención de aprobaciones éticas, diseño del protocolo de investigación.



1	1	5-12	Firma de acuerdos con hospitales, recolección de datos y revisión RGPD	Equipo de campo y Equipo de datos	Obtención datos, revisión de ley y anonimización
2	2	1-6	Etiquetado de datos por expertos y Procesamiento de los mismos	Equipo de datos	Etiquetado de datos con ayuda de expertos, procesamiento preliminar y verificación de calidad de los datos recogidos.
2	2	1-6	Desarrollo inicial del modelo	Equipo de ML	Desarrollo y ajuste preliminar de modelos de aprendizaje automático.
2	2	7-12	Integración y optimización de modelos	Equipo de ML	Integración de modelos de texto y video, optimización y validación cruzada.
3	3	1-6	Desarrollo de la aplicación y pruebas de software	Equipo de desarrollo	Implementación de la interfaz de usuario, backend y pruebas funcionales.
3	3	7-12	Pruebas de carga y seguridad	Equipo de desarrollo	Pruebas de carga, rendimiento y seguridad para garantizar la estabilidad y seguridad.
4	4	1-4	Implementación piloto y capacitación	Equipo de desarrollo	Pruebas piloto en hospitales seleccionados, capacitación de personal médico.
4	4	5-9	Evaluación de impacto y ajustes	Equipo de evaluación	Recopilación de feedback, evaluación de impacto, ajustes basados en datos de uso.
4	4	10-12	Escalabilidad y mantenimiento de la aplicación	Equipo de gestión	Preparación y soporte de la aplicación, asegurando la escalabilidad y funcionamiento.

### 3.4. Identificación de puntos críticos y plan de contingencia.

Tabla 2. Plan de contingencia.

Objetivo específico	Puntos críticos	Plan de contingencia
		Establecer nuevas redes de



OE1	No recopilar suficientes datos en el plazo de tiempo establecido.	contactos con centros sanitarios o búsqueda de bases de datos públicas que se aproximen a las condiciones del estudio.
OE2	Ausencia de generalización de los algoritmos de machine learning que no permitan una óptima clasificación de los pacientes.	Aplicación de técnicas de aumento de datos como la modificación de las características del audio, alteraciones en las imágenes y transformaciones textuales con el objetivo de mejorar el entrenamiento y los resultados.

### 3.5. Resultados previos del equipo en la temática de la propuesta.

El equipo de investigación propulsor del proyecto consta de amplia experiencia en el campo del aprendizaje automático. Todos los integrantes han demostrado con anterioridad su interés del mismo al campo de la salud mental, y han realizado una búsqueda bibliográfica exhaustiva, lo que les ha brindado un gran conocimiento del tema. El equipo al completo avala la viabilidad de la propuesta.

### 3.6. Recursos humanos, materiales y de equipamiento disponibles para la ejecución del proyecto.

A continuación se mencionan los recursos disponibles para el proyecto:

- *Recursos humanos:* Se contratará a un grupo de trabajo de expertos de la salud mental, concretamente psiquiatras colegiados. Este grupo se dedicará tanto a la confección del cuestionario, como a la realización de las entrevistas y el etiquetado. Se espera colaborar con múltiples hospitales públicos mediante convenios.
- *Recursos materiales y equipamiento:* Se tendrá acceso a los recursos informáticos de la Escuela Politécnica Superior (EPS) para la creación del modelo y su entrenamiento. A su vez se utilizará equipamiento de grabación de audio y vídeo de calidad para la creación de la base de datos.

Todo el proyecto se llevará a cabo de manera híbrida entre la EPS y los distintos hospitales con los que se han establecido convenios.

## 4. IMPACTO ESPERADO DE LOS RESULTADOS

### 4.1. Impacto esperado en la generación de conocimiento científico-técnico en el ámbito temático de la propuesta.

Como consecuencia de la tendencia creciente de problemas de salud mental, se prevé que *Diagnosti-CAT* tendrá una amplia aceptación y un gran impacto dentro de la comunidad científica. Debido a la creación de una novedosa herramienta de análisis multimodal para cuantificar la probabilidad como la gravedad de la depresión, la ansiedad, el trastorno de bipolaridad o la esquizofrenia. Además, de



considerar la ausencia de propuestas similares, esta herramienta mejora la accesibilidad a un tratamiento médico.

No obstante, con el fin de conseguir una adecuada divulgación se asistirá a congresos así como se publicarán artículos en revistas científicas enfocadas en el análisis de video y texto dentro del ámbito de la salud mental. De esta forma, se pretenderá potenciar el intercambio de resultados obtenidos al implementar nuestra metodología, al igual que, enriquecernos de conocimientos y datos aportados por otros investigadores dentro de la comunidad científica.

#### 4.2. Impacto social y económico de los resultados previstos.

Según los datos recogidos por la Organización Mundial de la Salud (OMS) uno de cada ocho personas padece algún problema de salud mental, entre ellos destaca la ansiedad, la depresión, la esquizofrenia o el trastorno de bipolaridad. No obstante, la principal preocupación de la salud mental deriva principalmente en los años perdidos por discapacidad, los costes sanitarios o la estigmatización social. Además de enfatizar la baja financiación de los gobiernos en la atención sanitaria enfocado en la salud mental, en concreto menos del 2%. Igualmente, esto contribuye a la ausencia de especialistas encargados del diagnóstico de estas patologías, de acuerdo con las cifras de la OMS estima 1 psiquiatra por cada 200.000 habitantes.

Por estas razones en este proyecto se ha diseñado *Diagnosti-CAT*, una herramienta que identifica la depresión, la ansiedad la esquizofrenia y el trastorno de bipolaridad mediante la realización de un cuestionario en un centro médico. *Diagnosti-CAT* es una herramienta que promueve la accesibilidad al diagnóstico médico precoz, al igual que, permitir realizar dicha evaluación en zonas médicas no especializadas como puede ser regiones rurales, entornos remotos, entre otros. Asimismo, esta aplicación incrementará la calidad de vida de los pacientes debido a que las patologías mentales son un factor de discapacidad tanto social como a nivel productivo.

Con el fin de demostrar los beneficios de *Diagnosti-CAT* se implementará en zonas rurales dentro del ámbito nacional con el objetivo de afirmar que *Diagnosti-CAT* permite mejorar y dar prioridad aquellos pacientes con un problema de salud mental de más severidad.

#### 4.3. Plan de comunicación científica e internacionalización de los resultados (indicar la previsión de publicaciones en acceso abierto).

Con el propósito de conseguir una adecuada difusión de *Diagnosti-CAT*, se publicarán tres artículos acerca de la predicción como la gravedad de las cuatro patologías bajo estudio. En concreto se divulgará en revistas como Journal of Medical Internet Research Mental Health, IEEE Transactions on Affective Computing, ACM Transactions on Interactive Intelligent Systems (TiiS), entre otras.

Del mismo modo, se asistirá a congresos nacionales como internacionales para alcanzar un mayor impacto y propagación de nuestra metodología. Entre ellos cabe destacar, el Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB), Conferencia Internacional sobre Computación Afectiva e Interacción Inteligente (ACII) o Conferencia internacional del IEEE sobre informática sanitaria (ICHI).

**Comentado [1]:** chrome-extension://efaidnbmnnnibpcajpcgclefindmkaj/https://iris.who.int/bitstream/handle/10665/356118/9789240051966-spa.pdf?sequence=1

**Comentado [2]:** [https://mental.jmir.org/?\\_hstc=236487962.dd91613c666bfcc3827f6ea6b0ab505.1715898644427.1715898644427.1715898644427.1&\\_hssc=236487962.1.1715898644428&\\_hsfp=2198788217](https://mental.jmir.org/?_hstc=236487962.dd91613c666bfcc3827f6ea6b0ab505.1715898644427.1715898644427.1715898644427.1&_hssc=236487962.1.1715898644428&_hsfp=2198788217)

**Comentado [3]:** <https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=5165369>

**Comentado [4]:** <https://dl.acm.org/journal/tiis>

**Comentado [5]:** <https://seib.org.es/caseib/>

**Comentado [6]:** <https://acii-conf.net/>

**Comentado [7]:** <https://ieeexplore.ieee.org/xpl/conhome/1803080/all-proceedings>



#### 4.4. Plan de divulgación de los resultados a los colectivos más relevantes para la temática del proyecto y a la sociedad en general. *En los casos en que sea de aplicación:*

Asimismo, con el fin de conseguir un adecuado impacto de la utilidad de la herramienta diseñada para la identificación de las cuatro patologías bajo análisis se procederá a su difusión mediante conferencias, talleres, boletines informativos o redes sociales en los centros sanitarios colaboradores, al igual que, instituciones educativas o sociales.

De este modo, se pretenderá reducir la estigmatización social que puede ocasionar padecer un problema de salud mental, permitiendo un tratamiento más precoz.

#### 4.5. Plan de transferencia y valorización de los resultados.

Debido a la necesidad clínica y el gran potencial e innovadora idea diseñada, *Diagnosti-CAT* se encontrará protegida bajo una patente en España. Con el fin de conseguir la transferencia se establecerán relaciones con centros sanitarios privados como públicos, mutuas o especialistas de salud mental que promoverán la implementación de *Diagnosti-CAT*. Asimismo, se dispondrá la opción de leasing con los centros suscritos, facilitando de esta forma una mayor accesibilidad ante actualizaciones y mejoras de *Diagnosti-CAT*.

#### 4.6. Resumen del plan de gestión de datos previsto.

Un elemento clave de *Diagnosti-CAT* es la obtención, etiquetado y tratamiento de los datos a utilizar. Con el objetivo de crear una base de datos representativa y equilibrada, se han establecido convenios con varios hospitales nacionales. Estas colaboraciones permitirán conseguir grabaciones clínicas con un esquema fijo mediante el empleo de *Diagnosti-CAT*. Esta plataforma está compuesta por 20 preguntas fijas redactadas previamente por un conjunto de expertos que permitirán al usuario responder de manera libre.

Como criterio de selección de los voluntarios se ha seleccionado un rango de edad de 18 a 60 años. Además, de incluir únicamente a sujetos sin antecedentes de patologías mentales previas o pacientes con diagnóstico en depresión, ansiedad, trastorno de bipolaridad o esquizofrenia.

La grabación consistirá en el registro de la parte superior del cuerpo del entrevistado con la finalidad de detectar expresiones faciales, al igual que, gestos realizados por el paciente durante el procedimiento y en la obtención del contenido de audio de las respuestas respondidas. No obstante, el audio será transformado a texto mediante la utilización de la aplicación de Google Cloud Platform. El objetivo de las grabaciones será extraer tanto el contenido visual como textual que constituirán la entrada a nuestro modelo. Mediante ambos componentes se estimará la probabilidad así como la gravedad de presentar alguna de las cuatro patologías bajo estudio: depresión, ansiedad, trastorno de bipolaridad o esquizofrenia. De esta manera, *Diagnosti-CAT* proporcionará una realimentación al sanitario acerca de la salud mental del usuario sometido a esta prueba.

Todos los datos obtenidos así como resultados extraídos por esta herramienta serán almacenados en un repositorio bajo la regulación del Reglamento General de Protección de los Datos (RGPD) y la ley Orgánica de Protección de Datos y Garantía de Derechos Digitales (LOPDGDD). Sin embargo, antes de la adquisición de estas grabaciones se deberá disponer el consentimiento informado de cada uno de los pacientes participantes al estudio, al igual que la aprobación del comité ético de los hospitales. También se procederá a la anonimización de los datos para la protección de los datos personales de los usuarios participantes.



#### 4.7. Efectos de la inclusión de género en el contenido de la propuesta.

Como se comentó anteriormente, los criterios de inclusión para la participación en este estudio serán el rango de edad y la presencia de la patología de interés o su ausencia. Por este motivo, este proyecto no se centrará en el impacto en la dimensión del género en los resultados alcanzados.

### 5. JUSTIFICACIÓN DEL PRESUPUESTO SOLICITADO

A la hora de desglosar los costes se ha realizado una división en base a la tipología del mismo, por lo que estos pueden ser costes de personal, costes de equipamiento o costes asociados a divulgación y difusión.

Comenzando con los costes de personal, para este proyecto se necesitará contratar a cuatro psiquiatras titulados para realizar el proceso de etiquetado de los datos tanto de imagen como de texto. Esto supondrá un coste de 40.000€ en un periodo de cinco meses y se calcula como: 16 horas/mes \* 125€/hora (precio de mercado) = 2.000€/mes; 2.000€/mes cada psiquiatra \* 5 meses de trabajo = 10.000€/periodo \* 4 psiquiatras = 40.000€.

En relación a los costes de equipamiento, estos hacen referencia a los costes de los activos necesarios para realizar el proyecto, entre los que encontramos: Equipos de recolección de datos, 5 máquinas capaces de capturar el audio y la imagen de los pacientes en buena calidad con un coste unitario de 500€; Equipos informáticos, 3 ordenadores con la suficiente potencia para el procesamiento de datos y elaboración del modelo con un precio unitario de 1.800€; Monitores, 3 con un coste unitario de 150€; Servidor, capaz de almacenar de manera segura datos de alta sensibilidad con un coste de 200€/mes durante todo el periodo 48 meses; Patente, con el fin de garantizar la propiedad intelectual de nuestra herramienta en la Oficina Española de Patentes y Marcas (OEPM) con un coste de 1200€ por el registro más 20€ anuales de tasas.

Por último, los costes asociados a la divulgación y difusión de nuestro proyecto. En este apartado podemos subdividir en: Campañas de captación para hospitales, con el objetivo de encontrar entidades colaboradoras, se realizarán 4 con un coste de 500€ cada una; Talleres con el objetivo de convencer a psiquiatras y psicólogos en los beneficios del uso de nuestra herramienta, se llevarán a cabo 5 de ellos con un coste de 500€ cada uno; Congresos de ámbito nacional e internacional; Publicación de tres artículos en revistas internacionales de alto impacto en modo de acceso abierto. En el ámbito nacional se acudirá al Congreso Anual de la Sociedad Española de Ingeniería Biomédica por parte de 4 de los integrantes del equipo en los años 2024 y 2025, el coste máximo para estos no excederá los 5.000€ anuales. Para los congresos internacionales, se acudirá al Affective Computing and Intelligent Interaction en 2024 en Glasgow por parte de 3 integrantes del equipo y el coste máximo no excederá los 9.500€. Por último, en 2025 3 integrantes acudirán al International Conference on Healthcare and Informatics en Praga con un coste máximo no superior a los 7.500€. Además, se divulgará los resultados obtenidos en revistas como Journal of Medical Internet Research Mental Health, IEEE Transactions on Affective Computing, ACM

**Comentado [8]:** [https://mental.jmir.org/?\\_hstc=236487962.dd91613c666bfbc3827f6ea6b0ab505.1715898644427.1715898644427.1715898644427.1&\\_hssc=236487962.1.1715898644428&\\_hsfp=2198788217](https://mental.jmir.org/?_hstc=236487962.dd91613c666bfbc3827f6ea6b0ab505.1715898644427.1715898644427.1715898644427.1&_hssc=236487962.1.1715898644428&_hsfp=2198788217)

**Comentado [9]:** <https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=5165369>





Transactions on Interactive Intelligent Systems (TiIS) para conseguir una mayor visibilidad de la herramienta creada.

**Comentado [10]:** <https://dl.acm.org/journal/tiis>

Con todo esto obtendremos un coste total de 96.580€ en los cuatro años de duración del proyecto.

## 6. Justificación de no procede

Hemos decidido no contratar a un estudiante de doctorado para este proyecto, y a continuación se presenta una explicación detallada:

### 1. Necesidades del proyecto

En cuanto a las necesidades del proyecto, la configuración actual de nuestro equipo ya cuenta con las habilidades y conocimientos necesarios para completar el proyecto. En el ámbito de la psiquiatría profesional, ya hemos contratado a cuatro psiquiatras. El trabajo relacionado con el aprendizaje automático y la programación puede ser realizado por otros miembros del equipo. Por lo tanto, los miembros actuales del equipo pueden manejar todas las tareas del proyecto sin la necesidad de contratar a un doctorado adicional.

### 2. Costos financieros

Durante los cuatro años de duración de nuestro proyecto, el costo de contratar a un estudiante de doctorado sería de al menos 84,000 euros. Nuestro objetivo es mantener los costos por debajo de los 100,000 euros. Considerando los salarios de otros miembros del equipo y los costos necesarios para el equipo y otros gastos del proyecto, hemos decidido no contratar a un estudiante de doctorado.

### 3. Costos de tiempo

Contratar a un estudiante de doctorado no solo implica un costo económico, sino que también requiere una inversión significativa de tiempo y recursos humanos para su formación y supervisión. Para nuestro proyecto, es crucial que todos los participantes se concentren en el trabajo de investigación para completar el proyecto de manera eficiente. Por lo tanto, hemos decidido no contratar a un estudiante de doctorado.

De esta manera, podemos asegurar que el proyecto se llevará a cabo de la manera más eficiente posible, utilizando los recursos y el personal ya disponibles en el equipo.

## 8. ANEXOS

### Anexo 1. Presupuesto

Tabla 3. Presupuesto.

Descripción	Tarea	Coste €	Periodo	Cuánta	Total coste €
Personal					
Psiquiatras	Etiquetado de datos	2000€/mes	5 meses	4	40.000,00 €
Equipamiento					
Equipo de recolección de datos	Compra de equipo para recolección de	500 x		5	2.500,00 €

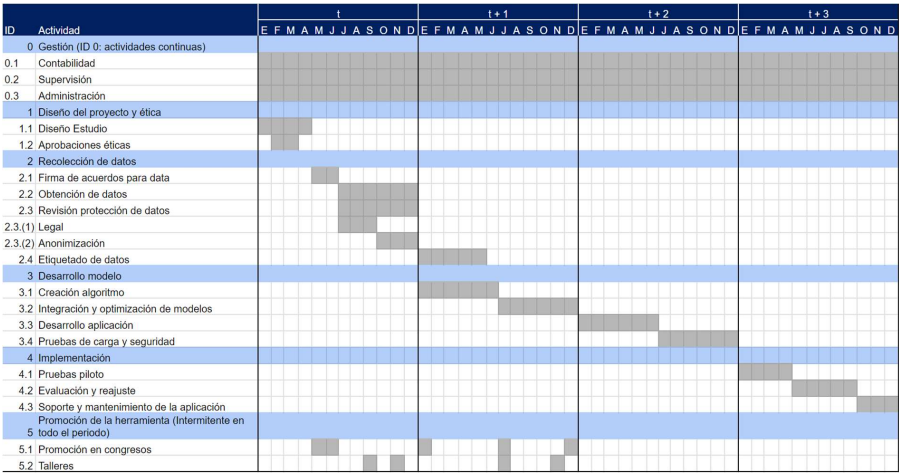


	datos multimodal				
Equipos informáticos	Compra de equipos para procesamiento de datos, creación del modelo y validación	1800 x		3	5.400,00 €
Monitores	Compra de monitores para estudio	150 x		3	450,00 €
Servidor	Compra de servidor con gran seguridad para preservar datos	200€/mes	Todo el proyecto	1	9.600,00 €
Patente	Registro y tasas	1200€ + 20€/año	Todo el proyecto	1	1.280,00 €
Congresos, divulgación y difusión					
Campaña de captación	Difusión de la idea para encontrar socios	500€ x		4	2.000,00 €
Talleres	Taller para convencer a psiquiatras de hacer uso de la herramienta	500€	Todo el proyecto	5	2.500,00 €
CASEIB 2024. Congreso Anual de la Sociedad Española de Ingeniería Biomédica	Divulgación nacional	100€	13, 14 y 15 de noviembre de 2024 y 2025	4	Máx. 5000€ (por año)
ACII 2024. Affective Computing and Intelligent Interaction	Divulgación internacional (Glasgow)	150€	15, 16, 17 y 18 Septiembre	3	Máx. 9500€
ICHI 2025. International Conference on Healthcare and Informatics	Divulgación internacional (Praga)	150€	8 y 9 de julio	3	Máx. 7500€
Journal of Medical Internet Research, IEEE Transactions on Affective Computing, ACM Transactions on Interactive Intelligent Systems (TiiS)	Publicación de artículo	1.950€		1	Máx. 5850 €
Total costes periodo					96.580,00 €



Anexo 2. Gantt

Tabla 4. Gantt



Los psiquiatras contratados intervendrán en la fase de obtención de datos con un plazo de duración de cinco meses. En el resto de etapas que componen el diagrama de Gantt participarán el resto de integrantes del grupo.

**Comentado [11]:** aqui a lo mejor sería interesante delimitar en que etapas intervienen unos y otros es decir el punto 5 del diagrama Gantt intervienen el componente IP4, IP3

Anexo 3. DAFO

Tabla 5. DAFO

Matriz de factores



## BIBLIOGRAFÍA:

- [1] Oluwashola David Adeniji, Samuel Oladele Adeyemi, and Sunday Adeola Ajagbe. An improved bagging ensemble in predicting mental disorder using hybridized random forest-artificial neural network model. *Informatica*, 46(4), 2022.
- [2] Carla Agurto, Mary Pietrowicz, Raquel Norel, Elif K Eyigoz, Emma Stanislawski, Guillermo Cecchi, and Cheryl Corcoran. Analyzing acoustic and prosodic fluctuations in free speech to predict psychosis onset in high-risk youths. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 5575–5579. IEEE, 2020.
- [3] Usman Ahmed, Gautam Srivastava, Unil Yun, and Jerry Chun-Wei Lin. Eandc: An explainable attention network based deep adaptive clustering model for mental health treatment. *Future Generation Computer Systems*, 130:106–113, 2022.
- [4] Asmaa Alayed, Manar Alrabie, Sarah Aldumaiji, Ghaida Allhyani, Sahar Siyam, and Reem Qaid. An arabic intelligent diagnosis assistant for psychologists using deep learning. *International Journal of Advanced Computer Science and Applications*, 14(6), 2023.
- [5] Shaykhah A Almaghrabi, Dominic Thewlis, Simon Thwaites, Nigel C Rogasch, Stephan Lau, Scott R Clark, and Mathias Baumert. The reproducibility of bio-acoustic features is associated with sample duration, speech task, and gender. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:167–175, 2022.



- [6] Nadim AA Atiya, Quentin JM Huys, Raymond J Dolan, and Stephen M Fleming. Explaining distortions in metacognition with an attractor network model of decision uncertainty. *PLoS Computational Biology*, 17(7):e1009201, 2021.
- [7] Casey C Bennett, Mindy K Ross, Eugene Baek, Dohyeon Kim, and Alex D Leow. Predicting clinically relevant changes in bipolar disorder outside the clinic walls based on pervasive technology interactions via smartphone typing dynamics. *Pervasive and Mobile Computing*, 83:101598, 2022.
- [8] Thomas Borger, Pablo Mosteiro, Heysem Kaya, Emil Rijcken, Albert Ali Salah, Floortje Scheepers, and Marco Spruit. Federated learning for violence incident prediction in a simulated cross-institutional psychiatric setting. *Expert Systems with Applications*, 199:116720, 2022.
- [9] Constantino Álvarez Casado, Manuel Lage Cañellas, and Miguel Bordallo López. Depression recognition using remote photoplethysmography from facial videos. *IEEE Transactions on Affective Computing*, 2023.
- [10] Chih-Yuan Chuang, Yi-Ting Lin, Chen-Chung Liu, Lue-En Lee, Hsin-Yang Chang, An-Sheng Liu, Shu-Hui Hung, and Li-Chen Fu. Multimodal assessment of schizophrenia symptom severity from linguistic, acoustic and visual cues. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023.
- [11] Jetli Chung and Jason Teo. Single classifier vs. ensemble machine learning approaches for mental health prediction. *Brain Informatics*, 10(1):1, 2023.
- [12] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [13] Longteng Duan, Wei Shao, and Linqi Song. Facial expression recognition based on data augmentation and swin-transformer. Pages 1–5, 2022.
- [14] Ricardo Flores, ML Tlachac, Avantika Shrestha, and Elke Rundensteiner. Temporal facial features for depression screening. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*, pages 488–493, 2022.
- [15] Rita Francese and Pasquale Attanasio. Emotion detection for supporting depression screening. *Multimedia Tools and Applications*, 82(9):12771–12795, 2023.
- [16] Yanrong Guo, Chenyang Zhu, Shijie Hao, and Richang Hong. Automatic depression detection via learning and fusing features from visual cues. *IEEE Transactions on Computational Social Systems*, 2022.
- [17] Lang He, Chenguang Guo, Prayag Tiwari, Rui Su, Hari Mohan Pandey, and Wei Dang. Depnet: An automated industrial intelligent system using deep learning for video-based depression analysis. *International Journal of Intelligent Systems*, 37(7):3815–3835, 2022.
- [18] Pufang Huang. A mental disorder prediction model with the ability of deep information expression using convolution neural networks technology. *Scientific Programming*, 2022:1–8, 2022.



- [19] Hyeonbin Hwang, Soyeon Kim, Parque Wei-Jin, Jiho Seo, Kyungtae Ko, and Hyeon Yeo. Vision transformer equipped with neural resizer on facial expression recognition task. Pages 2614–2618, 2023.
- [20] Ngumimi Karen Iyortsuun, Soo-Hyung Kim, Hyung-Jeong Yang, Seung-Won Kim, and Min Jhon. Additive cross-modal attention network (ACMA) for depression detection based on audio and textual features. *IEEE Access*, 2024.
- [21] Attila Zoltán Jenei and Gábor Kiss. Severity estimation of depression using convolutional neural network. *Periodica Polytechnica Electrical Engineering and Computer Science*, 65(3):227–234, 2021.
- [22] Shehroz S Khan, Pratik K Mishra, Nizwa Javed, Bing Ye, Kristine Newman, Alex Mihailidis, and Andrea Iaboni. Unsupervised deep learning to detect agitation from videos in people with dementia. *IEEE Access*, 10:10349–10358, 2022.
- [23] Rezaul K Khandker, Md Rakibul Islam Prince, Farid Chekani, Paul Richard Dexter, Malaz A Boustani, and Zina Ben Miled. Digital-reported outcome from medical notes of schizophrenia and bipolar patients using hierarchical BERT. *Information*, 14(9):471, 2023.
- [24] Lin Sze Khoo, Mei Kuan Lim, Chun Yong Chong, and Roisin McNaney. Machine learning for multimodal mental health detection: A systematic review of passive sensing approaches. *Sensors*, 24(2):348, 2024.
- [25] Rohan Kumar Gupta and Rohit Sinha. An investigation on the audio-video data based estimation of emotion regulation difficulties and their association with mental disorders. *IEEE Access*, 2023.
- [26] S Lalitha, Deepa Gupta, Mohammed Zakariah, and Yousef Ajami Alotaibi. Mental illness disorder diagnosis using emotion variation detection from continuous english speech. *Computers, Materials & Continua*, 69(3), 2021.
- [27] Bishal Lamichhane, Nidal Moukaddam, Ankit B Patel, and Ashutosh Sabharwal. Econet: Estimating everyday conversational network from free-living audio for mental health applications. *IEEE Pervasive Computing*, 21(2):32–40, 2022.
- [28] Keke Li and Weifang Yu. A mental health assessment model of college students using intelligent technology. *Wireless Communications and Mobile Computing*, 2021:1–10, 2021.
- [29] Shuang Li and Yu Liu. Deep learning-based mental health model on primary and secondary school students' quality cultivation. *Computational Intelligence and Neuroscience*, 2022, 2022.
- [30] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10012–10022, 2021.
- [31] Aritra Marik, Soumitri Chattopadhyay, and Pawan Kumar Singh. A hybrid deep feature selection framework for emotion recognition from human speeches. *Multimedia Tools and Applications*, 82(8):11461–11487, 2023.
- [32] Ridha Mezzi, Aymen Yahyaoui, Mohamed Wassim Krir, Wadii Boulila, and Anis Koubaa. Mental health intent recognition for arabic-speaking patients using the mini international neuropsychiatric interview (MINI) and BERT model. *Sensors*, 22(3):846, 2022.



- [33] Kirill Milintsevich, Kairit Sirts, and Gal Dias. Towards automatic text-based estimation of depression through symptom prediction. *Brain Informatics*, 10(1):4, 2023.
- [34] Alice Othmani and Muhammad Muzammel. An ambient intelligence-based approach for longitudinal monitoring of verbal and vocal depression symptoms. In *International Workshop on Predictive Intelligence In Medicine*, pages 206–217. Springer, 2022.
- [35] Mostafa Parchami and Sarwan Ali. Self-attention convolutional neural network for emotion recognition in speech. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7144–7148. IEEE, 2020.
- [36] Furkan Rahman, Mohammad Mahmudul Alam, and Asif Mahbub. Performance analysis of random forest model for mental health diagnosis from wearable device's data. *Informatics in Medicine Unlocked*, 29:101067, 2022.
- [37] William Ratti, Jose David Herrera, Steven Bedrick, Emily M Bender, Andrew G Oyler, and Aaron K Stevens. An ensemble learning approach to clinical depression detection using semantic fusion. *Pervasive and Mobile Computing*, 98:101775, 2023.
- [38] Ricardo Reyes-Ortiz, Yair Wisniak, Eugenio Martinez-Villalpando, Ricardo Campisano, and Jose A Rodriguez-Resendiz. Machine learning and deep learning for the identification of suicide and self-harm thoughts in social media posts: Systematic review. *JMIR Formative Research*, 7:e39872, 2023.
- [39] Xiangyu Shen, Huiying Ma, Hang Wang, Shichang Sun, and Bo Lei. Multi-modal cross attention network for early diagnosis of Alzheimer's disease. *IEEE Journal of Biomedical and Health Informatics*, 27(5):2625–2635, 2023.
- [40] Muhammad Iqbal Sodiq, Muhammad Adnan, and Aatif Jamal. On a method to recognize real-life facial expressions in videos. *IEEE Access*, 11:57208–57222, 2023.
- [41] Bo Tang, Qibin Sun, Yao Zhang, and Pei Liu. Speech emotion recognition using attention-based LSTM. *IEEE Access*, 6:71514–71522, 2018.
- [42] Simon Weiss, Harald Pirker, and Thomas Deselaers. Neural network-based visual emotion analysis for monitoring mental health. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 680–690, 2023.
- [43] Wing-Kin Wong and Ivy Lee. DepKB: Knowledge-driven automatic depression diagnosis based on medical context. *Healthcare Analytics*, 4:100116, 2023.
- [44] Kunlin Wu, Yunzhi Bai, Han Zhang, Zongheng Liu, Qiang Liu, Jionglong Su, Bin Tang, and Ling Shao. Visual-audio self-supervised learning framework for affective behavior analysis. *IEEE Transactions on Affective Computing*, 2023.
- [45] Yao Wu, Feng Zheng, and Huachun Tan. PS-HCRF: A hybrid model for emotion recognition from physiological signals. *IEEE Access*, 10:97278–97290, 2022.
- [46] Yichen Wu, Xiaoying Jin, Xiaowen Wu, Li Xue, and Hui Wu. Intelligent service system for elderly patients with mental illness based on deep learning. *Frontiers in Psychology*, 13, 2022.
- [47] Jeong Yeol Yun, Mahesan Niranjan, and Bernhard Pfahringer. Empirical study on deep learning approaches for automatic recognition of mental illness on social media. *International Journal of Environmental Research and Public Health*, 17(1):1–23, 2020.



- [48] Yi-Ching Wang, Yi-Wei Chen, Yi-Ping Hung, and Mei-Chen Yeh. Anomalous pupil behavior detection system for schizophrenia. *IEEE Transactions on Affective Computing*, 2023.
- [49] Hongda Xiao and Deyi Xue. Online mental disorder analysis system based on deep learning. *IEEE Access*, 9:155851–155857, 2021.
- [50] Yuan Yuan, Chaowei Sun, and Guanghui Wang. Automatic monitoring of mental health status in the digital twin framework. *IEEE Access*, 11:54114–54122, 2023.
- [51] Jiannan Zhao, Peng Ren, Qiguang Miao, Ruochi Zhang, Xiangyu Zhao, Jianhong Qiu, and Xiang Zhang. A novel EEG-based model for depression recognition using high gamma activity and basic wavelet entropy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:2546–2556, 2023.
- [52] Wenzhi Zhong, Tongsuo Liu, and Zhengwu Zhang. A novel vision transformer for intelligent mental disease recognition. *IEEE Transactions on Computational Social Systems*, 2023.